

Dokumentazioa eta eduki digitalen kudeaketa

Documentación y gestión de los contenidos digitales

Teresa Agirreazaldegí Berriozabal
(koord./coord.)

eman la zabalazazu



Universidad del País Vasco Euskal Herriko
Unibertsitatea

ARGITALPEN
ZERBITZUA
SERVICIO EDITORIAL



KAZETARITZA SAILA
DEPARTAMENTO DE PERIODISMO
DEPARTMENT OF JOURNALISM

**Dokumentazioa eta
eduki digitalen kudeaketa**

*Documentación y gestión
de los contenidos digitales*

Dokumentazioa eta eduki digitalen kudeaketa

Documentación y gestión de los contenidos digitales

Teresa Agirreazaldegí Berriozabal
(koord./coord.)

eman la zabal azuz:



Universidad del País Vasco Euskal Herriko Unibertsitatea

ARGITALPEN
ZERBITZUA
SERVICIO EDITORIAL

© Servicio Editorial de la Universidad del País Vasco
Euskal Herriko Unibertsitateko Argitalpen Zerbitzua

ISBN: 978-84-9860-486-3

Depósito legal/Lege gordailua: BI - 3.475-2010

Fotocomposición/Fotokonposizioa: Ipar, S. Coop.
Zurbaran, 2-4 - 48007 Bilbao

Impresión/Inprimatzea: Itxaropena, S.A.
Araba Kalea, 45 - 20800 Zarautz (Gipuzkoa)

Aurkibidea/Índice

Aurkezpena	9
Presentación	13

1. atala

Informazio eta dokumentazio sistemen kudeaketa digitala

1.^a parte

Gestión digital de los sistemas de información y documentación

I. Los contenidos digitales: la revolución soterrada...	
<i>Jokin Olaizola Azkarate</i>	19
II. Euskal Herriko Dokumentazio Zerbitzuak: begirada bat.	
<i>Teresa Agirreazaldegí, Idoia Camacho, Carmen Peñafiel eta Milagros Ronco</i> ..	51
III. Bizkaiko Foru Liburutegia eta aro digitala.	
<i>Ana Ares Martínez</i>	79
IV. Euskomedia informazio sistema.	
<i>Iñaki Arrieta Baro</i>	85

2. atala

Enpresako dokumentazio zerbitzuak aro digitalean

2.^a parte

Servicios de documentación de empresa en la era digital

V. Fagor sozietate kooperatiboaren Dokumentazio Zerbitzua: 20 urte erne eta tinko.	
<i>Juan Carlos Guridi</i>	95

VI. Retos y respuestas de la gestión de la información en la Corporación Tecnológica Tecnalía. <i>Edorta Aranguena, Pilar Ruiz e Izaskun Urrutia</i>	101
VII. La evolución del concepto de gestión documental en las organizaciones potenciada por la digitalización de los documentos y las herramientas de colaboración. El caso práctico de Bahía de Bizkaia Gas. <i>Ainhoa Arpide Etxano</i>	121
3. atala	
Internet sarea artxibatzen	
3.ª parte	
<i>Archivando la red Internet</i>	
VIII. Euskal hauteskundeetako webak Ondarenet-en. <i>Francisca Pulgar Vernalte eta Sonia Marcos Maciá</i>	135
IX. Las webs de las elecciones vascas en Ondarenet. <i>Francisca Pulgar Vernalte y Sonia Marcos Maciá</i>	145
X. La experiencia catalana archivando la red: el repositorio Padicat (Patrimonio Digital de Cataluña) de la Biblioteca de Catalunya. <i>Ciro Llueca y Daniel Cócera</i>	155
4. atala	
Informazio digitalaren unibertsoa	
4.ª parte	
<i>El universo de la información digital</i>	
XI. Semantikan oinarritutako bilaketak: Kyoto proiektua. <i>Iñaki Alegria eta German Rigau (IXA taldea)</i>	169
XII. Informazio kudeaketa, sindikazioaren eta tresna folksonomikoen bidez. <i>Gorka Julio</i>	183
XIII. Una visión de la gestión de contenidos web y el rol del gestor de información digital. <i>Fernando Fernández de Aránguiz</i>	187
XIV. El acceso al conocimiento global: la contribución de Google. <i>Luis Collado</i>	201
Apéndice: Los contenidos digitales: la revolución soterrada... (esquemas)	211

Aurkezpena

Euskal Herriko Unibertsitatean Kazetaritzako, Ikus-entzunezko Komunikazioko eta Publizitate eta Harreman Publikoko lizentziaturak ezarri zirenetik, Dokumentazioa irakasgai modura dago bertako curriculumean. Horri esker, milaka ikasle trebatu dira informazio iturrien sarreran eta dokumentu eta baliabide profesional eta zientifikoan erabileran. Eremu hori bultzatzeko asmoz, eta Unibertsitatea eztabaidarako eta ezagutzak trukatzeko gunea dela sinetsita, 2009ko apirilaren 29an eta 30ean Dokumentazio eta Eduki Digitalen Kudeaketari buruzko II. Jardunaldiak antolatu ziren Kazetaritza Sailean. Dokumentazioko ikertzaileak eta profesionalak foro bakarrean biltzea zen helburua, arlo horrek azkenaldian zer bilakaera izan duen eta Euskal Herrian zer-nola erabili den jakiteko asmoz. Txostengileen kalitateak eta jardunaldiek entzuleengan piztu zuten interesak —konferentzia-gela ikaslez, ikertzailez, irakaslez eta profesionalez gainezka egon zen— bertan aurkeztutako txosten eta material guztiak biltzera eta liburu batean eskaintzera bultzatu gaituzte, eta, hala, dokumentua den heinean, Leioan azaldutako diskurtsoaren jarraipen informatiboa izango da.

Hamalau txostenen sarrera errazteko asmoz, lau atal handitan egitaratu dugu liburua: I) Informazio- eta dokumentazio-sistemen kudeaketa digitala; II) Enpresako dokumentazio zerbitzuak aro digitalean; III) Internet sarea artxibatzen; eta IV) Informazio digitalaren unibertsoa.

Informazio- eta dokumentazio-sistemen kudeaketa digitala izenburu orokorren barruan lau lan bildu dira. Lehenengoan, Jokin Olaizola Azkarate Eusko Jaurlaritzako webguneen kudeatzaileak mundu digitalaren aurkezpen oso bat egiten du. Mundu hori erraza, automatizatua eta ohiko menpekotasunetik aske dagoela ematen badu ere, gutxitan azaltzen den konplexutasuna du; izan ere, hor potentzialki dena elkarrekin lotuta dago, eta denak bata bestearen menpe daude. Lotura horiek osotasunean ulertu behar direla dio egileak, zerikusi eta ondorio guztien jakitun

izan ahal izateko, bai gizabanako-herritar-kontsumitzaile ikuspegitik, bai igorlearen edo edukien sortzailearen ikuspuntutik, edo, harantzago joanda, dokumentalistaren ikuspegitik.

Lanaren bigarren kapituluan, Teresa Agirreazaldegi, Idoia Camacho, Carmen Peñafiel eta Milagros Ronco Euskal Herriko Unibertsitateko irakasleen *Dokumentazio zerbitzuak Euskal Herrian: begirada bat* txostena jasota dago. Bibliotekak eta dokumentazio zerbitzuak, diren motatakoak direla (orokorrak, espezializatuak, publikoak, pribatuak, barrukoak, eta abar), informazio- eta ezagutza- gizarterako oinarritzko azpiegiturak dira. Zerbitzuak Euskal Herrian nola ezarri diren aztertzen du ikerketak, haien ezaugarriak eta baliabide material eta pertsonalak aztertuta, baita ondarea, lana, eskaintzen diren zerbitzuak eta erabiltzaileak ere.

Hirugarren kapituluan, *Bizkaiko Foru Liburutegia aro digitalean*, Ana Ares Martínez funts bibliografikoen eta Foru Liburutegia sustatzeko buruak azaltzen du ehun urte baino gehiago duen liburutegi horrek digitalizazio erronkari nola egiten dion aurre. Halaber, gizartearen eskakizun berriei erantzuteko bibliotekako egituran, kudeaketan eta antolakuntzan izaten ari diren aldaketak azaltzen ditu.

Euskomedia Fundazioko proiektuen arduradun Iñaki Arrieta Baroren *Euskomedia informazio sistema* txostena dugu laugarren kapituluan. Fundazio hori sarean euskal kulturaren eta zientziaren erreferente bat izatea nahi dute, eta, horretarako, zenbait dokumentazio bilgune sortu dituzte: euskomedia.org — Auñamendi entziklopedia, *Euskonews* aldizkaria eta dokumentazioa eta argazkiak jasotzen dituzten aldizkari eta funtsak biltzen ditu—, 3digitala.com irudi bankua eta OAI Hedatuz gordailua, 2008an sortua, Euskal Herrian egiten den ikerketaren hedapena bultzatzeko.

Liburuaren bigarren atalak, *Enpresako dokumentazio zerbitzuak aro digitalean*, hiru kapitulu biltzen ditu; lehenengoan, *Fagor Koop. E.-ren Dokumentazio Zerbitzua: Hogeitau urte erne eta tinko*, Fagor kooperatibako dokumentazio zentroko arduradun Juan Carlos Guridik hogeitau urte baino gehiago duen zentro horren nondik norakoak azaltzen ditu, ematen dituen zerbitzuak, erabiltzaileen ezaugarriak eta erabiltzen dituzten lanabesak azalduz, besteak beste.

Seigarren txostenaren izenburua *Tecnalia teknologia korporazioko informazio kudeaketaren erronkak eta erantzunak* da. Tecnalia Espainiako estatuko teknologia enpresa handiena da, zortzi euskal teknologia zentroz osatua, bakoitza jatorri eta sektore arlo batekoa. Edorta Aranguenak, Pilar Ruízek eta Izaskun Urrutiak —Azi, Robotiker eta Labein enpresetako dokumentazio zerbitzuetako arduradunak dira, hurrenez hurren— erakunde horien beharrei aurre egiteko landu dituzten konponbideak azaltzen dituzte: dokumentu-kudeaketarako sistema baten diseinua, neurrira egina; kalitateko informazio baliabideen hautaketa eta optimizazioa, eta informazio-konpetentzietan trebatzeko ekintzak. Teknologia zentroetako dokumentazio arloko profesionalek aurre egin behar diete informazioaren inguruko erronka berriei, eta berez ezagutzaren langileak diren erabiltzaileen behar konplikatuei erantzun behar diete.

Atal horri amaiera emateko, Ainhoa Arpide Etxanok, Bahía de Bizkaia Gas SL konpainiako erosketen eta dokumentazio-kudeaketaren saileko arduradunak, agi-

ritegien digitalizazioak ematen dituen aukerak aurkezten ditu; izan ere, digitalizazioak aldaketak ahalbidetu ditu bai erabiltzaileen pentsamoldean, bai erakundeek agiriak kudeatzeko antolakuntzan. Agiritegia jada ez da dokumentuen kudeaketaren azken gunea, baizik eta denbora errealean parte hartzen du erakundearen jardueraren barruan.

Internet sarea artxibatzen da liburuaren hirugarren atala, eta Interneteko edukiak gordetzeko eta babesteko Eusko Jaurlaritzako Kultura Sailak eta Kataluniako Liburutegiak martxan jarri dituzten egitasmoak azaltzen dira: jatorriz digitala den informaziorako sarrera ziurtatzea da helburua. *Euskal hauteskundeetako webguneak Ondarenet-en* txostenean, Francisca Pulgar Vernalte Eusko Jaurlaritzako Liburutegi Zerbitzuko arduradunak eta Sonia Marcos Maciá ODEI enpresako teknikari dokumentalistak 2009ko Eusko Legebiltzarrerako hauteskundeetan parte hartu zuten alderdi politikoen webguneak berreskuratzeko, zaintzeko eta hedatzeko helburua duen proiektua azaltzen dute. Bildutako datuak edonork kontsultatu ahal izango ditu Ondarenet-en. Euskal Autonomia Erkidegoko ondare digitalaren fitxategia dugu Ondarenet, formatu digitalean sortutako baliabideak berreskuratzeko, zaintzeko eta hedatzeko helburuarekin Eusko Jaurlaritzaren Kultura Sailak martxan jarritako egitasmoa.

Hamargarren kapituluan, Ciro Lluccak eta Daniel Cócera Kataluniako ondare digitala jasotzen duen Padicat biltegia aurkezten dute. Kataluniako Bibliotekak sortu zuen egitasmoa, 2005ean, eta katalanez kaleratzen diren web orriak gordetzea da horren helburua, baita CAT domeinuan beste hizkuntzetan plazaratutakoa, edota beste domeinu batean azaltzen dena baina Kataluniarekin lotura duena. Lau urte igaro ondoren, Kataluniako Bibliotekak 4.000 baliabide digital baino gehiago eskaintzen ditu modu irekian Padicaten, sistematikoki jasoak, 400 instituziorekin egindako hitzarmenei esker eta Kataluniako gizartea ulertzeko baliagarriak diren gertaeren bilketa sistematikoari esker; hauteskundeak, kasu.

Liburuaren azken atalak —*Informazio digitalaren unibertsoa*— lau kapitulu jasotzen ditu. Dokumentazioaren kudeaketa digitalean datu-baseen informazioa atzitzeko orduan, hizkuntza ingeniariak hainbat aukera ematen ditu: ele askotako atzitzeak, multzokatze semantikoak, antzekotasunean oinarritutako atzitzeak, galdera-erantzun sistemak, informazioaren ondorioztatzea... *Semantikan oinarritutako bilaketak: Kyoto proiektua* kapituluan, Iñaki Alegria eta German Rigau Euskal Herriko Unibertsitateko irakasle eta IXA Taldekoek semantikaren inguruko aukeretan sakontzen dute, Europako beste unibertsitate eta enpresa batzuekin lantzen ari diren Kyoto proiektuaren ikerketa-ildoak azalduz.

Hamabigarren kapituluan, Elurnet enpresako ingeniari tekniko Gorka Julioren txostena jasota dago: *Informazioaren kudeaketa, sindikazio eta tresna folksonomikoen bidez*. Sindikatzeak informazio iturri ugarietara harpidetzeko aukera ematen digu, betiere informatuta egon gaitezen: gure esparruko berriak ezagutu, gure enpresa, talde, elkarteri buruz zer esaten den jakin, alertak jaso, eta abar. Baina sindikatzeak, informazioa jasotzeko aukera emateaz gain, informazio hori trukatzeko aukera ere ematen du, talde publiko zein pribatuekin, era parte-hartzaile batean; eta gure edukiak gure esparrutik kanpo zabaltzeko aukera izango dugu.

Fernando Fernández de Aranguizek, ODEI enpresako proiektu arduradunak, webguneetako edukiak kudeatzeko ikuskera bat eta informazio digitalaren kudeatzailearen eginkizuna aurkeztu du; hain zuzen, eduki digitalak nolakoak diren, webguneetako edukien kudeatzaileak zer diren (WCM) eta informazio digitalaren kudeatzailearen eginkizuna eta ezagutzak zein diren azaltzen du.

Liburua Luis Colladok aurkeztutako txostenarekin bukatzen da. Egilea Espainiako eta Portugalgo Google Book Search, Scholar eta News Archive Searcheko arduraduna da, eta Google informazioa eta edukiak nola gehitzen ari den agertzen du, eta nola horrela jakintzarako sarrera unibertsal bihurtzen ari den eta sarean dauden edukiak bilatzen, aurkitzen, analizatzen eta sailkatzen laguntzen ari den.

Presentación

La Documentación como materia forma parte del curriculum de las Licenciaturas de Periodismo, Comunicación Audiovisual y Publicidad y Relaciones Públicas desde la implantación de estos estudios en la Universidad del País Vasco; y a través de ella miles de alumnos se han formado en el acceso a las fuentes de información y en la utilización y organización de documentos y recursos profesionales y científicos. Para impulsar esta área, y en el convencimiento de que la Universidad debe ser un espacio de debate y de intercambio de conocimiento, desde el Departamento de Periodismo se han organizado los días 29 y 30 de abril de 2009 las II Jornadas sobre Documentación y Gestión de los Contenidos Digitales. El objetivo de las jornadas era reunir en un foro a investigadores y profesionales de la Documentación, para conocer de primera mano los últimos desarrollos de esta disciplina y su implantación en el País Vasco. La calidad de los ponentes y el interés despertado en la audiencia, con una sala abarrotada de estudiantes pero también de investigadores, docentes y profesionales, nos han empujado a recopilar todas las ponencias y materiales allí presentados y ofrecerlos en este formato libro que, en tanto que documento, tiene la virtud de proporcionar continuidad informativa al discurso allí expuesto. Para facilitar el acceso a las catorce aportaciones que contiene, hemos estructurado el libro en cuatro grandes partes: I) Gestión digital de los sistemas de información y documentación; II) Servicios de documentación de empresa en la era digital; III) Archivando la red Internet; y IV) El universo de la información digital.

Bajo el título genérico de «Gestión digital de los sistemas de información y documentación», se recogen cuatro trabajos. En el primero de ellos, «Los contenidos digitales: la revolución soterrada», Jokin Olaizola Azkarate, gestor web en el Gobierno Vasco, hace una presentación exhaustiva del mundo digital, un mundo que a simple vista puede parecer fácil, automatizado y libre de las servidumbres

tradicionales, pero que encierra un nivel de complejidad raramente explicitado; y donde, potencialmente, todo se relaciona con todo, y todos dependen de todos. El autor señala que es preciso comprender estas vinculaciones en su integridad para llegar a ser conscientes de todas sus implicaciones y consecuencias, tanto desde la perspectiva del individuo-ciudadano-consumidor, como desde la óptica del emisor o creador de contenidos o, yendo más allá, del documentalista.

En el segundo capítulo de la obra Teresa Agirreazaldegí, Idoia Camacho, Carmen Peñafiel y Milagros Ronco, profesoras de la Universidad del País Vasco, ofrecen una mirada a los servicios de documentación del País Vasco («Dokumentazio zerbitzuak Euskal Herrian: begirada bat»). En sus diferentes manifestaciones (generalistas, especializados, públicos, privados, internos, etc.), las bibliotecas y servicios de documentación son infraestructuras y servicios claves en la sociedad de la información y del conocimiento; la investigación examina la implantación de estos servicios en el País Vasco, analizando sus características, recursos materiales y personales, y proporcionando datos sobre fondos, servicios ofertados y usuarios, entre otros.

En el tercer capítulo («Bizkaiko Foru Liburutegia eta aro digitala») Ana Ares Martínez, jefa de la sección de fondos bibliográficos y promoción de la Biblioteca Foral de Bizkaia, expone cómo esta biblioteca de más de cien años de historia hace frente al reto digital, y señala los cambios que se están produciendo en su estructura, gestión y organización para adaptarse a las nuevas exigencias sociales.

Iñaki Arrieta Baro, responsable de proyectos de la fundación Euskomedia, presenta en el capítulo cuarto el sistema de información de Euskomedia. Fundación que pretende ser en la red un referente de la ciencia y la cultura vascas, para lo que ha creado varias plataformas de documentación: euskomedia.org que engloba la *Enciclopedia Auñamendi*, la revista *Euskonews*, y revistas y fondos en los que se recogen fotos y documentación; 3digitala.com, que es un banco de imágenes; y el repositorio OAI Hedatuz —para impulsar la expansión de las investigaciones hechas en el País Vasco—.

La segunda parte de la obra tiene como temática los «Servicios de documentación de empresa en la era digital», y consta de tres capítulos. En el primero de ellos, Juan Carlos Guridi, responsable del centro de documentación de Fagor sociedad cooperativa, presenta este centro de documentación que ha cumplido veinte años, explicando qué servicios presta, qué usuarios lo utilizan y qué herramientas manejan.

La siguiente ponencia proviene de la corporación tecnológica Tecnalía, la mayor empresa tecnológica del estado, integrada por ocho centros tecnológicos vascos de distinto origen y orientación sectorial. Bajo el título «Retos y respuestas de la gestión de la información en la Corporación Tecnológica Tecnalía», Edorta Aranguena, Pilar Ruíz e Izaskun Urrutia describen las soluciones que los centros de documentación de las empresas Azti, Robotiker y Labein han desarrollado en relación con necesidades clásicas y comunes a las entidades a las que dan servicio: el diseño a medida de un sistema de gestión documental, la selección y optimización de recursos de información de calidad y las acciones de formación en com-

petencias informacionales. Los profesionales de la documentación de los centros tecnológicos tienen que afrontar los retos que implican el nuevo escenario informacional, y responder a las complejas necesidades de unos usuarios que son —por definición— trabajadores del conocimiento.

Concluye esta segunda parte referida a los servicios de documentación de las empresas, con el capítulo «La evolución del concepto de gestión documental en las organizaciones potenciada por la digitalización de los documentos y las herramientas de colaboración. El caso práctico de Bahía de Bizkaia Gas», en el que Ainhoa Arpide Etxano, responsable del departamento de compras y gestión documental de BBG, presenta las oportunidades que ofrece la digitalización de los archivos, que ha posibilitado un cambio de mentalidad en los usuarios así como un cambio de planteamiento en la gestión de la documentación dentro de las organizaciones: el archivo no es el punto final de la gestión documental, sino que convive con la actividad de la organización en tiempo real.

En la tercera parte del libro y bajo el título «Archivando la red Internet», se presentan las iniciativas puestas en marcha por la Consejería de Cultura del Gobierno Vasco y la Biblioteca de Catalunya, destinadas a preservar los contenidos publicados en Internet, con el objetivo de asegurar el acceso a esa información. En «Webs de las elecciones vascas en Ondarenet», Francisca Pulgar Vernalte, responsable del servicio de bibliotecas del Gobierno Vasco, y Sonia Marcos Maciá, técnica documentalista de la empresa ODEI, presentan el proyecto de recuperación, preservación y difusión de las páginas webs de los partidos políticos que han participado en las elecciones al Parlamento Vasco en 2009. Estas páginas pueden ser consultadas desde Ondarenet, archivo electrónico del patrimonio digital vasco creado por el Departamento de Cultura del Gobierno Vasco para la recuperación, conservación y difusión de los recursos nacidos en formato digital.

En el capítulo «La experiencia catalana archivando la red: el repositorio Padicat (Patrimonio Digital de Cataluña) de la Biblioteca de Catalunya», Ciro Lluca y Daniel Cócera exponen el proyecto Padicat, puesto en marcha en 2005, por la Biblioteca de Catalunya, con el objetivo de preservar las páginas web relacionadas temáticamente con Cataluña, publicadas en lengua catalana y en otras lenguas, bajo el dominio .CAT y otros dominios geográficos o temáticos. Tras cuatro años de funcionamiento la Biblioteca de Catalunya ofrece de forma abierta más de 4.000 recursos digitales, capturados sistemáticamente gracias a acuerdos firmados con 400 instituciones, así como por la recopilación sistemática de diversos acontecimientos clave para comprender la sociedad catalana.

La última parte del libro consta de cuatro capítulos agrupados bajo el título «El universo de la información digital». Las técnicas de ingeniería lingüística aplicadas a la gestión digital de la documentación abren nuevas opciones de acceso a la información contenida en las bases de datos (como el acceso multilingüe, la agrupación semántica, el acceso basado en similitud, los sistemas de pregunta-respuesta, la inferencia de información...). Los profesores de la Universidad del País Vasco Iñaki Alegria y German Rigau, pertenecientes al Grupo Ixa, profundizan en las posibilidades basadas en la semántica, exponiendo las líneas de investigación

que están desarrollando, junto con otras universidades y empresas, en el proyecto europeo Kyoto («Semantikan oinarritutako bilaketak: Kyoto proiektua»).

En el capítulo duodécimo, Gorka Julio, ingeniero técnico de la empresa Elurnet, presenta la gestión de la información a través de la sindicación y de las herramientas folksonómicas («Informazioaren kudeaketa sindikazio eta tresna folksonomikoen bidez»). La sindicación nos da opción de suscribirnos a muchas y distintas fuentes de información, para que siempre permanezcamos informados, ya que cada vez es más importante hacer una correcta gestión de la información, conocer las noticias de nuestro entorno, nuestra empresa o grupo, conocer lo que se dice sobre la sociedad, recibir avisos... Pero, además, de todo ello, podremos intercambiar información, tanto con grupos públicos como privados, y propagar contenidos fuera de nuestros ámbitos.

Fernando Fernández de Arangiz, responsable de proyecto en la empresa ODEI, presenta «Una visión de la gestión de contenidos web y el rol del gestor de información digital», donde este autor analiza cómo son los contenidos digitales, qué son los Gestores de Contenidos Web (WCM) y, también, cuáles son las funciones y conocimientos del gestor de información digital.

La obra termina con el capítulo «El acceso al conocimiento global: la contribución de Google», de Luis Collado, responsable de Google Book Search, Scholar y News Archive Search para España y Portugal, que expone cómo Google incorpora información y contenidos que están universalizando el acceso al conocimiento, y de qué manera está ayudando a encontrar, descubrir, analizar y categorizar los contenidos accesibles en la Red.

1. atala

Informazio eta dokumentazio
sistemen kudeaketa digitala

1.^a parte

*Gestión digital de los sistemas
de información y documentación*

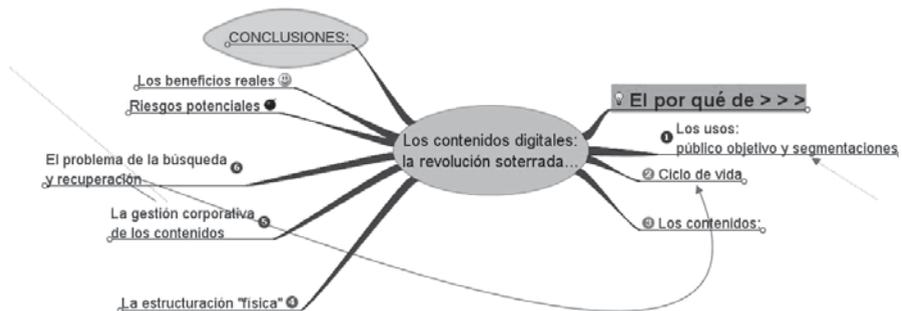
I

Los contenidos digitales: la revolución soterrada...

Jokin Olaizola Azkarate

Gestor Web en el Gobierno Vasco
Web kudeatzailea Eusko Jaurlaritzan

Resumen: Contra lo que puede parecer un mundo fácil, automatizado y libre de las servidumbres tradicionales, los contenidos digitales encierran un nivel de complejidad raramente explicitado. Así, se trata de un nuevo universo donde potencialmente todo se relaciona con todo, y todos dependen de todos. Vinculaciones que es preciso comprender en su integridad, para llegar a ser conscientes de todas sus implicaciones y consecuencias, tanto desde la perspectiva del individuo-ciudadano-consumidor, como desde la óptica del emisor o creador de contenidos o, yendo más allá, del documentalista. Empezando por el acceso a unos contenidos en crecimiento exponencial e indiscriminado, los problemas asociados a su perdurabilidad, los poderosísimos algoritmos de análisis y las actualmente «cuasi» ilimitadas capacidades de almacenamiento y proceso, la influencia sobre las personas, las implicaciones legales... etc.



Los contenidos digitales han dejado de ser «islas» de información para convertirse en un sistema poliédrico donde las diferentes facetas ejercen su influencia y reclaman sus especificaciones, y que es necesario abordar y gestionar de una manera integral. Y aunque, bien mirado, la mayoría de dichas especificaciones ya tienen un precedente «tradicional», ni por sus volúmenes, ni por su inmediatez, ni por sus efectos multiplicadores, ni por su globalización, pueden ser comparables con experiencias previas. En este sentido, se impone una visión multidisciplinar que, además de cumplir con los diferentes requisitos exigibles, sepa también dar respuesta en cada momento a las necesidades y demandas de las personas y entidades.

Nota: Los gráficos correspondientes a este artículo se pueden consultar a mayor tamaño en el apéndice incorporado en la página 211.

Laburpena: Mundua erraza, automatizatua eta ohiko morrontzetatik libre dela iruditu arren, eduki digitalek gutxitan aipatzen den zailtasun maila dute. Unibertso guztiz berria da, gauza guztiek elkarrekiko harremana izan dezakete eta, aldi berean, denak beste denen menpean egon daitezke. Harreman horiei dago-kienez eta, bereziki, beraien ondorio eta inplikazio guztiak ezagutzeko, ezinbestekoa da lotura horiek sakon eta erabat ulertzea, bai kontsumitzaile-herritar-gizabakoaren ikuspuntutik, bai edukiak igortzen edo sortzen dituenaren ikuspegitik, edo, urrunago joanda, dokumentalistarenetik. Hasteko, edukiak hazkunde esponentzial eta bereizi gabekoa dute, edukiok iraunkortasun arazoak dituzte, analisi-rako algoritmo indartsuak daude, gorde eta prozesatzeko ahalmenak ez du ia mugarik gaur egun, gizakiei eragiten diete, lege inplikazioak dituzte, eta abar.

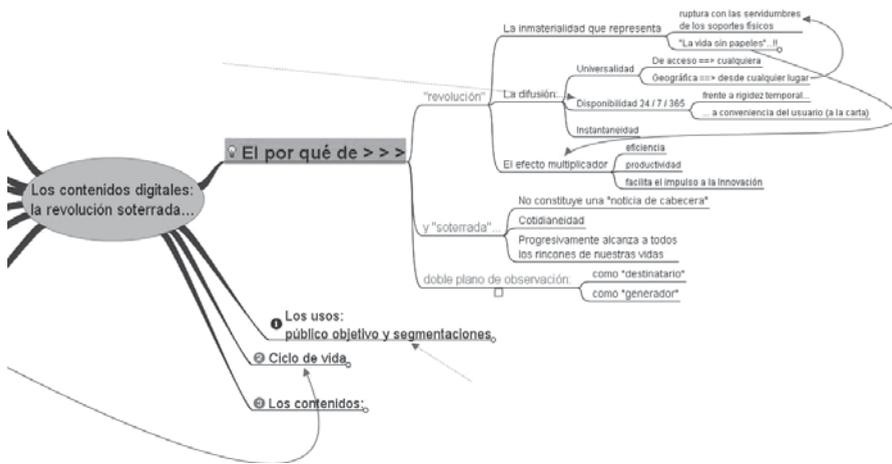
Eduki digitalek argibide/informazio uharteak izateari utzi diote eta sistema poliedriko bat osatu dute. Sistema berri horretan, alde guztiek beren eragina dute, eta beren zehaztapena eskatzen dutenez, osotasunez heldu eta osotasunez kudeatu behar dira guztiak. Eta nahiz eta, ondo aztertuta, zehaztapen gehienek baduten «tradiziozko» aurrekariren bat, ezin dira aurreko esperientziekin alderatu, ez kopuruan, ez berehalakotasunean, ez eragin biderkatzaileetan, ez globalizazioan. Ildo horretan, ezinbestekoa da diziplina anitzeko ikuspegia izatea, eskatutako betekizun guztiak gordetzeaz gain une bakoitzean pertsona eta erakundeen behar eta eskakizunei erantzun ahal izateko.

A QUÉ NOS ENFRENTAMOS...: SUS SIGNIFICADOS E IMPLICACIONES

En la actualidad, los contenidos digitales han adquirido ya una importancia y unas características tales, que se hace necesario un receso que nos permita abordar la cuestión en todas sus dimensiones. Y fruto del mismo surge la reflexión de hallarnos ante una «revolución soterrada». Una revolución, porque afecta a la naturaleza propia de los nuevos contenidos, tanto por la inmaterialidad que representan, como por las extraordinarias posibilidades divulgativas que encierran, o por el formidable efecto potenciador de la creatividad, la eficiencia, la productividad o la innovación, todas ellas clave en el futuro desarrollo de nuestras sociedades.

En segundo lugar, porque esta revolución está ya sucediendo ante nuestros ojos sin que en muchas ocasiones seamos conscientes de tal circunstancia, ni de sus implicaciones. Efectivamente, raramente constituye un titular de informativo, pero su desarrollo se está produciendo de manera paulatina e inexorable, tanto así como sus repercusiones.

Por otro lado, para que esta doble circunstancia esté presente en nuestras vidas, deben concurrir un número de elementos que actúan como eslabones de una larga cadena, y como tal, muestra también su debilidad: con que un sólo eslabón falle, la cadena deja de cumplir su cometido. En este contexto, la presente ponencia tiene por objetivo desglosar y analizar las partes que integran la tupida maraña de elementos y relaciones que constituyen hoy en día los contenidos digitales y su gestión, para de esa forma, disponer de la visión holística que nos permita comprender su alcance y consecuencias.

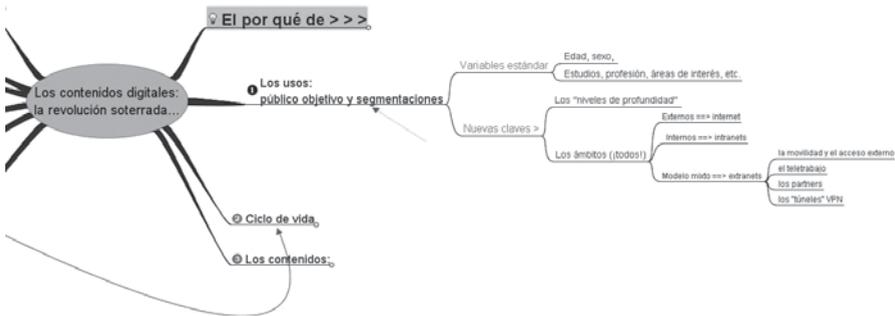


Con este fin, se abordarán los aspectos relacionados con los públicos objetivo a los que van dirigidos, el ciclo de vida de los contenidos, las características propias de los mismos, las condiciones de su estructuración física, su gestión cotidiana, los problemas asociados a la búsqueda y recuperación, así como en dos capítulos específicos, los riesgos potenciales y los beneficios reales que acompañan a toda nueva tecnología. Esta exploración se llevará a cabo a su vez indistintamente tanto desde el punto de vista del generador de contenidos, como del destinatario de los mismos, en la medida que ambas realidades son partes consustanciales e indisolubles de los individuos en nuestra sociedad actual.

LOS USOS: PÚBLICO OBJETIVO Y SEGMENTACIONES

Una percepción un tanto «candorosa» y que en muchas ocasiones yace en el subconsciente de los creadores de contenidos digitales, consiste en la asunción de que éstos son de interés para la totalidad de los usuarios del planeta Tierra, sin modulaciones de ningún tipo. Sin embargo, junto a las habituales segmentaciones basadas en parámetros de edad, sexo, país, preferencias, etc., se va configurando también la necesidad de realizar otras nuevas según el grado de interés y conocimiento de cada público objetivo. Así, podríamos hablar de la configuración de unos «niveles de profundidad» en los contenidos en función de las necesidades, preocupaciones y querencias de los usuarios.

Una segunda derivada del planteamiento inicial se plasma en la tendencia habitual de dirigir los contenidos exclusivamente hacia el mundo Internet en general. Sin embargo, junto a éste, coexisten otras esferas igualmente relevantes: me estoy refiriendo a las Intranets corporativas y los públicos «internos», y por ende, las Extranets, como modelo mixto que aglutina ambas realidades. Lógicamente,



cada una de ellas condiciona de alguna manera a los contenidos digitales, de forma que es importante tener consciencia de sus requisitos y modos de funcionamiento a fin de poder crear y gestionar los contenidos para que cumplan la finalidad a la que están destinados¹.

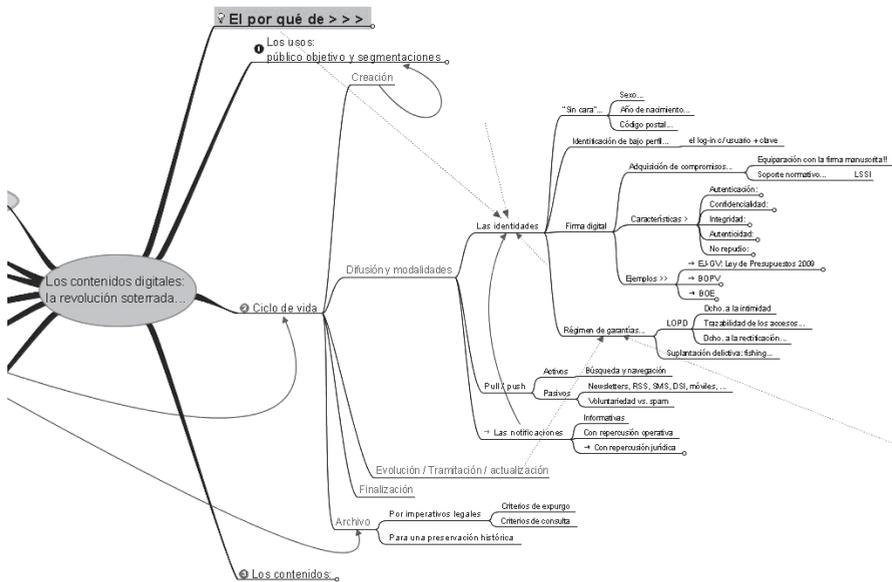
En este sentido, las necesidades de los individuos en su faceta de miembros de la sociedad, serán diferentes a las que tengan como profesionales que trabajan en el seno de una organización comercial y que, a su vez, serán diferentes si son miembros de una comunidad investigadora, etc. Al mismo tiempo, el ámbito de las Extranets ofrece unas extraordinarias potencialidades de desarrollo, pero comporta también un mayor nivel de complejidad en la gestión de los contenidos. Y para hacernos una idea de magnitud, baste citar los escenarios de movilidad laboral (telecomunicaciones, dispositivos,...), del teletrabajo, como una derivada de la anterior (con todas sus repercusiones operativas, culturales y personales) o la integración de *partners* o socios-colaboradores externos en los ciclos productivos. En todos los casos, los contenidos habrán de ser tratados y gestionados de acuerdo a unos requerimientos específicos y adaptados a cada circunstancia.

CICLO DE VIDA

Adentrándonos por las etapas básicas del ciclo de vida de los contenidos digitales, podremos constatar en primer lugar, que la creación de contenidos digitales vendrá condicionada por las segmentaciones de públicos objetivos que veíamos en el apartado anterior. Pero al mismo tiempo, podremos observar que al plantearnos la difusión de los mismos, empieza a actuar también la cuestión de las identidades...

El nivel elemental viene representado por el usuario absolutamente anónimo, «sin cara», y del que, sin embargo, podemos llegar a conocer bastantes aspectos, tanto de forma colaborativa, como deductiva a partir del rastro de su actividad en

¹ Conviene no perder de vista que, en última instancia, además de por una difusión genérica del conocimiento (enriquecimiento personal-colectivo), dicha finalidad viene descrita también por la exigencia de satisfacer las necesidades de los destinatarios en particular.



red. En el primer caso, la petición de facilitar p.ej. tres simples datos como puedan ser el año de nacimiento, el sexo y el código postal de su población, abren la vía al conocimiento del usuario. Si a éste se suman sus gustos y/o áreas de interés inferidas de sus pautas de comportamiento, el cuadro comienza a generar perfiles de usuarios. El siguiente escalón en esta línea viene representado por la asignación individual de un identificativo y clave de usuario para el acceso a determinados contenidos, aunque bien puede seguir permaneciendo en el anonimato. Sin embargo, el salto cualitativo verdaderamente trascendente se produce cuando se identifica jurídicamente al individuo que accede a los contenidos digitales. En primer lugar, como autenticación que nos identifique sin lugar a equívocos, y en segundo lugar, como reconocimiento de su capacidad jurídica para la toma de decisiones.

En este sentido, el elemento fundamental y desencadenante de este nuevo modelo ha sido la concepción de la firma digital, tanto en sus aspectos técnico-operativos, como en los de desarrollo normativo de aceptación, regulación y/o adecuación. El resultado ha sido la implantación de un sistema que asegura los cinco parámetros básicos de un sistema de confianza: la autenticación, la confidencialidad, la integridad, la autenticidad y el no repudio². Este sistema ha venido acompañado de

² Descripciones:

- **Autenticación:** cómo identificamos de forma inequívoca al acceder a un servicio sin que otro nos pueda suplantar.
- **Confidencialidad:** cómo conseguir que las comunicaciones no puedan ser escuchadas por otros y enviar documentos que sólo puedan ser leídos por el destinatario indicado.

una actuación en el marco normativo que ha posibilitado la instauración de entidades públicas de certificación que actúan como terceros independientes³, además de otorgar validez jurídica a la firma digital y establecer sus requisitos. En este sentido, este entramado garantiza sistemas de comprobación transparentes para todas las partes involucradas en una interacción.

Un aspecto muy relevante vinculado a la noción de la identidad de los usuarios lo constituye la existencia también de un régimen de garantías. Así, y al margen de las provenientes de la esfera de la propia firma digital, está la faceta correspondiente a la protección de datos, donde se concretan derechos básicos como la salvaguarda de la intimidad, de rectificación, responsabilidades en la custodia de datos, la trazabilidad de los accesos, etc. Todo ello configura un totalmente nuevo escenario de desarrollo y actuación en los contenidos digitales, que en el siguiente capítulo tendremos ocasión de explorar.

En cuanto a las modalidades de la difusión, cabe constatar que junto a la tradicional exploración y búsqueda constante para acceder a la información originada por el usuario («pull»⁴), coexisten y adquieren en la actualidad mayor peso los sistemas automatizados de difusión de información que nos mantienen al día de los cambios y novedades surgidos en aquellas áreas de interés a las que previamente nos hayamos suscrito («push»⁵). Así, los sistemas de suscripción se han multiplicado desde los correos electrónicos, a los sistemas RSS de sindicación de contenidos, los sistemas de mensajería electrónica, los SMS para móviles, etc.

Por otro lado, de la integración de las dos anteriores (firma digital y sistemas «push») aflora una nueva área destinada a adquirir una alta relevancia y trascendencia; nos referimos a las notificaciones electrónicas. Es importante reconocer el diferente papel que éstas pueden desempeñar. Así, en un primer nivel básico se

-
- **Integridad:** cómo asegurar que un documento no ha sido modificado.
 - **Autenticidad:** cómo poder demostrar ante terceros que un documento es auténtico, el equivalente de la firma autógrafa.
 - **No repudio:** cómo asegurar que un documento ha sido reconocido y aceptado por una o varias personas.

Fuente: *Izenpe*, <http://www.izenpe.com> (marzo 2009)

³ La figura del «Prestador de servicios de certificación» alude a la persona física o jurídica que expide certificados electrónicos o presta otros servicios en relación con la firma electrónica. Viene contemplado y definido a nivel europeo por la Directiva 93/1999, que es la que ha establecido el marco común aplicable a todos los países de la Unión Europea por el que los certificados cualificados son merecedores de confianza por cualquier tercero que confía, y sus certificados otorgan a la firma electrónica avanzada a la que acompañan el mismo valor que tiene la «firma manuscrita» o «firma ológrafa». En el marco español, se define en la Ley 59/2003 de Firma Electrónica.

⁴ Denominación dada al servicio de Internet donde el origen o requerimiento inicial para la transmisión de la información tiene su origen en la parte del «receptor».

⁵ En contraste con el anterior, el término *Push* describe aquél servicio de Internet donde la iniciativa o el origen para la realización de una determinada interacción proviene del «emisor» o servidor central. En estos casos, los servicios son configurados como suministradores automatizados de información desencadenados en función de determinados parámetros. También tiene su variante sin suscripción o involuntaria, habitualmente utilizada en la publicidad intrusiva y no deseada.

agrupan las de carácter meramente informativo. Un segundo nivel viene representado por aquellas notificaciones que conllevan una repercusión operativa, como por ejemplo, los avisos de averías, alarmas, etc. En un último nivel están aquellas que cuentan con repercusión legales de toda índole, como por ejemplo, las notificaciones judiciales, de procedimientos administrativos, etc.

Continuando por las siguientes fases del ciclo de vida de los contenidos, llegamos a la etapa evolutiva, donde experimentarán actualizaciones y/o modificaciones para, seguidamente, dar paso a la finalización de su vida activa. A partir de ahí, llegará la disyuntiva elemental del borrado o su mantenimiento/preservación. En este sentido, serán los motivos que originaron su creación los que fundamentalmente determinarán también la pauta a seguir. Así, habrá contenidos que nacen con una vocación de permanencia en red («*on-line*») y, por tanto, ésta constituya un fin en sí mismo. De la misma manera, habrá otros que se originan para una función o actuación coyuntural y que, cumplida la misma, no tenga interés su conservación y puedan ser por tanto borrados. Sin embargo, también los habrá sobre los que recaigan consideraciones legales que, al margen de una desconexión de la red («*off-line*»), obliguen a su preservación. En este caso, existirán también determinados criterios como p.ej. de preservación y trazabilidad de todas las actualizaciones que ha sufrido, de expurgo o de consulta, que habrá que disponer, controlar y gestionar adecuadamente⁶.

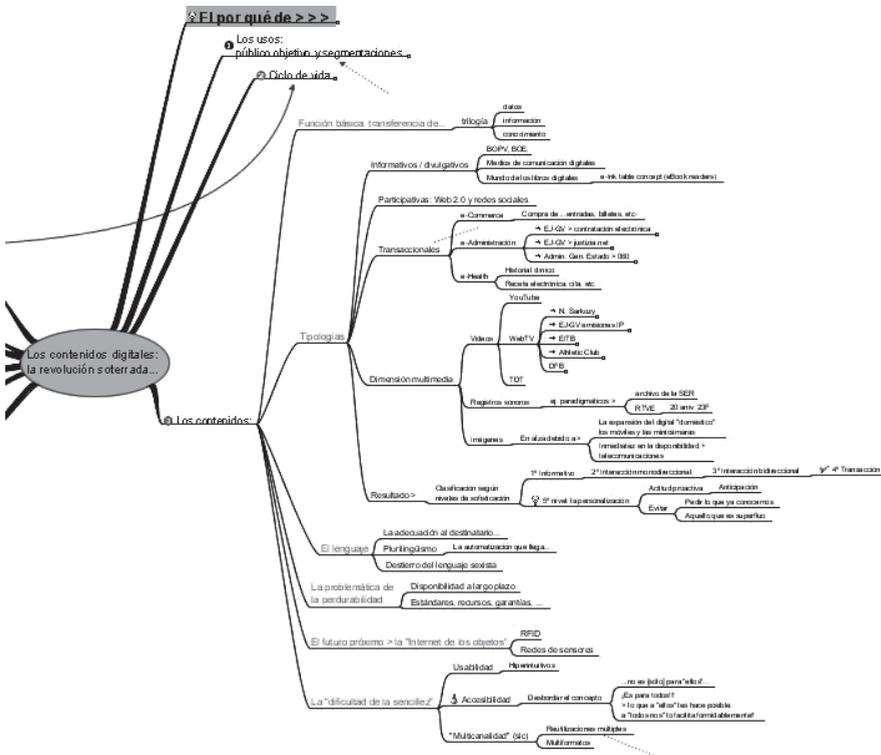
LOS CONTENIDOS

Al hablar de los contenidos propiamente, me parece obligada una mención a los orígenes y subrayar su función básica, es decir, la transferencia de la ampliamente conocida trilogía de datos, información y conocimiento, y a los que, en última instancia, bien se podría añadir una dimensión específica que recoja el ocio virtual o «*entertainment*».

Sobre la base de la finalidad de los contenidos digitales, podemos establecer una elemental agrupación en tres grandes tipologías. Así, la primera haría referencia a los de tipo informativo o divulgativo, normalmente de carácter estático, y donde quedan englobados aquellos de carácter «tradicional» como son las noticias de los medios de comunicación, publicaciones, boletines oficiales, catálogos industriales, etc. Mención expresa, por lo que de «revolución» supone, merece el mundo de los libros digitales y el acceso íntegro a sus contenidos (e-Books)⁷. Y de forma complementaria, referencia al despegue de la tecnología de la «tinta electrónica» o «e-ink», que ha dado paso a la llegada de dispositivos de lectura ligeros y portátiles, a imagen y semejanza del libro físico tradicional (los «ebook readers»)

⁶ Lo relativo a la problemática técnica asociada al archivado, la posponemos a los próximos capítulos, cuando se trate específicamente de los contenidos y su estructuración física.

⁷ Bien sea en forma gratuita, o mediante adquisición. Por citar sólo tres ejemplos, podemos referir el recientemente inaugurado World Digital Library o, anteriormente, el Google Library Project —una vez solventados la cuestión de los derechos de autor (www.googlebooksettlement.com)—, o la sección «eDocs» del Book Amazon.



Un segundo gran bloque tipológico vendría representado por aquellos contenidos de carácter dinámico, muy orientados al establecimiento de vínculos o intercambios de relaciones, donde la bidireccionalidad y el contacto personal juegan un papel primordial... Ejemplos paradigmáticos lo constituyen las redes sociales o lo que se ha dado en llamar Web 2.0. En ellos empieza a aparecer nítidamente la noción de interactividad, así como un nivel de identificación primario. El tercer gran apartado estaría constituido por aquellos contenidos enfocados expresamente a la realización de algún tipo de transacción. En éstos, el mutuo intercambio de datos está enfocado a la adquisición de compromisos para la realización o la obtención de algo. Como ejemplos representativos pueden citarse las enfocadas a adquisiciones en red (todo tipo de entradas a actos, billetes de viajes, ofertas de productos, etc.), los expedientes electrónicos en tiempo real o las recetas electrónicas. En este terreno, además del objetivo y dinamismo propio del contenido, la clave reside en el establecimiento de un sistema de garantías y confianza mutua, tanto más robusto cuanto mayor sea el grado de confidencialidad y/o responsabilidad requerido por las partes. Consecuentemente, se accede a otro estadio en los requisitos de identificación de los interlocutores, donde la autenticación y/o la firma digital entran en juego. De esta manera, en función de la tipología, las características, condicionantes y requisitos de los contenidos variarán ampliamente.

A todo ello resulta hoy en día imprescindible añadir la dimensión multimedia de los contenidos digitales. Ciertamente, la extraordinaria expansión del digital «doméstico» (foto, vídeo e infografías), tanto en forma de minicámaras como, más aún, de terminales móviles multimedia o software 3D, unida a unas características técnicas más que aceptables —hasta no hace mucho exclusivas del entorno profesional—, y junto a la facilidad e inmediatez de las telecomunicaciones, han diseñado un nuevo escenario donde la adopción de este medio está cruzando el umbral de lo novedoso-escaso para adentrarse en el terreno de lo común-habitual. En este sentido, la proliferación de sitios concebidos como repositorios de difusión audiovisual como el paradigmático YouTube, además de su propia y extraordinaria relevancia intrínseca como fenómeno de comunicación, han constituido también un poderoso acicate para la incorporación cada vez más frecuente del vídeo y WebTV como contenidos de los portales. A todo lo cual conviene añadir y no olvidar, las potencialidades ofrecidas por una TDT que despega definitivamente y cuyas posibilidades de interacción con el mundo Internet no son en absoluto desdeñables.

Sin embargo, la presencia de esta dimensión multimedia en el conjunto de los contenidos digitales no vendrá sin dificultades añadidas. Empezando por la gestión de la multiplicidad de canales y formatos existentes (mpeg, rm, wmv, flv/fsw, mov, avi, 3gp,...) que condicionan al usuario a la instalación de nuevos complementos de software en su equipo, continuando por la complejidad de alcanzar un único código fuente estandarizado para cualquier navegador a la hora de «embeber»⁸ la información multimedia en las páginas de los portales, y finalizando, en el caso del «streaming» vídeo en directo, por la necesidad de dotarse de una planificación de parrillas de programación⁹. A las que hay que añadir otros contratiempos obvios, derivados de la gestión de un «elemento más» en el conjunto del pequeño rompecabezas que puede llegar a representar hoy en día un portal complejo. Por último, en relación con esta dimensión multimedia, resulta trascendental subrayar —una vez más— el poderoso atractivo que lo audiovisual ejerce sobre las personas, atractivo que se ve potenciado cuando está unido a la inmediatez. Y como contrapunto, reconocer también el innegable interés que despiertan los registros históricos actualmente al alcance de nuestras manos.

A la vista de todo lo anterior, resulta obvio que en un portal de Internet se pueden combinar contenidos de todos los tipos, cumpliendo cada uno su misión específica. Pero puesto que raramente los portales nacen desarrollados al cien por cien, la progresiva incorporación de nuevas funcionalidades y servicios les lleva a ir pa-

⁸ «Vídeo embebido» es el término con el que se conoce la forma de presentar la pantalla del vídeo en el interior de la maquetación de una página web. En esencia, se trata de un código fuente inserto en el código HTML de la página y donde se establecen las características de esa ventanilla de presentación: tamaño, posición, atributos, nombre y ruta de acceso al vídeo, etc.

⁹ «Streaming vídeo» hace referencia a la técnica del suministro de la información del vídeo a través de Internet. Mediante ella, se establece un pseudo-canal virtual con el usuario que haya solicitado su visión, por donde se le va remitiendo la imagen de manera que empiece a visionarlo en tiempo real, sin necesidad de esperar a la descarga completa del mismo. Resulta imprescindible cuando se trata de realizar la emisión «en directo» de un evento.

sando por diferentes etapas. Así, esta evolución ha venido clasificándose tradicionalmente en cuatro estadios consecutivos que han servido de marco de referencia en función del grado de sofisticación conseguido en los servicios ofertados:

- 1.º Informativo: constituye el nivel elemental.
- 2.º Interacción monodireccional: representado típicamente por los interfaces de tipo formulario de cumplimentación, donde se facilita a los usuarios la aportación de datos para un determinado servicio, pero que no le exime de medidas complementarias de tipo tradicional.
- 3.º Interacción bidireccional: donde a la recogida de datos, le sigue una entrega de información relacionada con aquella.
- 4.º Transacción: caracterizado por la adición de que el intercambio de datos tiene por finalidad la consecución de una transacción entendida como un trato entre dos o más partes. El cambio cualitativo es muy importante debido a que, contrariamente a los estadios anteriores, en éste se llega a identificar a los intervinientes, se formalizan «pactos» en forma de contratos o convenios, y se asumen compromisos y responsabilidades. Junto a esta característica, y ya más en el ámbito específico de la e-Administración, también se hallan presentes otros aspectos fundamentales tales como el análisis y la reflexión sobre la pertinencia y utilidad real de lo que se solicita, el principio de evitar la repetición sistemática de toma de datos ya conocidos o el de liberar al ciudadano de la carga de la prueba, que pasa a ser responsabilidad de las administraciones su verificación (por supuesto, electrónica).

Centrado ya más en la administración electrónica, sobre esta clasificación se ha añadido un quinto estadio representativo de otro salto cualitativo significativo:

- 5.º Personalización: En ella se hace referencia a una etapa en la que el proveedor del servicio adopta una actitud proactiva a la hora de facilitar a sus clientes o interlocutores¹⁰ el acceso a determinadas prestaciones o la cumplimentación de requerimientos. Se fundamenta en el conocimiento que el proveedor tiene de los mismos procedente de actuaciones previas, y desde donde se pueden anticipar determinadas necesidades.

En el camino de evolución de un portal por los diferentes estadios, si bien las tres primeras no presentan excesivas dificultades, el acceso a las dos últimas etapas de sofisticación lleva aparejadas importantísimas repercusiones en el seno de la entidad propietaria del mismo. Por un lado, en el nivel técnico de los contenidos digitales, por la complejidad de las arquitecturas y aplicaciones requeridas y, donde asimismo, figu-

¹⁰ Eterno debate sobre la idoneidad del uso del término «cliente» —proveniente de la esfera de lo privado— para dirigirse a los «ciudadanos» —referentes por antonomasia de la esfera de lo público—. Personalmente, aún admitiendo una hipotética falta de exactitud u ortodoxia gramatical, considero acertado su utilización por lo que conlleva de cambio cultural, de filosofía y de mentalidad a la hora de afrontar la prestación de un servicio público.

rarán compromisos ineludibles de confidencialidad y protección de datos. Por otro, en el plano organizativo interno, donde será necesario «repensar» cómo se hacen las cosas en busca de una mecanización y automatización avanzada de los trámites. De ambas cuestiones tendremos ocasión de hablar en siguientes capítulos.

Por otro lado, estas repercusiones también se extienden a la conservación y gestión de nuestros contenidos archivados¹¹. Ya habíamos tenido ocasión de ver anteriormente las posibilidades de borrado o de conservación («*on-line*» u «*off-line*») al llegar a la finalización del ciclo de vida de un contenido. En este sentido, y refiriéndonos al concepto de «borrado», es necesario hacer una advertencia previa sobre la engañosa percepción que provoca este término; en realidad, y especialmente en el caso de contenidos estáticos y que estén alojados en servicios de albergue ajenos, no tenemos ningún control sobre el efectivo borrado de ficheros en la totalidad de sus sistemas (backups incluidos...) Por otro lado, y aún estando albergados en servidores propios, los poderosos algoritmos rastreadores de los buscadores generan un alto número de duplicados en memorias temporales («*cache*») sobre las que nuevamente, tampoco tenemos certezas. De esta manera deberíamos considerar que, una vez publicados con una permanencia mínima, podrían quedar en algún lugar como esculpido sobre piedra, aún habiéndolos borrado.

Desde una óptica tecnológica, no se atisban en la actualidad problemas o amenazas para la cuestión de su conservación provenientes de las meras capacidades físicas de almacenamiento¹². Sin embargo, donde sí será necesario prestar atención es en lo concerniente a la dificultad técnica asociada a la conservación funcional y a largo plazo de datos distribuidos en múltiples sistemas de aplicaciones (aquellos contenidos asociados fundamentalmente a los estadios de interacción bidireccional, transacción y personalización de los portales). Así, la simple copia de seguridad o backup no será suficiente para garantizar la replicación de un entorno y unas funcionalidades concretas. Siendo como tales imprescindibles, deberán de estar acompañadas de planes para la disponibilidad tanto de infraestructuras y entornos operativos, como de cada uno de los sistemas funcionales integrados en el portal; a los cuales, por otro lado, será necesario prever las correspondientes tareas de mantenimiento y actualización individuales, dotar de medidas de seguridad para la preservación de su integridad (no modificación tras su finalización), etc. Tratamiento aparte tienen los sistemas de «captura» y con capacidad de «reconstrucción» funcional nula (sin navegación) o limitada (enlaces relativos de tipo estático), donde la dificultad residirá sin duda más en la calidad del tratamiento informático necesario que lo haga viable, que en la reproducción. En cualquiera de los casos, a todo

¹¹ «Conservación» entendida como la acción de mantener algo o cuidar de su permanencia (RAE, 22.^a edición). Frente a ésta, la «preservación», como acción de proteger, resguardar anticipadamente a una persona, animal o cosa, de algún daño o peligro (cf. *ibídem*), tiene una connotación defensiva o de protección, más relacionada con los «riesgos potenciales» planteados más adelante.

¹² Superada la problemática tradicional asociada a los metros lineales de estanterías, los terabytes de almacenamiento no sólo han dejado de ser una quimera inalcanzable, sino que empiezan a estar disponibles a precio asequible para los actores más modestos del mercado, ofreciendo alternativas inmediatas a las necesidades crecientes de los contenidos digitales.

ello será preciso añadir también, desde un punto de vista organizativo, la gestión de otros criterios como controles de acceso, los expurgos, etc.

En este contexto integral, un apartado que cada vez adquiere más relevancia, es el referido al lenguaje. En primer lugar, como vehículo de comunicación que ha de estar ajustado al destinatario. Veámos en apartados anteriores toda la casuística asociada a las identidades de los usuarios receptores de los contenidos, así como la referida a los diferentes grados de interés de los usuarios, que aconsejaban la adopción de unos «niveles de profundidad» en los contenidos en función de aquéllos. Al mismo tiempo, un segundo factor también en clara progresión de importancia, viene definido por el plurilingüismo, tanto en su vertiente local, cuando nos referimos a comunidades con más de un idioma oficial, como en la universal, donde la globalización acarrea la imperiosa necesidad de posibilitar y facilitar la comunicación con otras culturas y mercados. En cualquiera de los casos, supone una dificultad añadida representada tanto por la deseable simultaneidad en la publicación de los contenidos, sin o con mínima penalización en los plazos, como por la problemática asociada a otras funcionalidades como la navegación entre idiomas, versionados diferenciados, interrelación con los usuarios, etc. En este sentido, los nuevos desarrollos en herramientas de automatización de traducciones ofrecen ciertamente un esperanzador horizonte. Finalmente, la necesidad de tomar en consideración —e incorporar— la cultura del uso de un lenguaje apropiado, donde los desafectos sexistas, machistas o de otro tipo queden descartados.

Complementariamente, pero como elemento cada vez más imprescindible en la elaboración de contenidos, figura la cuestión de la accesibilidad, entendida como «*acceso de todos a la Web, independientemente del tipo de hardware, software, infraestructura de red, idioma, cultura, localización geográfica y capacidades de los usuarios*»¹³. En este sentido, y aunque sea largamente desconocido, no sólo no es una «carga» destinada exclusivamente a determinados colectivos minoritarios, sino que es una construcción beneficiosa para todos: lo que a dichas personas les hace posible, ¡a «todos nos» lo facilita formidablemente! Por otro lado, en el ámbito de un sector tan importante por su influencia y extensión como la administración pública, empieza a haber ya compromisos y plazos legales concretos para su efectiva puesta en práctica. De manera que, primero por convicción ética, después por demanda social y, también, por imperativos legales crecientes, la tarea de cumplir con escrupulosidad los criterios técnicos para asegurar la accesibilidad de los contenidos a todos los usuarios ha dejado de ser algo secundario o prescindible. En cualquier caso, la accesibilidad ha de considerarse como una disciplina o especialización propia y compleja, que requerirá de una atención permanente y nada desdeñable.

Por último, y apuntando hacia el futuro no muy lejano, es necesario mencionar el mundo —y la nueva revolución— que viene bajo la denominación de «Internet de las cosas» («Internet of Things»). Si bien hasta ahora hablábamos de los contenidos como algo intangible, propiciado por el formato digital, próximamente debere-

¹³ Definición de la *Web Accessibility Initiative* (WAI) surgida en el seno de la *World Wide Web Consortium* (W3C) <http://www.w3c.es/divulgacion/a-z/>

mos readaptar esta noción para incluir también a los objetos físicos. Porque gracias a nuevas tecnologías como la RFID¹⁴, las redes de sensores, el posicionamiento por GPS o las comunicaciones inalámbricas de corto alcance, aquéllos pasarán a ser visibles y accesibles desde Internet. Esta interrelación entre el ciberespacio y el mundo real abre las puertas a nuevas posibilidades de interacción tanto a nivel empresarial como individual. Ejemplos como la ejecución y automatización remota de procesos (gestión de almacenes, puntos de venta, etc.) o la asistencia y reparación a distancia de maquinaria (lectura y monitorización de parámetros, test de diagnóstico de averías, etc.), pueden dar una muestra de las potencialidades que encierra.

Para terminar de dibujar el panorama sobre la idiosincrasia de los contenidos digitales, tiene sentido puntualizar dos tópicos hondamente asimilados en el subconsciente colectivo. El primero hace referencia a la percepción de simplicidad que rodea a la Sociedad de la Información. Si bien es cierto que el mundo virtual que conlleva es de fácil e intuitiva utilización, concluir sin embargo de dicha circunstancia que la gestión de los contenidos digitales es igual de simple, constituye un gran error. Para que dicho uso sea así de fácil, hay una gran complejidad residente entre bambalinas. Constituye lo que bien podríamos denominar la «dificultad de la sencillez».

El otro tópico «mental» al que aludíamos es el referido a una cierta percepción inmovilista de los contenidos. Sin embargo, los tiempos de la creación de contenidos desde una perspectiva de «colgar y olvidar», hace tiempo que dejaron de tener sentido y han dado paso —y lo harán más todavía— a una perspectiva de «publicar e interactuar».

LA ESTRUCTURACIÓN «FÍSICA»

La implantación y extraordinaria expansión de Internet como la «red de redes» vino definida por varios elementos fundamentales: en lo relativo a la parte de contenidos, estuvo marcada por la noción del Hipertexto y la emergencia de un nuevo lenguaje a él asociado, el HTML. En cuanto a la parte de infraestructura, por el sistema IP para la comunicación de datos a través de una red de paquetes conmutados, junto con el protocolo HTTP de transferencia de hipertextos¹⁵.

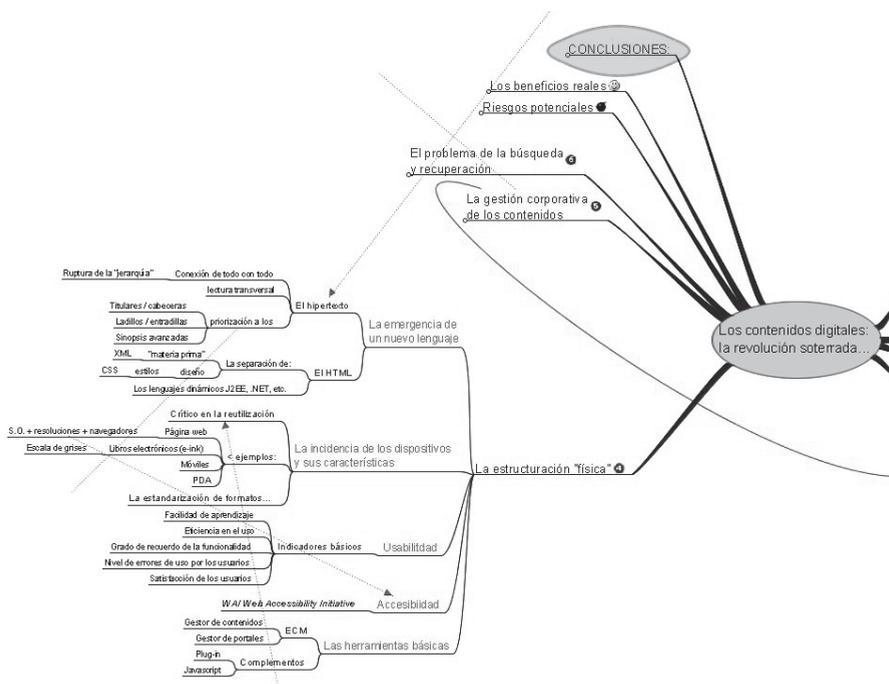
¹⁴ Radio Frequency Identification - Identificación por radio frecuencia. Hace referencia a dispositivos electrónicos miniaturizados y de bajo coste, que incorporan la posibilidad de almacenar (sin consumo eléctrico) una cierta información básica en forma de campos de códigos para identificación o para indicación de estados, y que, a su vez, tienen la propiedad de responder a un estímulo electrónico externo mediante la emisión de dichos datos (lo que se conoce como modo transpondedor). Debido a sus características de miniaturización y bajo coste, se posibilita su utilización masiva insertos en etiquetas que se adhieren a los objetos.

¹⁵ Hipertexto es el nombre dado a los textos que permiten al lector saltar o navegar en una pantalla de ordenador a otro/s texto/s relacionado/s.

HTML HyperText Markup Language o lenguaje de marcas de hipertexto.

IP Internet Protocol.

HTTP Hypertext Transfer Protocol o protocolo de transferencia de hipertexto.



Pues bien, los dos primeros fundamentos han marcado una nueva forma de concebir y actuar con los contenidos digitales que han aparejado múltiples consecuencias. Comenzando por el hipertexto, su puesta en escena ha posibilitado una conexión de «todo con todo», que ha significado la ruptura en la secuenciación de la lectura. Esta ruptura de la «jerarquía» discursiva ha facilitado enormemente el acceso rápido a grandes volúmenes de información. Sin embargo, ha generado también importantes cambios en las pautas de lectura. Así, nos estamos habituando a un modo de lectura «transversal» o «entre líneas-párrafos», donde el usuario va buscando información selectivamente a partir de titulares, entradillas y/o sinopsis avanzadas (y cortas), descartando rápidamente el resto de la información, hasta llegar al contenido/s deseado/s. En este sentido, la conclusión obvia apunta a la importancia capital de una adecuada estructuración del contenido de forma que incluya los datos claves en los elementos antes citados.

Por otro lado, con el lenguaje HTML se ha pasado de la tradicional concepción «unitaria», donde la unidad básica de información estaba representada por «el» documento, a una arquitectura virtual donde esa misma unidad básica está conformada a su vez por múltiples sub-elementos individuales¹⁶, que pasan a gestionar

¹⁶ XML, CSS u hojas de estilo, ilustraciones JPG, GIF, PNG, ficheros con información adicional — PDF, documentos, hojas de cálculo, presentaciones —, etc.

de forma separada los componentes textuales, visuales y estéticos. Este planteamiento, que conceptualmente se expresa de forma sencilla, encierra una complejidad ciertamente creciente. Empezando por la incorporación de lenguajes para la generación dinámica de HTML, tales como JSP, J2EE, .NET, ASP, etc., que representan la apertura a la inteligencia de las aplicaciones (bases de datos, servidores de aplicaciones, etc.) y a la interactividad con los usuarios; y continuando por la problemática asociada a asegurar la operatividad de funcionalidades que no siempre están soportadas en todos los dispositivos (o se comportan de forma diferente). Y a lo que hay que sumar la necesidad de «reconstruir» en tiempo real el todo a partir de las partes, contando con que cada una de las piezas puede tener ubicaciones en unidades o repositorios separados, que han de estar en todo momento disponibles y con tiempos de respuesta aceptables, etc.¹⁷ Mención a parte merecen otros recursos técnicos disponibles y que posibilitan determinadas funciones no soportadas en su estado actual por los navegadores y/o los servidores web como son los complementos / plug-in¹⁸ y los javascripts. En ambos casos tratándose de pequeñas aplicaciones, en los primeros, significa la necesidad de su descarga e instalación; en los segundos, se trata de código que para ser interpretado, también requiere de la disponibilidad previa de software-base. Habitualmente ofrecen prestaciones audiovisuales que, complementándolo, van sin embargo más allá de las posibilidades del lenguaje HTML y añaden extraordinarias posibilidades creativas. En cualquier caso, e independientemente de las posibles ventajas e inconvenientes que puedan representar, en muchas ocasiones suponen una fuente de conflicto con los criterios de accesibilidad.

Una consecuencia derivada de esta segregación de los elementos constituyentes de los contenidos es que abre la puerta a que puedan ser adaptados a la utilización desde diferentes soportes y plataformas de difusión, mediante el oportuno ajuste en la «reconstrucción» de los mismos a las diferentes características y requisitos técnicos que de dichos soportes necesariamente imponen¹⁹.

Un concepto fundamental que emerge inmediatamente de la experiencia de interactuar con los portales y contenidos de Internet es el de usabilidad, que fue definido por Jakob Nielsen como el atributo de calidad que mide lo fáciles que son de usar las interfaces Web²⁰. El principio subyacente está vinculado con la filosofía que prioriza permanentemente el diseño centrado en el usuario. Así, a la hora de medir el «grado de usabilidad» de un sistema, en esencia se contemplan cinco indicadores: facilidad de aprendizaje, eficiencia en el uso, grado de recuerdo de

¹⁷ Especialmente crítico cuando ya la mayoría de los servicios web están soportados en conjuntos de aplicaciones software operando desde diferentes servidores e íntimamente integrados entre sí. Es la debilidad del eslabón roto de la cadena que se ha comentado con anterioridad.

¹⁸ Por citar algunos muy extendidos: Acrobat Reader, Windows Media Player, Flash, etc.

¹⁹ Ténganse en cuenta, a modo orientativo, la siguiente casuística: ordenadores, PDAs, teléfonos móviles, lectores e-Book, TV-TDT declinados en forma de marcas y modelos; a su vez, todos ellos con sus respectivos —y diferentes— sistemas operativos, navegadores, pantallas (tamaños, resoluciones y colores), capacidades de ancho de banda, agregadores de información para RSS...

²⁰ «Usability 101: Definition and Fundamentals - What, Why, How» (<http://www.useit.com>).

su funcionalidad, nivel de errores cometidos por los usuarios en su utilización y la satisfacción de los mismos. Como se podrá apreciar, no se trata de apreciaciones subjetivas, sino que se está haciendo referencia a una medida empírica y que es de carácter relativo respecto a los objetivos fijados de usabilidad. Así pues, es un aspecto que condiciona de manera clara la estructuración y construcción de los contenidos.

Junto a la usabilidad, otro gran elemento a considerar es la ya previamente citada cuestión de la accesibilidad de los contenidos. Así, apoyándose sobre el principio de separación entre «materia prima» y «diseño», la construcción bajo estándares accesibles permite funcionalidades tales como la escalabilidad de los textos, la variabilidad de los contrastes de los colores, la interpretación correcta por sintetizadores de voz, la identificación de idioma, la detección de términos o citas textuales en otros idiomas, la estructuración lógica de la información, la intuitiva utilización de dispositivos diferentes —y mucho más críticos en sensibilidad— que el ubicuo ratón, etc. Para que esto sea posible, resulta imprescindible el uso escrupuloso y disciplinado de las correspondientes directivas técnicas. Indudablemente, frente a la libertad absoluta, estas pautas que impone la accesibilidad suponen una cierta limitación y/o rigidez a la hora del tratamiento creativo, estético o funcional de los contenidos; sin embargo, existen ya estándares y normativas que correctamente aplicados, proporcionan altos grados de accesibilidad. Ahí están los trabajos de la *Web Accessibility Initiative* (WAI) en el seno de la *World Wide Web Consortium* (W3C) para demostrarlo. En cualquier caso, la accesibilidad se perfila como un área tecnológica de alta especialización y que requerirá de mayores esfuerzos y cuidados para su correcta y constante implementación.

En este marco de trabajo, las herramientas básicas que posibilitan el tratamiento de los contenidos digitales están habitualmente englobadas bajo el término genérico de ECM «*Enterprise Content Management*», estando normalmente disponibles en las Intranets corporativas o mediante Extranets. Entre los módulos o funcionalidades básicas que aportan se pueden diferenciar dos: la gestión de portales y la gestión de contenidos. Los gestores de portal son los que permiten crear y administrar el contexto o marco donde son presentados los contenidos (el «lienzo» donde se «pintan»). Proporcionan el esqueleto del portal de Internet, le aportan una identidad, una visión de conjunto y le dotan de una lógica de navegabilidad. Los gestores de contenidos están, en cambio, orientados a las funciones específicas de edición, maquetación-plantillas, catalogación, establecimiento de enlaces, publicación, control de versionado, etc. de los propios contenidos de tipo estático. En el caso de los contenidos dinámicos, son las propias aplicaciones las que incorporan las plantillas adecuadas que configuran las interfaces de interacción hombre-máquina, tanto sean para búsquedas, tomas de datos, etc., como para presentación de resultados de las acciones solicitadas.

Y si antes hablábamos que la segregación de la información en sus elementos constitutivos permite la «adaptación» técnica de los contenidos a las características de los soportes, la utilización del hipertexto, junto a la inteligencia de las aplicacio-

nes y la posibilidad de una gestión multicontextual, posibilita una amplia «reutilización» de los diferentes contenidos desde diferentes entornos, situaciones y/o momentos, en función de las necesidades concretas de los usuarios del sistema. Esto, unido a la búsqueda de sinergias con otros actores —internos y externos— concurrentes en nuestro ámbito de interés, es un atributo fundamental a la hora de conseguir una prestación de servicios de forma óptima y eficiente.

En este punto relativo a la búsqueda de sinergias y la utilización del hipertexto, cabe mencionar la cuestión de a dónde «apuntar»: o bien a un contenido concreto (individualizado caso por caso) o, por el contrario, remitir siempre al acceso principal del portal donde esté albergado. Éste último caso resulta, obviamente, el más sencillo pero, a la vez, el más pobre de cara a nuestros usuarios, que agradecerán ser encaminados directamente al documento de interés. Sin embargo, esto supone para el gestor asumir un riesgo inherente al hipertexto no siempre suficientemente explicitado y que es pertinente subrayar: los enlaces rotos. El establecimiento de enlaces que apunten a contenidos ajenos —¡o propios!— conlleva la posibilidad de que cuando éstos cambien o desaparezcan en su forma presente, el resultado será un enlace roto y un mensaje de error. En este sentido, el recurso profuso de links a contenidos concretos aconseja mantener una tarea de comprobación periódica de los mismos.

LA GESTIÓN CORPORATIVA DE LOS CONTENIDOS

Si nos colocamos en la posición del usuario individual que crea y publica contenidos digitales en un contexto doméstico, probablemente el número de tareas a realizar resulte asumible para una sola persona. Sin embargo, desde una perspectiva específica de «emisores» de contenidos a nivel corporativo, la gestión de los contenidos digitales presenta un panorama que, lejos de simplificarse, se vuelve cada vez más complejo. Así, sin perder de vista lo mencionado en el capítulo anterior respecto a las características intrínsecas de su estructuración física, resulta imprescindible abordar esta cuestión de la gestión de contenidos contemplando el ciclo de vida en su globalidad y, como tal, darle respuesta en su conjunto. En primer lugar, porque el ciclo de vida de contenidos ya no será más el simple encadenamiento de acciones aisladas, sino antes al contrario, estará basado en la participación de múltiples y diferentes actores para su elaboración (...y su posterior actualización, y finalización), que, a su vez, se producirá en diferentes momentos, con diferentes funciones y responsabilidades y donde, además, la gestión de los tiempos y/ plazos cobrará progresivamente más importancia. Y al mismo tiempo, porque factores tales como su número o el aumento del grado de interactividad ligado a los mismos contribuyen también a la necesidad de un *modus operandi* estructurado y específico a la hora de su gestión.

Esta situación nos encamina invariablemente hacia una operatividad basada en la Gestión por Procesos, que se podría explicar de forma sencilla como una secuencia de actividades que transforman un conjunto de entradas en salidas —en

zar cuáles son susceptibles de ser totalmente automatizables y cuáles requieren la intervención humana (individual o colegiada) en algún punto del proceso²³.

En última instancia, de lo que estamos hablando es de un escenario donde se producen interacciones entre los contenidos digitales con unas reglas de negocio, en un marco de procesos estructurados y flujos de trabajo. Obviamente, el salto cualitativo que representa en complejidad la operación de este entramado hace recomendable el recurso a herramientas específicas de gestión, comúnmente identificadas bajo la denominación genérica de «*Business Process Management*» (BPM) o administración de los procesos del negocio.

Adentrándonos ahora en el área del análisis vinculado a los contenidos digitales y sus resultados, un elemento fundamental a integrar en la gestión de los contenidos lo constituye el «*feed-back*» o retorno de información por parte de los usuarios²⁴. Acerca de esta cuestión, tiene sentido distinguir entre la realimentación implícita, que viene encarnada por las estadísticas de uso, de aquella otra explícita, representada por los sistemas de votaciones y/o ranking valoración, los canales de recepción de opiniones de los usuarios, sus demandas, sus peticiones individuales, etc. En el caso de la implícita, su materia prima básica la representan las «huellas» dejadas por los usuarios en su navegación e interacción con los contenidos²⁵. La segunda nos conduce al despliegue de sistemas CRM²⁶ específicos para la administración de la relación con los clientes²⁷.

Saltando a un ámbito más especializado, y pasando ya al plano estratégico de la organización, llegaremos a los sistemas de «*Business Intelligence*» (inteligencia empresarial o de negocio) destinados a descubrir patrones de comportamiento en el análisis de las ingentes cantidades de datos extraídos del uso dado a los contenidos digitales.

En este maremágnum de líneas de trabajo y derivaciones vinculadas a la gestión de contenidos, no debe tampoco olvidarse lo referido a lo ya previamente señalado respecto a la problemática asociada a la preservación y gestión de los contenidos archivados.

Por último, considero relevante mencionar el papel del plano jurídico en la gestión corporativa de los contenidos. La actual Sociedad de la Información, entre otras características, destaca también por ser una sociedad donde es enormemente fácil la trasgresión de derechos legales de personas y/o entidades, y donde,

²³ Véanse las pólizas «express» de seguros, los préstamos personales instantáneos, etc.

²⁴ Ya veíamos a este respecto en el apartado dedicado a «Los contenidos», que éstos cada vez serán menos del tipo «colgar y olvidar», para pasar a ser del tipo «publicar e interactuar».

²⁵ A título meramente de ejemplos representativos, se pueden citar los siguientes: el n.º de páginas vistas, n.º de transacciones completadas (y su proporción respecto al anterior), documentos más visitados, rutas más seguidas, palabras y términos de búsqueda más utilizados, franjas horarias de acceso, tiempos de permanencia, puntos de abandono, etc.

²⁶ Customer Relationship Management.

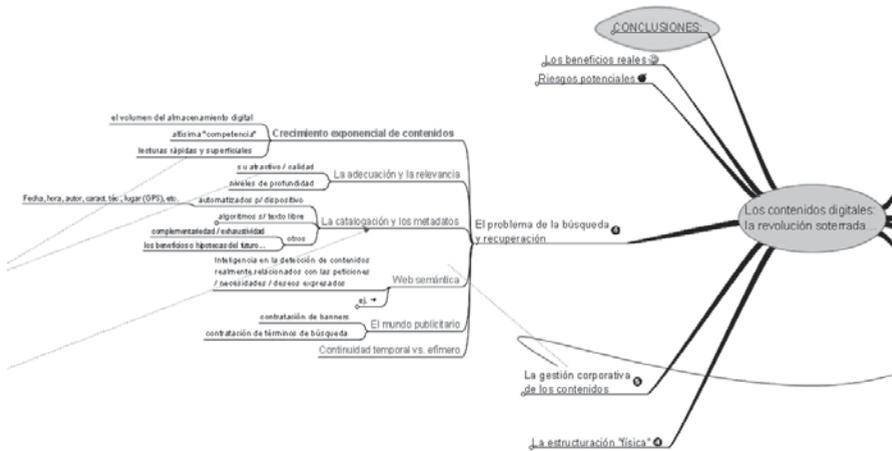
²⁷ Quizá ahora se pueda apreciar mejor y en toda su extensión lo que apuntábamos respecto a relevancia de la dualidad «identidades vs. anonimato», en el apartado «Ciclo de vida > Difusión y modalidades».

al mismo tiempo, se contraponen una tendencia garantista en auge que se manifiesta, tanto en una mayor sensibilidad autorreguladora, como en una actuación activa de los órganos judiciales. La conjunción de ambas circunstancias, junto con el desarrollo de una cultura social más asociativa, reivindicativa y celosa de sus derechos, está provocando el alza en la tendencia querellante. A la vista de todo ello, resulta patente la necesidad de reforzar las precauciones para que la correcta utilización y protección de datos de terceros esté siempre presente durante toda la cadena de valor en la vida de los contenidos, desde la planificación hasta la ejecución de los procesos involucrados. Comenzando por la elemental gestión de los derechos intelectuales y de explotación, y finalizando con la preceptiva custodia y protección de datos personales. Tanto más cuanto que nos adentremos en contenidos relacionados con el comercio electrónico, administración electrónica, etc. Descuidar este planteamiento conllevará el riesgo de incurrir en responsabilidades legales por el incumplimiento de preceptos legales en la gestión y custodia de datos personales.

EL PROBLEMA DE LA BÚSQUEDA Y RECUPERACIÓN

Asociado al ingente e incesante crecimiento de contenidos digitales en red, se generan varios fenómenos que es preciso tener en cuenta y gestionar en consonancia. Empezando por la constatación de la elevada «competencia» entre contenidos por ocupar las primeras posiciones de las búsquedas, dado que lo que no figure entre las primeras diez o veinte referencias, es *de facto* como si no existiese (la prioridad se ha trasladado de la *presencia* a la *visibilidad*). Se puede continuar con la conocida dicotomía entre adecuación-pertinencia-relevancia y cobertura: la búsqueda con muchos términos (exhaustividad) acaba sin ningún resultado y, por el contrario, la utilización de pocos términos y de carácter común produce un exorbitante número de resultados. A lo que es necesario añadir la certeza de que un porcentaje altísimo de los mismos no se ajustará a las expectativas, aunque haya podido cumplir los requisitos de búsqueda utilizados. En este contexto, la progresiva utilización de nuevas técnicas y recursos permite afinar más en la realización de búsquedas, posibilitando un punto intermedio entre las situaciones extremas antes descritas. Al tiempo que, como contrapunto, el usuario va desarrollando un «olfato» o sexto sentido para la detección, a través de lecturas rápidas y esquemáticas, de aquellas referencias que interpreta como relacionadas con la búsqueda. Y a su vez, hay que contemplar que estas circunstancias se reproducen idénticamente en un doble contexto: el externo, de los grandes buscadores de Internet, y el interno, el de los buscadores propios de los portales corporativos de cierta entidad.

En este estado de cosas, y desde la perspectiva del administrador de contenidos digitales, la prioridad para que sus contenidos sean encontrados por los buscadores será actuar en primer lugar sobre aquellos recursos que están dentro de su alcance. A este respecto, la primera actuación deberá estar enfocada sobre los



atributos y las características de los contenidos. Así, resulta imprescindible que estén adecuadamente catalogados mediante los metadatos²⁸ internos asociados. Si bien algunos podrán venir incorporados automáticamente en origen por el dispositivo utilizado en su generación (fecha, hora, autor, parámetros técnicos, lugar —vinculado a coordenadas GPS—, etc.), otros habrán de documentarse de manera individual (los «METATAG» del propio lenguaje HTML: <Title>, <Description>, <Keywords>, <lang>, etc.), cuidando además, de que haya una correspondencia efectiva entre los descriptores utilizados y la indexación efectiva de los textos que —en su caso— componen el contenido. Al mismo tiempo, esta catalogación deberá realizarse mediante la incorporación de aquellos términos y expresiones que previsiblemente utilizarán los usuarios a la hora de realizar búsquedas al respecto. En conjunto, su utilidad no viene dada tanto por su capacidad de procurar una *presencia* en Internet —que ya está asegurada por los algoritmos de indexación—, sino en alcanzar una adecuada *visibilidad* en el mundo virtual. En definitiva, de su adecuada elaboración dependerán en una significativa medida los beneficios o las hipotecas del futuro.

La siguiente prioridad desde la perspectiva del administrador de contenidos digitales será la de tratar de posicionarlos correctamente en los buscadores externos, sobre los que tiene limitada capacidad de actuación por estar fuera de su alcance. Sin embargo, siempre podrá adoptar todas medidas tendentes a ayudar y facilitar sus labores de indexación automatizada. En este sentido, la cooperación mediante

²⁸ Habitualmente se define de forma sencilla metadato como datos que describen datos. Profundizando un poco más, los metadatos son datos asociados a un documento digital que recogen información fundamentalmente descriptiva (autor, título, etc.). También pueden incluir información de administración (creación del recurso, derechos, control de acceso ...), y preservación (tipo de formato, etc.). Fuente: *Glosario de la Universidad Carlos III de Madrid* <<http://e-archivo.uc3m.es/dspace/help/glosario.html>>

actividades como la configuración en los servidores de óptimas reglas de acceso a los robots de exploración, la remisión de sitemaps exhaustivos y en formatos estandarizados, la notificación de actualizaciones (y su frecuencia), hacer uso extensivo de enlaces a otros contenidos, etc., se presentan como tareas inexcusables a realizar de forma constante.

Otra área complementaria a examinar para la efectiva visibilidad y captura de tráfico en la red hacia nuestros contenidos, viene dado por el mundo publicitario o de patrocinio de términos. El primero, representado fundamentalmente por la contratación de banners en portales de intenso tráfico para que operen como punto de enganche y redireccionamiento de usuarios hacia nuestros contenidos. El segundo, mediante la contratación en los portales de búsqueda de palabras o términos de manera que, por cada ocasión que tengan una coincidencia con las búsquedas de los usuarios, generen una aparición colateral en los resultados ofrecidos a la interrogación. Un tema crítico en ambos casos viene vinculado a la verificación efectiva de las «exposiciones» y/o tráfico reenviado al servidor del anunciante y, por ende, con la configuración y explotación de las estadísticas de accesos (ya mencionada en la gestión corporativa de los contenidos).

Avanzando un poco más allá, para conseguir que una vez localizado-presentado-mostrado al usuario, éste acceda a continuación a nuestro contenido frente a, y/o descartando otros «competidores», la estructuración del contenido deberá estar optimizada para facilitarle esa fase —altamente crítica— de visionado rápido de los resultados de búsquedas. En este sentido, además de suministrar la información básica en el título y la descripción, es necesario hacerlo de manera efectiva y atractiva para evitar el descarte inadvertido; seguidamente, posibilitar una lectura rápida de cabeceras que permita alcanzar una síntesis sobre el conjunto de la información. Aspectos todos ellos también ligados a la cuestión de la usabilidad mencionado en el apartado de la estructuración física de los contenidos.

Sin perjuicio del enorme interés de todo lo referido al posicionamiento en los buscadores externos, resulta imprescindible destacar que la funcionalidad y operatividad de los buscadores internos de los portales corporativos de cierta entidad también revisten de una gran importancia. Y esto es así, en primera instancia, porque la «novedad» de los contenidos recientes siempre estará antes disponible en el buscador interno propio del emisor que en cualquier otro buscador externo. Complementariamente, porque cuando se trata de acudir a las fuentes de la información, los buscadores corporativos ofrecen el atractivo de encontrarse «libres» de mezclar contenidos de otras fuentes ajenas.

En este contexto, tiene especial importancia la adecuada configuración de las herramientas de indexación y los motores de búsqueda implantados en la infraestructura de soporte, con un énfasis añadido en la cuestión de los interfaces hombre-máquina. A este respecto, vale la pena incidir quizá no tanto en los de introducción de datos²⁹,

²⁹ Las interrogaciones sobre texto natural han probado su gran flexibilidad y aportan una indudable familiaridad al usuario. En cualquier caso, nada impide el complemento de otros campos de búsqueda avanzados, área ampliamente desarrollada.

sino en los de presentación. Así, una forma de combatir los elevados números de resultados en búsquedas viene dada por la inteligente agrupación y ordenación de los resultados por múltiples parámetros de discriminación, tales como fecha (con opciones creciente o decreciente), formatos (html, pdf, vídeos, fotos, ppt, etc.), metadatos, relevancia (según la parametrización del uso y/o adecuación), clasificaciones «ad hoc» («niveles de profundidad», tipologías como «ayudas y subvenciones», «noticias», «notas de prensa», «declaraciones», etc.) Otra aportación importante viene de la configuración de búsquedas en cascada, es decir, con la posibilidad de refinar la búsqueda con nuevos términos pero exclusivamente sobre resultados previos.

Profundizando más en las herramientas al uso para la realización de las búsquedas, llegamos a la noción de web semántica. Se trata de aportar una inteligencia en la detección de contenidos realmente relacionados con las necesidades expresadas. Para ello los portales han de incorporar elementos complementarios que representan los significados de las palabras, así como la relación entre ellas³⁰. Complementada con sistemas de lematización que trabajan sobre la base de las raíces morfológicas de las palabras, y la sintaxis de las mismas, se consigue una web extendida que abre la vía al procesamiento complejo y automático que permita razonar y entender correctamente tanto las peticiones de búsqueda de información, como el criterio de filtrado correcto de los resultados presentados³¹. La diferencia pues fundamental respecto a los motores de búsqueda tradicionales es que no trabajan únicamente sobre concordancia de cadenas de caracteres y/o clasificaciones de «popularidad» (representado normalmente por el número de enlaces que apuntan a un mismo contenido), sino mediante la adscripción e interpretación de conceptos, adquiriendo la habilidad de comprender las búsquedas por el contexto creado con las palabras utilizadas.

Junto a esto, la web semántica también puede combinarse y complementarse con otros sistemas de recuperación de información basados en algoritmos de «cálculo/estimación» del éxito (grado de utilización) de los contenidos por otros usuarios y/u otros contenidos visitados posteriormente por los mismos. Adicionalmente, la web semántica representa también una aportación básica en el desarrollo del plurilingüismo y los sistemas automatizados de traducción que hacíamos referencia al hablar del lenguaje en los contenidos.

En cualquier caso, todo lo relacionado con la búsqueda y recuperación de contenidos constituye otra compleja disciplina técnica más, que es necesario comprender en su integridad y aplicar de forma constante para facilitar la búsqueda y recuperación de los contenidos digitales en red.

³⁰ Se abre por aquí el acceso al mundo de las ontologías informáticas y las tautologías.

³¹ Por ejemplo, una búsqueda de «Enfermedades forestales en el continente norteamericano», en una web semántica debería ser capaz de incluir contenidos relativos a las infecciones de árboles, robles, plagas en bosques, etc., en ámbitos geográficos como Norteamérica, Alaska, Estados Unidos (también USA), Canadá, México, al tiempo de excluir las existentes en, p.ej., Hawai.

actuaciones de mantenimiento hardware y software sin interrupción del servicio, etc. Junto a ellos, los imprescindibles planes de contingencia ante lo «inverosímil», desde cortes prolongados de suministro eléctrico, a incendios, inundaciones, etc...³³, o las políticas de copias de resguardo que ofrezcan una garantía de restitución (los *backup*)³⁴. Y al mismo tiempo, las condiciones de seguridad establecidas tanto en lo referente a medidas «defensivas» frente al software malicioso (*firewalls* o «cortafuegos», DMZ o «zonas desmilitarizadas», antivirus, antispam, etc.), como en lo concerniente a los accesos físicos. Aspectos todos ellos más que suficientes para reflexionar sobre en manos de qué organización depositamos nuestro patrimonio intangible: qué servicios nos proporciona, en qué condiciones y con qué garantías³⁵.

En un plano más próximo al ámbito específico de la generación de contenidos y refiriéndonos a las arquitecturas de software, citar tres aspectos de especial relevancia: la integración de sistemas, la adopción de estándares abiertos y el intercambio de datos. El primero, como respuesta a la creciente complejidad de los contenidos y sus interacciones; hoy en día ya no es concebible su gestión desde «islas» o sistemas aislados, máxime cuando estamos entrando en un contexto de implantación de la e-Administración, el *Business to Business* (B2B), la TDT interactiva o la telefonía móvil de la «n^a» generación. A este respecto es pertinente subrayar que, a diferencia de la simple yuxtaposición de funciones, la integración es lo que posibilita que un mismo elemento pueda llevar a cabo varias funciones para diferentes módulos o sistemas en la infraestructura. Y relacionado con la integración, la necesidad de un enfoque específico hacia la optimización en el diseño y el ensamblado de los contenidos de forma que se agilice tanto la descarga como la interacción con los mismos. En cuanto al segundo aspecto mencionado, insistir en los riesgos asociados a los aplicativos o tecnologías propietarias, que *de facto* pueden aparejar restricciones en las integraciones, al tiempo que una dependencia tecnológica y económica alejada de las tendencias del libre comercio³⁶. En este sentido, la adopción de estándares abiertos ha dejado de ser una conveniencia, para pasar a ser un requisito. Finalmente, la perentoriedad de asegurar el intercambio de datos no sólo pensando en el presente, sino también con la vista puesta en los sistemas futuros.

E inexorablemente ligado a todas las anteriores ataduras, el eslabón de las telecomunicaciones, que representan la conectividad con otras redes y el mundo exterior. A los consabidos requisitos de disponibilidad, redundancia, capacidades, etc., se une en este caso el carácter estratégico implícito que le convierte en objetivo de

³³ Baste recordar como ejemplo reciente (03.03.2009) el derrumbe del edificio sede del Archivo Histórico de la ciudad de Colonia, en Alemania.

³⁴ Viene al caso recordar un axioma de un anónimo *gurú* informático de finales de los años 80 que establecía dos tipos de usuarios: los que ya han sufrido una pérdida de datos..., ¡y los que están a punto de sufrirla! (porque, antes o después, pasará; la cuestión estriba en cómo de preparado le pille a cada uno).

³⁵ A este respecto, vale la pena apuntar aquí las problemáticas asociadas al *outsourcing* informático.

³⁶ En cualquier caso, como mínimo siempre se ha de garantizar el acceso desde los aplicativos Open Source o código abierto.

ataque o control desde muchas y diversas instancias³⁷. En este sentido, baste recordar la perentoria necesidad de los servidores de DNS (*Domain Name System*) que albergan el núcleo del sistema de IPs, los nodos de enrutamiento y por donde transcurren todas las comunicaciones de Internet.

Por otro lado, junto a las nuevas servidumbres ya indicadas, tampoco son nada desdeñables otros factores con potencialidades dañinas como las representadas por los colapsos, la brecha digital o lo referido a la intimidad de las personas. Los primeros, por lo que conllevan de potencialidad para provocar interrupciones en servicios que cada vez son más imprescindibles³⁸. La denominada «brecha digital», por lo que puede suponer de segregación de importantes sectores de la población. Y en cuanto a la intimidad de las personas, por tratarse de un tema cada vez más sensible, tanto más, cuanto que los contenidos digitales con información referida a situaciones y/o actividades personales están en constante crecimiento y con importantes previsiones de progresión³⁹.

Sumemos a esto las cláusulas abusivas a las que maquinalmente se presta consentimiento con una mezcla de buena fe, prisas, ligereza, ignorancia o fatalidad al momento de darse de alta en determinados y populares servicios de uso masivo vinculados al fenómeno de redes sociales⁴⁰, o la dificultad añadida representada por la superposición e internacionalización de ámbitos judiciales, y nos encontraremos en situaciones de auténtica indefensión de los usuarios.

A todos ellos será necesario dar respuesta, tanto en el primer caso mediante el estudio de los posibles puntos de estrangulamiento o cuellos de botella exis-

³⁷ Referencias recientes que acuden inmediatamente a la memoria pueden ser la Guerra del Cáucaso de agosto de 2008, donde ambos contendientes trasladaron sus hostilidades al ciberespacio o, anteriormente, el caso de la oleada masiva de ciberataques perpetrada desde Rusia contra Estonia en mayo de 2007 —en ambos casos con serias alarmas en los cuarteles de la OTAN—, o los filtros existentes en China para la libre navegación por Internet.

³⁸ Ejemplos característicos de saturaciones de servicio asociados a eventos masivos y concentrados en cortos lapsos de tiempo pueden ser el estado de las carreteras ante previsiones meteorológicas desfavorables, la publicación de resultados electorales, o de sorteos de pisos, o de ofertas públicas de empleo, la retransmisión de eventos deportivos, etc. Mención especial merece el efecto de vasos comunicantes que se produce cuando un determinado canal se satura o interrumpe (p.ej.: portal de Internet), derivándose las peticiones de información hacia los restantes canales disponibles (p.ej.: el telefónico).

³⁹ En este sentido, recordemos aquí los habituales formularios con campos de datos personales presentes en las páginas web para el alta y acceso a servicios en red, para la reserva y adquisición de viajes, estancias, conciertos, solicitudes electrónicas, etc., a los que ya se pueden añadir también las tramitaciones oficiales de expedientes en la e-Administración, consulta de situación de expedientes administrativos personales, mensajería instantánea, etc. Así, del simple cruce de información entre los datos de identificación (DNI o pasaporte), Seguridad Social y tarjetas bancarias (VISA, MasterCard, etc.), no sólo se obtendrían perfiles, sino que se abriría la vía potencial al acceso a un detalladísimo historial individual de las personas. Si añadiésemos posicionamientos geográficos disponibles mediante la triangulación de telefonía móvil y/o las omnipresentes cámaras digitales de videovigilancia, llegamos a un escenario que, con independencia de sus virtudes, también trae a colación ciertas reminiscencias al Big Brother de George Orwell en su novela *Nineteen Eighty-Four* o a las obsesiones y peripecias de Mel Gibson en *Conspiracy Theory (El complot, Richard Donner, USA, 1997)*

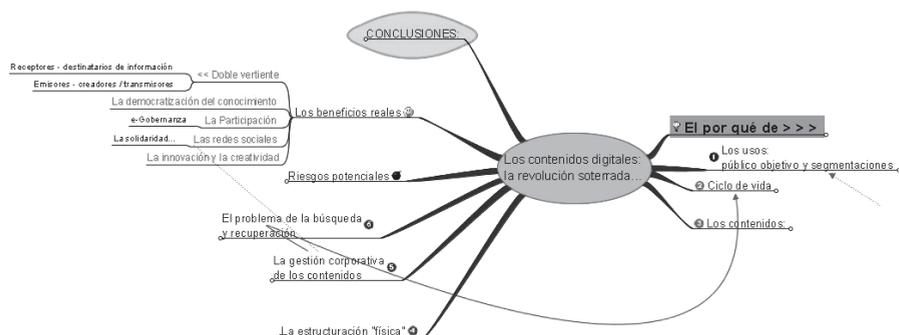
⁴⁰ P.ej., la cesión «inconsciente» de la letra pequeña referente a *todos los derechos de explotación* y en *exclusividad*, origen y raíz de innumerables problemas.

tentes en toda la cadena de la infraestructura y la realización de las oportunas pruebas de carga previas a la apertura del servicio, o bien, como en el caso de la brecha digital, contraponiendo actuaciones que permitan la e-Inclusión⁴¹, o la irrenunciable necesidad de garantías en la protección de nuestros datos y los de nuestros usuarios.

Como se podrá comprobar, no son pocas las áreas sobre las que es preciso actuar en su conjunto o, cuando menos, tener en consideración, a la hora prevenir riesgos potenciales.

LOS BENEFICIOS REALES

Y frente a las sombras proyectadas por los riesgos potenciales existentes, también es necesario contemplar en su justa medida, innegables beneficios tangibles aportados por los contenidos digitales en la doble vertiente de receptores-destinatarios de información, y de emisores-creadores-transmisores de información. Sea como fuere, y ante la obviedad y lo generalizado de los mismos, nos conformaremos con apuntar unas consideraciones universales al respecto.



En lo tocante a los «receptores», ya sean las personas como individuos o la sociedad como colectivo, cabe mencionar en primer lugar lo que, no por evidente, es menos importante: la democratización del conocimiento. Con ello nos estamos refiriendo a la posibilidad de cualquier persona al acceso a los datos e información que requiera, contando además con la libertad de elección y contraste de las fuentes.

⁴¹ Posibilitando primero, que «todos puedan», lo que implica accesibilidad por un lado, e inversión y libre competencia para la disponibilidad por otro; y seguidamente, que «todos sepan», lo que lleva aparejadas la usabilidad y la formación y, asociada a ésta, el impulso de una cultura potenciadora del deseo de aprender y/o curiosidad, junto a un «perder el miedo» a la operación con ordenadores y telecomunicaciones.

Simultáneamente, el desarrollo de redes sociales sin fronteras, capaces de trasladar las relaciones humanas y la solidaridad allí donde se necesite. Y al hilo de dichas redes sociales, la apertura hacia sistemas de Participación que pongan en valor los postulados de la Gobernanza⁴².

Ya desde un punto de vista de generadores o «emisores» de información, destacar por un lado lo que supone de ruptura con las precedentes dificultades y embudos a la hora de distribuir y hacer llegar la información a los interesados, comenzando por la posibilidad de expresar la propia voz⁴³. Por otro lado, lo que conlleva de eficiencia y eficacia a la hora de ofrecer productos y servicios a los usuarios, máxime cuando se plantea en forma de sistemas altamente automatizados y/o el concepto de autoservicio sin horarios (a la carta). En este sentido, destacar la aportación que suponen a la hora de desarrollar actividades económicas en un mundo globalizado y a partir de unos modestos medios propios hasta hace pocos años absolutamente impensables.

Finalmente, mencionar también la contribución que representan en tres esferas de capital importancia: la investigación+desarrollo, la innovación y la creatividad.

CONCLUSIONES

Tras este repaso por los diferentes aspectos involucrados con los contenidos digitales, se pueden extraer varias conclusiones que pasamos a desglosar:

1. La constatación de enfrentarnos a una realidad poliédrica⁴⁴, que nos conduce invariablemente a un punto de partida básico: la necesidad de abordar los contenidos digitales de forma pluridisciplinar y multipolar (ver esquema en el anexo). Abordarlos desde una óptica «reduccionista» será fuente más que probable de problemas.
2. En esencia, lo relacionado con los contenidos digitales no es tan «diferente», hay una persistencia de lo básico: los documentos, la catalogación, la comunicación, etc. Sin embargo, hay nuevos condicionantes a los que es necesario aplicar otros métodos: la gestión por procesos, flujos de trabajo, reglas de negocio, etc.
3. Interiorización en la cultura de la generación y gestión de contenidos digitales del plano de la salvaguarda legal de los derechos individuales de las personas y de las ocasionales responsabilidades jurídicas.

⁴² Como forma de gobierno de los estados caracterizada por atributos como la transparencia, la buena orientación, y que potencia buena parte de su legitimidad ante los ciudadanos.

⁴³ Siempre condicionado, recordémoslo, a las diversas vicisitudes ya comentadas en el capítulo anterior relativo a «Riesgos potenciales» (véase países con censura en el acceso a Internet o a determinados contenidos digitales, o bien los segmentos de población incluidos en la «brecha digital»).

⁴⁴ Considere el lector que, sin ánimo de exhaustividad, se ha hecho referencia a aproximadamente setenta cuestiones y disciplinas de alguna manera relacionadas con los contenidos digitales, y que presentan elevados grados de vinculación e interconexión entre sí.



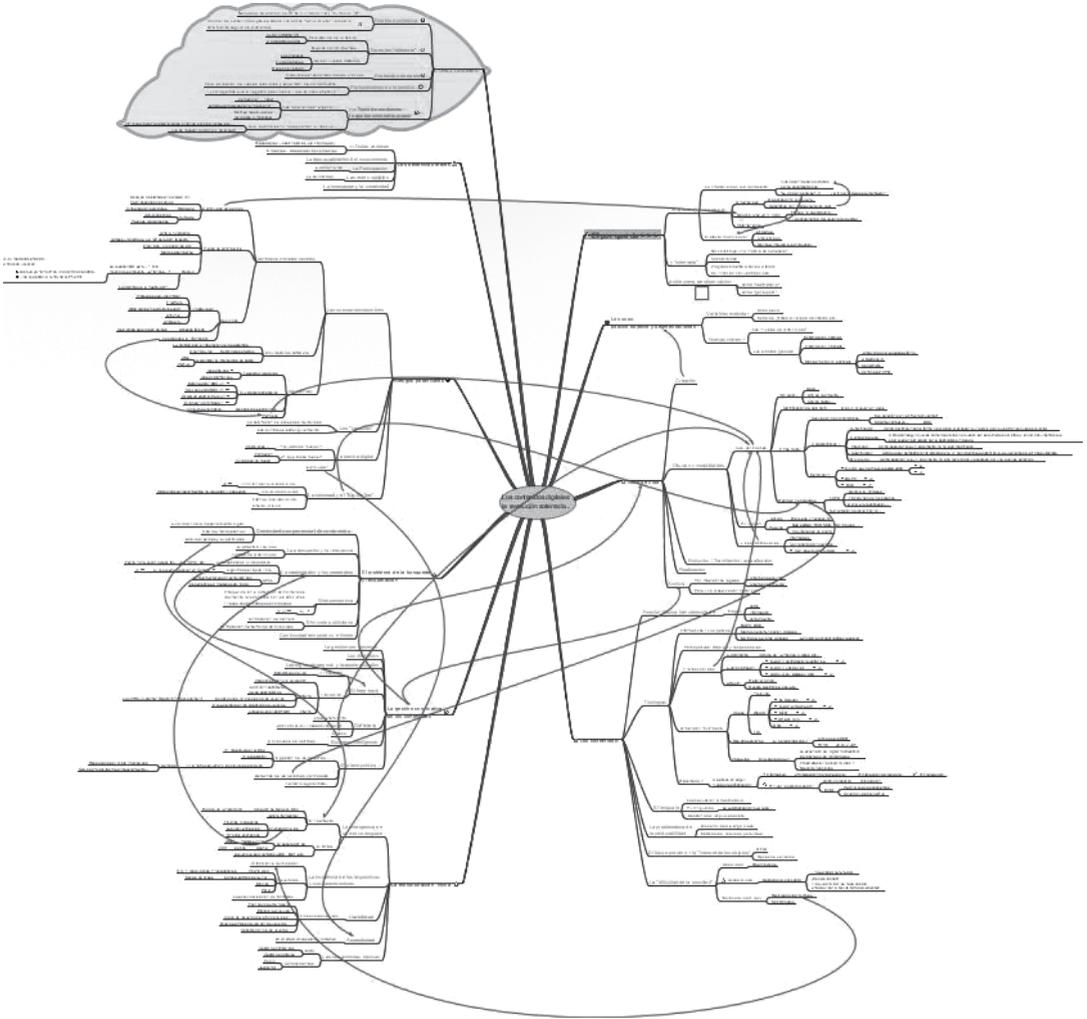
4. Que los árboles no nos escondan el bosque: las «apariencias» engañan..., de «sencillo»... ¡nada! La complejidad radica en el «back end» y puede llegar a ser extremada.
5. Profundicemos en lo positivo: creación de nuevas relaciones y expansión del conocimiento..., y consigamos que lo negativo resulte poco atractivo.

En resumen: si bien no estamos ante el fin de la Galaxia Gutenberg —como tampoco lo fue la televisión para la radio—, sí que nos enfrentamos a un profundo cambio de paradigma para el que es imprescindible una clara toma de conciencia integral de lo que tenemos entre manos.

Epílogo

A la vista de la maraña de ítems existentes en torno a los contenidos digitales y sus vinculaciones (sólo parcialmente representadas en el anexo), bien puede merecer la pena volver a leer y reinterpretar la segunda parte del título de la presente ponencia: «...la revolución soterrada...»

ANEXO I: MAPA



ANEXO II: ÁREAS DE VIGILANCIA Y SEGUIMIENTO

Accesibilidad	RFID
WCAG	Legalidad
Buscadores:	Protección de datos - dcho. intimidad
externos: posicionamiento	Propiedad intelectual
propios: parametrización	Gestión de derechos: CPL, GNU,
indexadores	Creative Commons,
Canales	Lenguaje
Multiplicidad de formatos	Adecuación
Reutilización / sinergias	Niveles de profundidad
Catalogación - taxonomías - metadatos	Plurilingüismo
Conservación (vs. preservación)	Modalidades de difusión
Backups	Push
Dispositivos	Pull
Ordenadores portátiles, Notebooks,	Notificaciones
PC sobremesa	Política de dominios y subdominios
Móviles última generación	Publicidad: contratación de
e-Book, e-Ink	Banners
Estadísticas de uso	Palabras de búsqueda
Páginas vistas, etc.	Públicos objetivos y segmentaciones
Cookies	Identidades
Inteligencia de Negocio (Business	Firma digital
Intelligence): análisis de tendencias,	Retorno de información - retroalimentación -
rutas de navegación, etc.	Feed-back:
Estructuración de la información:	CRM
Tipología contenidos	Estadísticas
Títulos-antetítulos-entradillas-ladillos-	Sistemas de votación
sinopsis	Encuestas
Dimensión multimedia	Seguridad
Firma digital	Firewalls
Funciones básicas: distribución de	Antivirus
Datos - información - conocimiento	Antispam
Entertainment	Perfiles de usuarios: privilegios y
Gestión por procesos - BPM Business	mantenimiento
Process	Servidores:
Management	Servicios web
Ciclo de vida	Servicios de correo
EFQM	Aplicaciones
Flujos de trabajo	Repositorios - Dataware
Reglas de negocio - Business rules o la	Sofisticación de los portales: grados
administración de los procesos del	Software:
negocio	Lenguajes informáticos: HTML, XML,
Herramientas: ECM	CSS, J2EE, .NET, etc.
Gestor de portales	Estándares abiertos
Gestor de contenidos	Hipertexto
Inteligencia de negocio - Business	Segregación componentes
Intelligence	Gestión de errores (calidad)
Internet de las cosas - «Internet of Things»	Plug-in

Javascripts	Usabilidad
Open source - software libre	Vídeo
Suscripciones	Streaming on-line, off-line,
Listas de distribución	Formatos, codecs, etc.
Difusión selectiva de información	Web 2.0 y redes sociales:
RSS Sindicación de contenidos	FaceBook, Twitter, Blogs, etc.
Telecomunicaciones	Foros
Protocolos: HTTP, IP, RTSP,	Wikis
Ancho de banda	Web semántica
Wifi	Morfología / lematización
Telefonía 3G	Semántica
Transacciones:	Sintaxis
e-Commerce	Ontologías / tautologías
e-Government	W3C

II

Euskal Herriko Dokumentazio Zerbitzuak: begirada bat

Teresa Agirreazaldegi, Idoia Camacho, Carmen Peñafiel eta Milagros Ronco

UPV/EHuko Kazetaritza Saileko irakasleak
Profesoras del Departamento de Periodismo de la UPV/EHU

Laburpena: Euren adierazpide guztietan (orokorrak, espezializatuak, publikoak, pribatuak, barrukoak, etab.), bibliotekak eta dokumentazio zerbitzuak, informazio- eta ezagutza-gizarterako oinarritzko azpiegiturak dira. Euskal Herrian mota horretako zer zerbitzu dagoen aztertzen du ikerketak, horien ezaugarriak, baliabide material eta pertsonalak analizatuz, eta ondare, egindako lan, eskainitako zerbitzu eta erabiltzaileen inguruko datuak eskainiz.

Resumen: Servicios de documentación del País Vasco: una mirada. En sus diferentes manifestaciones (generalistas, especializados, públicos, privados, internos, etc.), las bibliotecas y servicios de documentación son infraestructuras y servicios claves en la sociedad de la información y del conocimiento. La investigación examina la implantación de estos servicios en el País Vasco, analizando sus características, recursos materiales y personales, y proporcionando datos sobre fondos, servicios ofertados y usuarios, entre otros.

SARRERA

Informazio gutxi dago Euskal Herriko dokumentazio zerbitzuen inguruan. Nekez aipa daitezke jatorri profesionaleko txostenak, ikerketa sistematikoa urria da eta zaharkitua geratu da, azkar aldatzen ari den esparrua baitugu. Hutsune hori bete nahian, ikerketa proiektu zabal bat egiten ari gara. Gure helburua, alde bate-tik, Euskal Herriko Dokumentazioaren sektorea analizatzea eta ezagutzea da eta, bestetik, dokumentazio zerbitzuen direktorio publiko bat osatzea. Egitasmoa Euskal Herriko Unibertsitateak finantzatu du.

Ikerketa hau deskriptiboa izateaz gain, ebaluatiboa eta prospektiboa ere bada. Bi ataletan banatu dugu, lehenengo atalean komunikabideetako dokumentazio zerbitzuak aztertu ditugu, eta lortutako emaitzak hainbat argitalpenetan zabaldu dira jada (1) (2) (3); bigarren atalean, gainontzeko azpiegitura eta dokumentazio zerbitzuak analizatu ditugu. Ikerketaren datu adierazgarrienak azaltzen saiatuko gara hurrengo lerroetan.

Testu honi azken ukituak ematen ari ginela, gure ikerketan eragina duten bi albisteren berri jakin genuen. Lehenak zioen Espainiako Biblioteka Nazionalak bere *Biblioteka Espainiarren Direktoria* webgune berrian integratu duela. Direktorioaren arduradunarekin harremanetan jarrita, ekimena aspaldikoa dela jakin dugu,

1976koa hain zuzen ere, baina ikusgarritasuna datu basearen kokapen berriak eman dio, dudarik gabe. Bigarren albistea, Kataluniatik etorri da, eta dio Generalitateak Kataluniako liburutegi espezializatuen datu basea jarri duela bere webgunean, eta kontsultatu ahal dela. Aurreko albisteez egiaztatzen digute gure ikerketak baduela gaurkotasuna. Ezagutza zabaltzeaz gain, gaiak bere garrantzia baduela ohartarazteko eta direktorioaren beharra sentiarazteko balio badu, bere funtzioa ondo beteta egongo da.

AURREKARIAK

Direktorioak

Lehenago aipatu dugun moduan, Euskal Herriko dokumentazio zerbitzuen inguruan dauden ikerketak urriak dira; guztien artean aipagarriena *La Documentación en la Comunidad Autónoma del País Vasco (Dokumentazioa Euskal Autonomia Erkidegoan)* egitasmoa da. Proiektua 1993tik 1995era bitartean egin zen, Artxibo, Liburutegi eta Dokumentazio Zentroetako Profesionalen Elkarteak (ALDEE) eman zuen beka bati esker. Galdetegiak erabilita, EAeko 91 dokumentazio zentrotan datu bilduma osatu zuten ikerlariek, eta zentro bakoitzean datu hauek jaso zituzten: izena, helbidea eta landutako gaiak. Direktorioa ez zen inoiz argitaratu; ikerketaren emaitzak, ordea, Fesabid-en kongresuan (Artxibistika, Bibliotekonomia, Dokumentazio eta Museistakako Espainiako Federazioa) aurkeztu ziren (4), baita ALDEE elkartearen buletinaren ere (5).

Direktorioei dagokienez, aipagarriak dira 80ko hamarkadan Eusko Ikaskuntzak osatu eta publikatu zituen Euskal Herriko artxiboen (6) eta liburutegien zentsuak (7).

Bere aldetik, 1995ean, Espainako Biblioteka Nazionalak (BNE) bere *Directorio de Bibliotecas Españolas (Espainiako Biblioteken Direktorioa)* jarri zuen Interneten, Ariane OPAC-en (*online public access catalog*). Egitasmoa ez zen berria; 1976tik, Kultura Ministerioarekin batera, horren aurrekoa osatu zuen: *Espainiako Liburutegien Gida*. Gidaren kontsulta hasieran paperean eta gero ordenagailuetan egin zen, PIC egitasmoaren barruan (Informazio Kultureko Guneak). Garai hartan, *Aldizkako Argitalpenen Katalogoa* egiteko helburuarekin kontaktatu zuten liburutegiekin; eskola liburutegien zerrenda, ordea, Hezkuntza Ministeriotik jaso zuten. Gaur egun, ildo beretik dihardute *Prentsa Hemeroteken Bildumak* osatuz. 2008an, datu basea aplikazio berrian instalatu zuten, eta 2009an, webgune berrian (<http://www1.bne.es/DIBI-ConsultaWEB/inicio.jsp>).

Direktorioan 10.535 erregistro daude (2009-09-23), eta aldiro kontaktatzen dute liburutegi guztiekin datuak eguneratzeko; laster, ordea, zerbitzu bakoitzak ahal izango du Interneten zuzenean bere datuak gehitu eta aldatu (8). Arabako 224 erregistro daude, Gipuzkoako 362, Bizkaiko 349 eta Nafarroako 159. Lurralde guztiak kontuan hartuta, liburutegi hauek ditugu: eskolakoak 282, espezializatuak 306, orokorrak/publikoak 454, nazionalak 1 eta unibertsitatekoak 51. Bakoitzetik hurrengo datu hauek jasotzen ditu direktorioak:

- Bibliotekaren izena.
- Biblioteka mota.
- Sortu zen urtea.
- Erakundea.
- Zerbitzu mota.
- Helbide osoa.
- Telefonoa.
- Posta elektronikoa.
- Gaiak.
- Funtsak.
- Zerbitzuak.
- Harrera.
- Ordutegia.
- Datuen eguneratze data.

Bilaketetan hainbat irizpide erabil daitezke, eragile boolearrekin konbinatuta (*eta, edo, izan ezik*):

- Bibliotekaren izena.
- Helbidea: herria, lurraldea, elkarte autonomoa (zerrenda), herrialdea (zerrenda).
- Gaiak (700 baino gehiago agertzen dira zerrendan, bibliotekek proposatuta).
- Funtsak (40 mota).
- Funtsen mendea —hemerografikoak barne— (zerrenda).
- Zerbitzua (60 baino gehiago zerrendan).
- Biblioteken arteko mailegua.
- Biblioteka mota: guztiak, eskolako, espezializatu, orokorra/publikoa, nazionala/erregionala/elkarte autonomikoena, unibertsitarioa.
- Mendekotasun mota: guztiak, administrazio autonomikoa, administrazio lokal, administrazio publiko zentrala, herrialdeko administrazioa edo uhartekoa, partzuergoa, entitate pribatua, eliza, besteak).
- Funts hemerografikoak.

Sistemak estatistikak eskaintzen ditu, irizpide hauen arabera: lurraldeko banaketa, liburutegi mota, mendekotasunaren arabera eta harrera mota. Automatikoki osatutako estatistika horietaz aparte, BNEk ez du datuekin inongo azterketarik egin.

Dokumentazio zerbitzuak, liburutegiak eta artxiboak ezagutzeko, lanbide elkarrekin bide bat eskaintzen dute. Asnabi Nafarroako Liburuzainen Elkarteak Nafarroako liburutegiaren errolda du bere webgunean (<http://www.asnabi.com/bibliotecas-denavarra>), liburutegi bakoitzaren kontaktu datuekin; ez dago, ordea, dokumentazio zerbitzuen erroldarik, ezta artxiboarik ere. Frantziako ADBS elkarteak (Informazio eta Dokumentazioaren Profesionalek Elkarteak) dokumentazio zentroen urtekaria dauka Interneten, 400 helbiderekin (<http://www.adbs.fr/>), baina horietatik bi besterik ez dira Euskal Herrikoak. Horietaz gain, hainbat helbide lortzeko,

ALDEE, Sedic (Dokumentazio eta Informazioaren Espainiako Elkarte) eta Exit (informazioaren tratamenduan adituen direktorioa) erabili ditugu.

Katalogo kolektiboak

Direktorioek liburutegi eta dokumentazio zentroen datu orokorrak eskaintzen dituzte; katalogoek, ordea, haien funtsen berri zehatza ematen dute. Euskal Herrian hainbat katalogo kolektibo ditugu. Eusko Jaurlaritzaren webgunean bi katalogo kontsulta daitezke: *EIPS* eta *Bilgunea*. *EIPS*, *Euskadiko Irakurketa Publikoko Sarea*, EAEko udal liburutegietako *online* katalogoa da, eta *Bilgunea* Euskal Autonomia Erkidegoko hainbat erakunde publiko eta pribatutako bibliografia funtsek osatutako katalogo kolektiboa da. Azken horretan datuak urtean hainbatean kargatzen dira, eta koordinazioa Eusko Jaurlaritzaren Kultura Sailari dagokio. *Bilgunean* parte hartzen duten erakundeen artean Legezko Gordailua, Eresbil, EAEko ehun udal liburutegitik gora, Eusko Jaurlaritzako sailtako liburutegiak, Euskal Herriko Unibertsitatea, Deustuko Unibertsitatea, Mondragon Unibertsitatea, Sancho el Sabio, Arabako Foru Aldundia, Labayru Ikastegia eta *Ondare bibliografikoa* (Loiolako Santutegia, Gasteizko Seminarioa...) ditugu.

Nafarroako liburutegien funtsak ezagutzeko *Nafarroako liburutegi publikoen katalogo kolektiboa* eta *Nafarroako ondare bibliografikoaren katalogo kolektiboa* ditugu (inprimatutako agirien erregistro katalogafikoak, inprimategiaren hasieratik 1910. urtera artekoak), biak Nafarroako Gobernuaren webgunean kontsulta daitezke. Bestalde, Nafarroako Liburutegi Orokorra 1940tik lurraldeko ondare bibliografikoa gordetzen eta zabaltzen ari da, Liburutegi Publikoen Sistemaren koordinazio gunearen lana betetzen du, eta Legezko Gordailua kudeatzen du. 2005etik, Nafarroako Biblioteka eta Filmoteka 20.000 metro koadroko eraikuntza berri batean bildu dituzte.

Bestalde, *Frantziako Katalogo Kolektiboaren* webguneak *Biblioteka eta ondare dokumentalen erreperitorio nazionala* jasotzen du (<http://ccfr.bnf.fr/>), non 4.800 biblioteka eta 1.700 funtsen informazioa ematen den; horietatik hamalau helbide Euskal Herrikoak dira.

Administrazioa eta legedia

Azpimarratu beharra dago Euskal Autonomia Erkidegoan botere publikoen aldetik inizatiba eta lidergo falta dagoela arlo honetan. 1979tik Gobernuko organoetatik datorren legedi bibliotekarioa eskasa eta garrantzi gutxikoa da, ez dago bidea markatuko duen biblioteka lege bat, Espainiako Estatuko toki ia guztietan dagoen lez (9). 7/1990 Legeak, uztaillaren 3koak, Euskal Kultur Ondareari buruzkoak, udal liburutegiak soilik arautzen zituen, kanpoan utziz EAEko zentro inportanteenak, foru aldundietako liburutegiak eta biblioteka eta dokumentazio zerbitzu espezializatuak eta pribatuak. 11/2007 Legeak ere, urriaren 26koak,

Euskadiko Liburutegiei buruzkoak, kanpoan uzten ditu zentro horiek, baina Irakurketa Publikoaren Sarean sartzeko aukera ematen die, «beren zerbitzuak herri-tar orori zabaltzen badiete».

Profesionalek sarritan salatu dute erregulazioa falta dela (10) (11) (12), baita sektorearen lidergoa aurrera eramango duten erakunde enblematikoak ere. Joana Albret Bibliotekonomia Mintegiak 1998an jada honela adierazten zituen eremuaren egoera eta beharrak:

- a) Euskal Herriko biblioteka, dokumentazio zentro eta artxiboen egoera, oro har, kezagarria da: arazo eta gabezia handiak daude. Bereziki euskararen estatusari eta erabilerari dagokienez, egoera biziki txarra da.
- b) Alor honen arazo eta gabeziei aurre egiteko politika aktiboago eta eragin-korrago bat eskatzen dugu. Orobat, euskal erkidegoetan horretarako legeak egitea.
- c) Euskal Biblioteka Nazional bat eskatzen dugu, alegia, Euskal Herri osoko biblioteken arteko lotura eta koordinazioa bideratuko lukeen biblioteka. Herri agintariei eskatzen diegu lehenbailehen beharrezkoak diren elkarrizketak eta gestioak egin ditzaten proiektu historiko hori abiarazteko.
- d) Bereziki, euskarak pairatzen duen bazterketari konponbidea emateak lehen-tasun osoa behar luke izan. Euskal Agentzia Katalogafiko bat eskatzen dugu alor honetako arazoak bideratzeko (13).

Azken hogeitaz urteotan, Estatuko autonomia erkidego gehienek liburutegiei buruzko legeak onartu dituzte, baita legedi horien barruan liburutegi nagusiak sortu ere: Andaluzia, Aragoi, Asturias, Madril, Murtzia, Gaztela eta Leon, Valentzia... Bakoitzak bere lurraldean, Biblioteca Nacional delakoak estatu mailan dituen funtzioen antzekoak ditu. Kataluniako Liburutegi Legea (bigarrena, 1993koa) osatuena eta pentsatuena da Estatuko gainerakoek artean.

Nafarroan Biblioteca General de Navarra deiturikoa izango da liburutegi nagusia, eta Instituto Navarro de Bibliografía sortu dute liburutegi horren itzalean, Bibliografía Navarra ekoizteko xedearekin. Bitartean, Euskal Autonomia Erkidegoan aipatutako 11/2007 Legeak dioena noiz beteko itxaron beharko da, laugarren xedapen gehigarrian Liburutegien Legea indarrean sartu eta gehienez ere 18 hilabetera Euskadiko Liburutegia eratuko dela adierazten baitu. Baina oraindik ez da eratu.

IFLaren eskaerari erantzunez, J. Lor-ek eman zuen argitara 1997an bere «Guidelines for legislation for national services». Bertan ordura arteko definizioak aztertu zituen eta hiru dimentsio bereizi zituen biblioteka nazionalaren kontzeptuan: ondarea (biblioteka zahar edo klasikoek euren herrialdeetako altxor literarioei ematen diete garrantzia), azpiegitura (koordinazioa, lidergoa eta beste bibliotekei emandako zerbitzuak azpimarratzen dira) eta zerbitzu nazional hedatuak (batez ere garatu gabeko herrialdeetan, bertako azken erabiltzaileei emandako zerbitzuak nabarmentzen dira gehien) (14) (15).

Ikusten denez, Liburutegi nazionala da kultur komunitate baten erreferentzia bibliotekario nagusia bertako ondare bibliografikoaren bilketa eta kontserbazioa bermatzeko. Horrekin lotuta, hainbat zeregin ditu, nazioarteko erakundeek (Unesco, IFLA...) funtsezkotzat edo komenigarritzat jo dituztenak; zeregin horietan garrantzitsuenak lurraldeko ekoizpen bibliografikoa jasotzea, bibliografia eta katalogo kolektiboak elaboratzea, eta sistema bibliotekarioaren buru eta koordinatzaile izatea dira (16) (17). Liburutegi Nazionala ez izatea muga izugarria da herri baten garapen kulturalerako.

IKERKETA

Ikerketa bi zatitan antolatu genuen: komunikabideetako dokumentazio zerbitzuak 2008an aztertu genituen, eta gainontzeko dokumentazio egiturak, 2009an. Komunikabideetako dokumentazio zerbitzuen ikerketa esparrua mugatua, nahiko homogeneoa eta ikerketa taldearentzat gertukoa denez, gaia sakontasun handiarekin ikertu ahal izan genuen. Bigarren atala, ordea, oso esparru zabala eta heterogeneoa da: zentroak oso egitura desberdinekoak dira, sakabanatuta daude eta, barne egiturak direnean, ez dira ezagunak erakundetik kanpo. Euskal Herrian Bibliotekonomia eta Dokumentazio ikasketa propioak ez egoteak, ezta profesionalen elkarrik ere (18), mugatu egiten du unibertso hori ezagutzea. Zailtasunak zailtasun, ikerketan lortutako datuak lagungarriak izango zaizkigu dokumentazioaren unibertso askotarikoa eta zabala hobeto ezagutzeko.

Ikerketako gaia

Euskal Herriko dokumentazio zerbitzuak aztertu nahi izan ditugu, baina zer dira *dokumentazio zerbitzuak*? Zerk eragiten du dokumentazio zerbitzu izatea? Ondareak? Funtzioak? Profesionalek? Informazio eta Dokumentazio Zientzietan erabiltzen den terminologia aztertuz, honela definitu dugu ikerketako gaia:

Dokumentazio zerbitzuak erabiltzaile talde baten behar informatiboak betetzeko eratzen diren informazio sistemak eta egiturak dira.

Definizio horretan funtzioak dira ardatza, baita aitorten soziala eta profesionala ere. Historikoki egitura horiek hainbat izen izan dituzte: artxibo, biblioteka, dokumentazio zentro, informazio zerbitzu... Bibliotekak jakintzaren gotorleku izan dira historian zehar; xx. mendean, irakurketa publikoaren kontzeptuari lotuta agertu dira, hau da, herritarrok kulturaz jabetzeko daukagun eskubideari lotuta. Dokumentazio zerbitzuak, berriz, informazio zientifiko eta teknikoarekin lotuta daude sortzetik bertatik, hau da, xx. mendearen hastapenetatik, edo erabiltzaile talde zehatzekin, gehienetan ezagutza esparru espezializatuetakoak. Halere, liburutegien eta dokumentazio zerbitzuen arteko mugak ez dira beti oso garbiak izaten; batzuetan, liburutegian bertan kanpokoko baliabi-

deei sarrera eskaintzen duen zerbitzua da dokumentazioa; hala, dokumentazioaz baliatuz, liburutegiaren funtsak iristen ez diren tokietara iristeko. Artxiboek dokumentu administratiboak jasotzen dituzte, hau da, erakunde, enpresa edota pertsona pribatuek beren eginkizunetan sortutakoak, denborarekin historikoak bilaka daitezkeenak. Edonola ere, ikerketak honako hau utzi du agerian: zerbitzu baten izendapena nahiko arbitrarioa izan daiteke, eta horretan tradizioak eragin handia izan dezake.

Unibertsoa

Salbuespenak salbuespen ikerketan ez ditugu aztertu EAEko artxiboak, horien datuak eta funtsen deskripzioak Irargi Euskadiko Dokumentu Ondarearen Zentroaren webgunean jasota daudelako jada (<http://www.snae.org/>); *Artxiboen Errolda* atalean, hain zuzen. Udal liburutegiak ere ez ditugu kontuan hartu ikerketan, hiriburuetakoak salbu. EAEko udal liburutegien helbideak eta katalogoak *Euskadiko irakurketa publikoko sarean* kontsulta daitezke (19). Era berean, Nafarroako udal liburutegien funtsak ezagutzeko *Nafarroako liburutegi publikoen katalogo kolektiboa* dugu, eta liburutegi horien datuak webgune berean daude (20). Bestalde, Frantziako Katalogo Kolektiboaren webgunean kontsulta daitezke funtsak, baita *Biblioteca eta ondare dokumentalen errepertorioa*, eta Iparraldekoak Pirinio Atlantikoetako Departamenduan aurkituko ditugu (21). Sare baten barruan dauden informazio zerbitzuak ere ikerketatik kanpo geratu dira; aztertu ditugu, ordea, horien bulego zentralak. Eskola eta institutu etako liburutegiak ere ez ditugu aztertu. Zentro horiek guztiek berezko ezaugarri asko dituzte, eta merezi dute aparteko beste ikerketa batzuetarako gaia izatea.

Datu bilketa

Ikerketa egiteko, galdetegiaren metodologia erabili dugu, bi helbururekin: direktorioa osatzeko datuak jasotzea eta esparrua ikertzeko behar beste informazio bil-tzea. Galdetegiaren datuak bost ataletan sailkatu ditugu:

1. Enpresa edo erakundearen datuak:

- Izena.
- Lan arloa.
- Sarea (publikoa/pribatua).
- Webgunea.

2. Dokumentazio zerbitzuaren datuak.

- Izena.

— Helbidea:

- Zerbitzuaren webgunea.
- Telefonoa.
- Faxe.
- Harremanetarako pertsona.
- Posta elektronikoa.
- Sarrera eta erabiltzaileak.
- Ordutegia.

— Sortu zen eguna.

— Zentroaren historia laburra.

— Atxikipen hierarkikoa.

— Kontsulta kopurua.

— Langile kopurua.

— Langileen ezaugarriak: titulazio akademikoa eta generoaren arabeko banaketa.

— Lanpostuak: kopurua eta generoaren arabeko banaketa.

3. Funtsak, ekipamendua, zerbitzuak eta ekoizpenak.

— Zerbitzuak.

— Ekipamendua.

— Funtsak: gaiak, lurralde eremuak, denbora eremuak, ezaugarriak eta hizkuntza.

— Ekoizpenak: datu baseak, gidak, inbentarioak eta gainontzeko ekoizpenak (ezaugarriak eta hizkuntza).

4. Ematen dituen beste zerbitzu batzuk, hau da, galdetegian beren tokia aurkitu ez duten egitasmo interesgarriak.

5. Dokumentazio zerbitzuaren egitasmoak etorkizunerako.

Dokumentazio zerbitzuen sailkapena errazteko asmoz, jarduera sektoreen zereztrenda osatu eta galdetegiari gehitu genion. Honako jarduera sektore hauek proposatzen genituen:

— Administrazio publikoak.

— Aholkularitza, ingeniaritza eta enpresei zuzendutako mota horretako zerbitzuak.

— Aisialdia eta ikuskizunak.

— Defentsa, ordena publikoa eta segurtasuna.

— Ekonomia eta ogasuna.

— Enplegua, sindikatuak eta elkarte patronalak.

— Gizarte lana.

— Hezkuntza, unibertsitateak eta ikerketa.

— Industria (zehazteko eskatu zitzaien).

— Informatika eta teknologia berriak.

- Ingurugiroa, energia eta lurreko zientziak.
- Kirola.
- Komunikabideak.
- Kultura eta ondarea.
- Lanbide elkarteak.
- Merkataritza eta kontsumoa.
- Nazioarteko harremanak eta garapenerako kooperazioa.
- Nekazaritza, abeltzaintza eta arrantza.
- Osasuna
- Turismoa.

Erantzunak

Galdetegia banatzeko, 233 helbide dituen datu base bat osatu genuen; ondoren, 2008ko urrian hasi ginen galdetegia bidaltzen posta elektronikoaren bitartez, euskaraz zein gaztelaniaz edota frantsesez, aurkezpen gutun batekin batera. Erantzun ez zuten zentroetan beste bi saio egin genituen 2009ko lehenengo lau hilabeteetan (lehenengo, posta arruntaren bidez, eta gero, berriro ere, posta elektronikoz). Hainbat kasutan barruko zerbitzua delako ez dugu erantzunik jaso. Horrela adierazi digute informazioa publiko egiteari uko egin dioten arduradun batzuek.

Galdetegien erantzun maila %37koa izan da; antzeko beste ikerketa batzuekin alderatuta, nahiko ehuneko altua dela esan daiteke. Dokumentazio zentroa duten komunikabide guztiek erantzun dute, hau da, 15ek; eta gainontzeko zentroetatik 72 erantzun jaso ditugu. Guztira, gure ikerketa egiteko, 87 dokumentazio zerbitzuren datuak izan ditugu.

Datuak gainbegiratu, normalizatu, sailkatu, etiketen bitartez indexatu eta itzuli ditugu, gero Interneten dagoen direktorioan jartzeko, *Gidadok* webgunean (<http://www.gidadok.com/>). Helburua informazio hori jendeari oro har eskaintzea izan da.

DATUEN AZTERKETA KUANTITATIBOA

1. Dokumentazio zentroak, lurraldeka

Zentroak honela daude banatuta lurraldeka: 16 Araban (guztiak Gasteizen), 40 Bizkaian (horietatik 27 Bilbon), 21 Gipuzkoan (horietatik 13 Donostian), 5 Iparraldean (horietatik 2 Baionan) eta 5 Nafarroa Garaian (horietatik 4 Iruñean). Erantzun gehien Arabatik jaso dugu, lurralde horretan bidalitako galdetegien %56 erantzun baitituzte; beste ertzean Nafarroa dugu, non bidalitako %7 besterik ez dugun jaso. Erdian geratzen dira Bizkaia (%48), Gipuzkoa (%32) eta Iparraldea (%29).

2. Sortutako urtea

Zazpi zentro sortu ziren 1960. urtea baino lehenago: Ignacio Aldekoa kultur etxea, XIX. mendean (ez dute zehaztu urtea); Bizkaiko Abokatuen Elkargoko Biblioteka, 1838an; Bizkaiko Foru Aldundiko Liburutegia, 1907an; Azkue Biblioteka, 1919an; Baionako Euskal Museoa, 1924an; Radio Vitoria, 1934an, eta Koldo Mitxelena Biblioteka, 1939an. Zentro gehienak, ordea, 1976tik aurrera sortuak dira; garai emankorrena 1986tik 1990era izan zen; izan ere, urte horietan 17 zerbitzu osatu ziren. Euskal Herriko zentroetan garai horretan dokumentazioaren informatizazioa eta datu baseen hedapena gertatu zen. Aztertutako zentro berriena 2006koa da.

3. Egituraren izena

Ikertu ditugun egiturak izendatzeko gehien erabiltzen den terminoa *dokumentazio* hitza da: 47 agerraldi; horietatik 30ek *dokumentazio zentro* dute izena. Ondoren, *biblioteka* terminoa dator: 27 agerraldi; zentro horietatik 16k dute *biblioteka* izena. Hirugarren lekuan *artxibo* terminoa dugu: 13 agerraldirekin; eta 5 zentrok dute *artxibo* izena. *Zerbitzu* terminoa 9 bider agertzen da. Aurreko terminoak honela konbinatzen dira euren artean: *biblioteka* eta *dokumentazio* hitzak 11 bider agertzen dira batera; *artxibo* eta *biblioteka*, behin; *artxibo*, *biblioteka* eta *dokumentazio*, 4 bider; *informazio* eta *dokumentazio*, 3 bider, eta beste horrenbeste *artxibo* eta *dokumentazio*. Ikus dezagun koadroan:

Terminoak zentroen izenetan	
47 Dokumentazio	30 Dokumentazio Zentro
27 Biblioteka	16 Biblioteka
13 Arxibo	5 Arxibo
4 Fonoteka	
Terminoen konbinaketa	
11 Biblioteka eta Dokumentazio	
4 Arxibo, Biblioteka eta Dokumentazio	
3 Informazio eta Dokumentazio	
3 Arxibo eta Dokumentazio	
1 Arxibo eta Biblioteka	

Izendatzeko moduak asko dira, eta batzuetan betebeharrak ere halaxe izaten dira. Honela azaldu du arduradun batek kontu hori: «Zerbitzua *artxibo* izenarekin izendatu da beti, nire kontratuan nire kargua *liburuzain* hitzarekin ageri da, baina egin pixka bat denetik egiten dut, artxiboko lana, liburuzain, dokumentazioa...» (22).

8. Profesionalen ikasketak

Galdetegietan 263 langileren heziketaren datuak eskaini dizkigute:

25	Graduatu-ondokoa (%10)
142	Lizentziadunak (%54)
29	Diplomadunak (%11)
20	Teknikari espezialistak (%8)
15	Teknikari laguntzaileak (%6)
28	Batxilergoa (%11)
4	Beste batzuk (%2)

Ikusten denez, oso maila altuko langileak dira: %75ek unibertsitate ikasketak dituzte. Ikus dezagun, baina, banaketa hori generoaren arabera:

- Emakumeen artean, 18 graduatu-ondokoak dira, 100 lizentziadunak, 20 diplomadunak, 13 teknikari espezialistak, 11 teknikari laguntzaileak, 14k dute batxilergoa eta 2 *beste batzuk* taldekoak dira, hau da, %76k ikasketa unibertsitarioak dituzte, eta %56 lizentziadunak dira.
- Gizonezkoei dagokienez, 7 graduatu-ondokoak dira, 42 lizentziadunak, 9 diplomadunak, 7 teknikari espezialistak, 4 teknikari laguntzaileak, 14k dute batxilergoa eta 2 *beste batzuk* taldean daude.

Emakumeen nagusitasuna handiagoa da, unibertsitate ikasketak dituztenen artean; gizonezkoena, berriz, teknikari espezialistak eta batxilergoa ikasketa mailetan.

Ikerketaren bigarren atala egin genuenean, iruditu zitzaigun interesgarria izango zela ikasketen maila jasotzeaz gain, ikasketa horien jakintza arloak zehaztea. Hori dela eta, adierazle hori bigarren atalean bildutako datuetan dugu soilik.

Aztertutako 16 graduatu-ondokoetatik 9k, hau da, %56k, Dokumentazioko graduatu ondoko ikasketak egin dituzte; 7k, hau da, %43k, beste lerro batekoak egin dituzte: batek Historia, batek Artea, bik Kooperazioa, batek Psikologia, batek Filologia eta beste batek zehaztu gabekoa.

Lizentziadunen ikasketak honela daude banatuta: %36k Historia ikasi dute; %11k, Dokumentazioa; %10ek, Filologia; %8,6k, Filosofia eta Letrak; %7k, Kazetaritza; %6k, Zuzenbidea edota Soziologia; %3k, Ingeniaritza, Ekonomia, Pedagogia edota Artearen Historia, eta %1ek, Zientzia Politikoak, Biologia edota Kimika.

Diplomadunen ikasketak, berriz, honela daude banatuta: %36 Dokumentazioa, %27 Magisteritza, %18 Lan Harremanak, %9 Historia eta %9 Turismoa.

Datuak erakusten dutenaren arabera, dokumentazio zerbitzuetan lan egiten duten gehienek letretako lizentziatura tradizionalan ikasitakoak dira. Prestakuntza

espezifikoek, ordea, graduatu-ondokoetan eta diplomaturetan dute nagusitasuna. Hori langileen eraberritzearen prozesu baten erakuslea izan daiteke.

9. Lanpostuak

Dokumentazio zentroetako langileek hurrengo kategoriak betetzen dituzte: %19k, Zuzendari-Arduradun; %22k, Dokumentalista; %16k, Liburuzain; %6k, Artxibozain; %18k, Laguntzaile; %9k, Administrari, eta %9k, beste batzuk.

Sexuaren arabera, honela daude banatuta kategoriak:

- Emakumeak: %10, zuzendari-arduradun; %10, dokumentalista; %11, liburuzain; %3, artxibozain; %20, laguntzaile; %8, administrari, eta %5, beste batzuk.
- Gizonezkoak: %5, zuzendari-arduradun; %3, dokumentalista; %3, liburuzain; %1, artxibozain; %11, laguntzaile; %3, administrari, eta %6, beste batzuk. Ehuneko guztiak alderatuta, esan dezakegu gizonezko arduradunena altua dela, baita *beste batzuk* kategoriakoena; azken horiek eginkizun teknikoak dira, edota mandatari lanei dagozkie.

AZTERKETA KUALITATIBOA

Aztertu ditugun egiturak bi multzotan bana daitezke, independenteak direnak —hau da, gure kasuan 5 liburutegi handiak eta 3 artxiboak— eta erakunde baten atalak direnak —gainontzeko guztiak, 79—. Gure ikerketan, erakunde horiek 9 administrazioak, 2 instituzioak, 4 hezkuntza zentroak, 4 unibertsitateak, 5 ospitaleak, 5 elkarteak, 25 institutu eta fundazioak, 3 lanbide elkargoak, 4 museoak, 7 enpresak, 1 sindikatua eta 15 komunikabideak dira.

Gaiaren arabera, ikertutako zentroak honela daude banatuta: 9 Orotarikoak, 18 Kultura, 20 Gizarte Gaiak, 7 Industria eta Teknologia, 15 Albisteak eta 8 Osasuna. Aurreko sailkapenak murrizketa handiak dakartza ezinbestean. Hala, unibertsitateetako liburutegiek atal espezializatuak dituzte, baina sailkatzeko orduan, erakunde osoa kontuan harturik, *orotarikoak* sailean sartu ditugu; beste aldean, irratietako fonotekak *albisteak* kategoriaren barruan daude, erakundeari lehenetsia emanez, baina, zalantzarik gabe, musika grabazioak dira ugariak, eta alde horretatik, *kultura* atalean ere egon zitezkeen.

Zentro independenteak: Liburutegi handiak

Bost liburutegi handik hartu dute parte ikerketan, horietatik lauk 50 urte baino gehiago dituzte: zaharrena XIX. mendekoa da; berriena, berriaz, 2005ekoa (Miarrizteko mediateka). Guztiak sare publikoan edota publiko eta pribatuan daude (Bizkaiko Foru Aldundikoa, kasu). Langile asko dituzte: langile gutxien

duenak 24 ditu, eta gehien duenak, 57, ikasketa eta kategoria askotarikoak. Bil-tzen dituzten funtsak oso handiak dira, eta ehun mila liburu aise gainditzen di-tuzte; hainbat formatutan dituzte horiek, baina liburuek dute nagusitasuna. Sa-rerra publikoa dute, eta oso ordutegi zabala.

Zerbitzu asko eta askotarikoak eskaintzen dituzte: informazio eta orientazio bibliografikoa (bai lekuan bertan, bai modu birtualean), erreferentzia zerbitzua, ka-talogoen kontsulta, dokumentuen kontsulta, dokumentuen mailegua, liburutegien arteko mailegua, Internet bidezko mailegu berritzea, erosketa proposamenak ja-sotzeko sistemak, erreserbak, informazioaren hedapen hautatuaren zerbitzuak eta erabiltzaileen formazioa. Hainbat tokitan, galderak faxez, posta elektronikoz eta te-lefonoz ere jasotzen dituzte.

Datu baseen kontsulta eskaintzen dute, sarean eta euskarri optikoetan. Bil-duma digitalak gero eta ugariagoak dira, prentsa, liburuak eta dokumentu elektro-nikoak bilduta. Entzunezkoak eta ikus-entzunezkoak ere badira, DVDn eta beste formatu batzuetan. Horrez gain, mapak, grabatuak, dokumentu historiko garrantzi-tsuak, bilduma bereziak, inkunableak, eta abar ere aurki ditzakegu. Funtsak batik bat gaztelaniaz daude (edota frantsesez Miarritzeko mediatekan), neurri txikiagoan euskaraz, eta batzuk beste hizkuntza batzuetan.

Ekipamendu zabala dute: ikasketa aretoak, irakurketa eta kontsulta aretoak, erreferentziako zerbitzuak, hemerrotekak, ikertzaileen aretoak, Interneten konsul-tatzeko aukera ematen dute, wifi sareak, kopiagailuak, ordenagailuak (finkoak zein eramangarriak), inprimagailuak, mifrofilm eta mikrofitxen irakurgailuak, bideo eta DVD irakurgailuak eta telebistak.

Gai orokorrez gain, atal espezializatuak ere badituzte, batez ere euskal gaiak, herriko informazioa edota herrialdekoa jasotzeko. Erreferentzia sekzioa dute, baita hemeroteka ere, non eguneko prentsa paperean zein modu digitalizatuan eta *online* zerbitzuak —*Hemeroteka Plus* eta *Iconoce*— eskaintzen dituzten. Talde zehatz batzuentzako atalak ere egoten dira, hala nola itsuentzako eta ikusmen urri-tasunen bat duten pertsonentzako zerbitzuak, umeentzako atalak, *bebeteca*-k, gura-soentzako bibliotekak eta gazteentzakoak.

Irakurketa gidak eta dokumentu berrien gidak osatzen dituzte, eta gai lokalen in-guruko bibliografiak eta aldizkarien hustuketa ere egiten dituzte zenbaitetan. Luzapen bibliotekario edota kulturaleko ekintzak ere eskaintzen dituzte sarri, alegia, ikastaroak, bisita gidatuak, erakusketak, ospakizun aretoak, argitalpenak eta liburuen salmenta.

Etorkizunari dagokionez, ondorengo asmoak adierazi dituzte: webgunea berri-tzea, wifi sareak ezartzea, digitalizazioarekin segitzea, ONA tarjeta erabiltzea, al-dizkarien hustuketa egitea eta erabiltzaileak prestatzea.

Zentro independenteak: Artxiboak

Artxiboek dituzten funtsak bakarrak direla, horixe da horien berezitasuna. Eta funts horiek deskribatzeko, inbentarioak eta katalogoak osatzen dituzte. Era-biltzaileak espezialistak eta ikertzaileak izaten ohi dira, eta artxiboetan bertan bi-

bliotekak eta hemeroteka espezializatuak ere izaten dira, ikerkuntzarako tresna gisa. Bi eta hamar (Eresbil) profesional artean lan egiten dute ikertu ditugun artxiboetan; haiek artxibistikan eta artxiboko gaiaren inguruko adituak dira. Aztertu dugun artxibo zaharrena 1974koa da, eta berriena, 2000. urtekoa; bi pribatuak dira, eta bat, publikoa. Artxibo horien etorkizunerako erronka digitalizazioa dela adierazi digute.

Mendeko zentroak: Museoak

Museoetako zentroek funts espezializatuak dituzte, eta liburuak ez diren beste materialak ugariak dira bertan. Publikoak edota fundazioak dira. Sarrera zabala eskaintzen dute, baita ordutegia ere, baina ez liburutegi handiena bezainbestekoa. Dokumentazio zentroak museoekin batera sortu ziren; bi muturretan San Telmo museoa, 1902an sortua, eta Artium, 2002an sortua, ditugu. Langile kopurua 2-6 lagun tartean dago; denak lizentziadunak dira.

Katalogo bibliografikoak, Interneteko baliabideen gidak, erabiltzaileentzako gidak, dosierak, aldizkarien aurkibideak eta argazkitegiak eskaintzen dituzte. Hainbat ekitaldi ere antolatzen dituzte bertan, zentro espezializatuetakoa bilkurak, irakurketa klubak, erakusketak, tertuliak, liburuen aurkezpenak, lantegiak, hitzaldiak, liburuaren egunak, *bookcrossing*... Bekak ere eskaintzen dituzte. Etorkizunerako, egitasmo haiek dituztela adierazi digute: lankidetzaren profesionala bultzatzea, web 2.0 inguruko zerbitzuak eskaintzea, lantegiak antolatzea, Intraneta garatzea, museoaren artxiboa osatzea eta web orria gaurkotzea.

Ikerketan museo bateko dokumentazio sail bat ere aztertu dugu: Gasteizko Arte Ederretakoa. Kasu horretan, barruko zerbitzua da, eta hurrengo betekizun hauek ditu: museoeko bilduma katalogatzea eta horren dokumentazioa osatzea, bildumako piezen eta egileen inguruko informazioa biltzea, bildumako piezen argazkiak jasotzea, museoaren eta museologiaren inguruko informazioa biltzea, argazki artxiboa kudeatzea, eta egile eta lanen artxibo dokumentala kudeatzea.

Mendeko zentroak: Administrazioa eta erakundeak

Barruko langileentzat osatutako zentroak dira, egoitzarako sarbidea mugatua baita, eta hartzaile guztiei zuzendutako informazioa erakundearen webgunean jarzen dute. Ordutegi zabala dute, lanekoa, eta langile kopurua bi eta zortzi artekoa da. Zerbitzu zaharrena 1983koa da, eta berriena, 2006koa.

Honako zerbitzu hauek eskaintzen dituzte: kontsulta, mailegua, informazioa eta erreferentziak, informazioaren hedapen hautatuak, alertak, aholkularitza bibliografikoa, dokumentuak eskuratzea, buletinak, aholku teknikoak, txosten dokumentalak, teknologikoak eta sektorialak.

Funts hauek dituzte: liburuak, aldizkariak, liburuxkak, multimedia dokumentuak, dokumentu elektronikoak eta datu baseak (DVD, CD eta sarean). Erreposi-

torioak dituzte (*Euskadi.net*). Katalogoak, prentsa datu baseak, aurkibideen buletinak, dokumentu berrien buletinak eta adituen datu baseak osatzen dituzte, eta trebakuntza eta aholkularitza ere eskaintzen dute. Dokumentuen hizkuntza gaztelania da, eta ingelesa eta euskara ere bai, baina neurri txikiagoan.

Etorkizunerako, egitasmo hauek adierazi dituzte: web orriak indartu, informazioa eta dokumentazioa hedatzeko; aldizkako argitalpenen katalogo kolektiboa osatu (elektronikoak barne) Eusko Jaurlaritzako beste liburutegi espezializatuekin; Gazteria *Euskadiko Liburutegi Sistema Nazionalaren* barruan sartu; aldizkariak eta analitikak Absys kudeaketa sistema integralaren barruan sartu.

Aldaera: Eusko Jaurlaritza

Eusko Jaurlaritzako dokumentazio egitura barruko beharrei erantzuteko da, baina ikertzaile eta herritarrentzat ere zabalik dago. 1981ean osatu zuten, eta gaur egun 26 lagunek lan egiten dute bertan: zortzik artxihoan, hamarrek liburutegian eta dokumentazioan, seik argitalpenetan eta bik zuzendaritzan.

Honako funts hauek dituzte: artxihoan, Espainiako Estatuak transferitutako dokumentuak, EAEko Administrazioak sortutakoak eta dohaintzak; bibliotekan, hainbat funts hizkuntza anitzetan; eta horiez gain, euskal prentsaren hemeroteca digitala. Dokumentuen denbora-eremuak hauek dira: artxihoa (1566-), biblioteka (1548-), argitalpenak (1987-). Kontsultaz gain, dokumentuen eskuratzea, liburutegien arteko mailegua, prentsa dosierrak, dossier dokumentalak, dokumentuen erreproduktzioa, liburuen salmenta eta argitalpenen edizioa egiten dute. Datu base hauek osatzen dituzte: aldizkari ofiziala, prentsa, argazkiak, bibliotekako katalogoa. 2004an *Eusko Jaurlaritzaren liburutegietako aldizkako argitalpenen katalogo orokorra* publikatu zuten, argitalpenen katalogoa eta argitalpen berrien buletina, eta horien liburuxka.

Etorkizunerako, egitasmo hauek dituzte: artxiboaren kudeaketa, artxibo digitala eta kudeaketa administratiboa, e-administrazioari zuzenduta; eraikuntza berri bat izatea, robotizatua; biblioteka digitala, web 2.0 eta biltegi instituzionala eraikitzea; eta argitalpenetan liburu eletronikoak argitaratzea.

Aldaera: Eusko Legebiltzarra

1981. urtean osatu zuten, Azterlanen zuzendaritzapean, eta gaur egun 16 lagunek lan egiten dute bertan. Barrurako zerbitzua da, baina ikertzaile akreditatuak ere onartzen dituzte. Liburutegi eta artxiboen funtsak kontsultatu ahal dira, eta edozein motatako informazio eskaerari erantzuten diote. Dokumentuen eskuratzea ahalbidetzen dute, eta horien mailegua, bertakoa zein biblioteken artekoa. Kudeaketarako tresna informatikoak eta datu base espezializatuen sarrera eskaintzen dituzte. Bilduman euskal gaien funts historikoak, bolivartar bildumak, zuzenbide publikoak eta osagarriak dituzte. Bibliotekako katalogoaz gain, datu base hauek eskaintzen dituzte: ekimenen, osoko bilkuren, batzordeen eta buletinen datu ba-

seak. Horrez gain, argitalpen ez-ofizialen katalogoa eta *Infoberriak* (asteroko buletina) osatzen dituzte.

Aldaera: Arabako Biltzar Nagusiak

Dokumentazio eta Informazio Sistemen Zerbitzua 1986an osatu zen, eta bi lagunek lan egiten dute bertan. Hauek dira zerbitzuaren betekizunak: dokumentazio funtsen antolaketa, deskribapena eta hedapena; funts elektronikoen eta artxi-boko funtsen kudeaketa elektronikoa; liburutegi osoaren kudeaketa (eskuratzea, katalogazioa, mantentze lana eta hedapena); aholkularitza tekniko dokumentala; dosierrak; erreproduktzioa eta hedapena; espedienteen digitalizazioaren kudeaketa lanak eta horiek webgunean jartzea; webguneko informazioaren mantentze lan osoa, atal guztietako informazioa, Ganberako organoen saioen digitalizazio sistemen zuzendaritza eta kontrola; eta datu baseen inplementazioa eta kudeaketa (arauak, aktak, biografiak). Horrez gain, Ganberaren informatika zerbitzuen kudeaketa administratiboaren zuzendaritza eta ardura teknikoa dute.

Arabako Gobernuako organoen Akta Historikoak (1502-1877) gordetzen dituzte; gainontzeko funtsak, ordea, 1979tik aurrera sortutakoak dira. Testuak (Legebiltzarreko espedienteak), irudiak, ikus-entzunezkoak (saioen bideoak), entzunezkoak eta elektronikoak (aktak, espedienteak, ikus-entzunezkoak, soinuak) dira. Prentsa dosierrak automatizatuak daude 2003tik. Edukiak argitaratzeko sistema dute, Legebiltzarreko informazioa kudeatzeko eta hedatzeko osoko sistema, eta horren barruan, hurrengo datu baseak: saioen aktak (webgunean), arautegi instituzionala (Intranet), Legebiltzarreko jardunak (webgunean), mozioak, akordioak (webgunean eta Intraneten), funts bibliografikoak, aldizkako publikazioak (Intranet). Horiez gain, biografia (webgunean), argazkia eta prentsa (dosierrak eta albisteak) datu baseak dituzte. Halaber, akta historikoen hustuketa ere egiten ari dira (laburpenak, transkripzioak, irudiak).

Webgunean gaur egungo informazioari 1979tik 1999rako dokumentazio guztia gehitzea da etorkizunerako duten egitasmoa (testua, ikus-entzunezkoa eta grafikoa), baita webguneko kontsulta sistemak aldatzea ere, sistema intuitiboagoak eta sartzeko errazagoak ezartzeko.

Aldarea: Ingurugelak

Bost zentro daude EAEn, ordutegi zabala dute eta, sarritan, langileek dokumentazio gunea kudeatzeaz gain, beste ardura batzuk ere badituzte; esaterako, aholkularitza, proiektuen dinamizazioa eta formakuntza. Material didaktikoa dute, eta funtsetan eta katalogazioan euskara ohikoa da. Informazio buletinak osatzen dituzte, hala nola dokumentu berrien buletin elektronikoa irakasleentzat, nagusiki horiek erabiltzen baitituzte. Ingurumen dokumentazio zentroen sarean daude (RECIDA), eta horrek zenbait kontsultari erantzuten laguntzen die. Urdaibaikoak Eusko Jaurlaritzako katalogoan integartzeko asmoa adierazi du.

Aldaera: Osasun Saila eta Osasun Lurralde-Zuzendaritzak

Sarbide mugatua dute, baina inoiz salbuespena eginda, kanpoko profesionalak ere atenditzen dituzte. 80ko hamarkadan sortu ziren, baina horietako batean langile espezializatuak 1998an hasi ziren lanean; langile kopurua bat eta hiru artekoa da.

Zerbitzu hauek eskaintzen dituzte: mailegua, baita liburutegi artekoa ere; informazio bibliografikoa; bilaketa zientifiko espezializatuak; dokumentuen hautaketa eta eskuratzea; barneko erabiltzaileen prestakuntza espezializatuak; erabiltzaileen edozein informazio eskakizuni erantzutea; artxibo administratiboaren kudeaketa; aholkularitza; ezagutzaren kudeaketako bitartekaritza; aplikazioen prestaketa, eta baliabide elektronikoen aktibazioa. Honela deskribatzen du bere lana saileko arduradunak: «Kudeatzaile, liburuzain, dokumentalista eta artxibozain lanak egiten ditut, eta erabiltzaileen formakuntzaz, eta aplikazioen, tresnen eta sareko eta inprimatutako informazio iturrien inguruko erabakiak hartzeaz ere arduratzen naiz. Bizkaiko, Gipuzkoako eta Arabako Osasun Lurralde-Zuzendaritzekin lan egiten dugu. Webgunearen egitura eta mantentze lana koordinatzen dut. Osakidetza ospitaleetako liburutegiekiko koordinazioa egiten dut».

Funtsak eguneratuta daude; informazioak ez dauka hamar urte baino gehiago. Dokumentuak gaztelaniaz eta ingelesez daude; beste hainbat lan, berriz, Eusko Jaurlaritzak argitaratuak, euskaraz eta gaztelaniaz. Monografiak, aldizkariak, literatura grisa, euskarri berrietako dokumentuak dituzte, eta gero eta gehiago argitalpen elektronikoa eta harpidetza elektronikoa partekatutakoak. Osasun Saileko Liburute-giaren Katalogoak Gipuzkoako eta Bizkaiko lurraldeetako liburutegenak, Osasun Sailekoa eta Erizaintza Eskolakoa biltzen ditu. Saileko hiru liburutegiek partekatzen dituzte datu baseak eta aldizkari elektronikoen harpidetza, gastuak eta edukien sarbideak elkarrekin banatuta.

Etorkizunean ligurutegi digital eguneratuaren kontzeptuarekin segituko dute, eta Espaniako Osasun Sistema Nazionalaren Biblioteka Birtualean parte hartzeko asmoa ere adierazi dute. Era berean, Osakidetza ospitaleetako bibliotekekin lankidetzara osoa izan nahi dute, Osakidetza eta Osasun Saileko baliabide guztien kontsulta bermatuko duen sare instituzional bat eraikita.

Mendeko zentroak: Ospitaleak

Aztertutako guztiak zerbitzu publikoak dira; sarrerari dagokionez, mugatua edota zabala izan daiteke. Lau zentro 70eko hamarkadan sortu ziren; bat, berriz, 1999an. Orduetara oso zabala dute (12 eta 13 ordukoa). Langile kopurua batetik lura doa, baina hiru ospitaletan langile bakarra dago, hiruretan administrari kategoriarekin, nahiz eta horietako bat lizentziaduna izan.

Zerbitzu hauek eskaintzen dituzte: erreferentzia zerbitzua; informazio bilaketak; bilaketa bibliografikoak; erreprografia; kontsulta; mailegua, baita liburutegen artekoa ere; prestakuntza eta aholkularitza; zentro batean Dokumentuak Eskuratzeko Zerbitzua dute (SOD). Horrez gain, erabiltzaileei etengabeko trebakuntza eskaintzen zaie.

Funtsak honako dokumentu hauek osatzen dituzte: liburuak, aldizkariak, tesiak, literatura grisak eta datu baseak. Paperean zein formatu elektronikoan, eta gaztelaniaz, ingelesez eta fantsez daude, batik bat. Gurutzetako ospitaleko funts zaharrenak 1930ekoak dira. Barruko datu baseak osatzen dituzte; esate baterako, liburuen katalogoak (paperez) eta aldizkariaren katalogoak (paperean zein formatu elektronikoan). EAEko osasun arloko dokumentazio zerbitzuek sarean lan egiten dute, eta baliabide elektroniko guztiak erakundeko langileek kontsultatzeko moduan daude Intraneten, Osakidetzako Liburutegi Birtualean. Zentro bateko liburutegia berri egingo dute, baita programa eta funts elektronikoak ere.

Mendeko zentroak: Enpresak

Gure ikerketan parte hartu duten enpresak pribatuak dira, oso puntakoak, eta barneko dokumentazio zerbitzuak dituzte. Zerbitzu zaharrena 1982an sortutakoa da, eta berriena, 2001ean. Laneko ordutegia dute, eta langileen kopurua batetik laura bitartekoa da, gehienek letretako formakuntza dute: 2 graduatu-ondokoak dira (letretan eta medikuntzan); 12, lizentziadunak —8 letretakoak, 3 zientzieta-koak eta 1 teknologietako ikasketetakoak—, eta 3 administrariak dira.

Honako zerbitzu hauek eskaintzen dituzte: informazioaren eta dokumentazioaren bilaketa eta erosketen kudeaketa (artikulu zientifikoak, liburuak, aldizkari zientifikoak, datu baseak, atari espezializatuak), aldizkariaren harpidetza, bilaketa bibliografikoak, kontsulta, liburuen mailegua, bibliografiaren kudeaketa, aldizkariaren indizeen zerbitzua, prentsa laburpena, erakundearen proiektuen artxiboaren kudeaketa, legeen segimendua eta eguneratzea, zaintza eta alerta teknologikoak, prospektiba, industria sektoreen zaintza konpetitiboa, elkartetako kideen inguruko informazioaren kudeaketa, erabiltzaileen formakuntza dokumentazio zerbitzuko informazio baliabidetan; webguneko edukien eta blog korporatiboen kudeaketa. Ikertutako zerbitzu bat artxiboa da, eta erakundearen gestioak sortutako funts administratiboak ere kudeatzen ditu.

Datu base dokumentalen diseinu eta mantentze lana egiten dute (liburuen katalogoa, aldizkariaren katalogoa, arauak, artxiboa, eta abar) eta kanpoko datu baseen sarbidea kudeatzen dute. Ikerketa unitateentzat interesa duen informazioaren segimendua osatzen dute (prentsa, laguntzak, lehiaketak, legeak, ekitaldiak, arautegia...) eta Intraneten kudeaketa eta mantentzea. Hedapen informatiboa buletin bitartez egiten dute: dokumentu berrien buletinak eremu teknologiko bakoitzean (liburu, aldizkari, legeria, albiste, deialdi, agenda eta abar); informazio orriak, tratamendu azkarragoa eta zabalagoa behar duten auzien azken albisteak zabaltzeko (diru laguntzak...) eta sektore bakoitzari buruzko legeen buletinak (ingurugiro alorrekoa, elikagai alorrekoa...), Europako legeak, Espaniakoak, bertakoak...

Funts hauek dituzte: liburuak —formatu elektronikoan ere bai—, aldizkari zientifikoak eta teknikoak —paperean eta sarean; azken horiek gero eta ugariagoak dira—, gehienak nazioartekoak, eta CD-ROM eta DVD batzuk —batik bat, kongresuetako materiala—. Gehiena ingelesez eta gaztelaniaz dago. Zerbitzu hauek funts digitalak dagoeneko gehiago dira, besteak baino. Arduradun batek honela

azaldu du: «Enpresa kudeaketari buruzko 400 bat liburu ditugu, gaztelaniaz, ingelesez eta frantsesez. Paperean ez dugu ezer gehiago. Harpidetza bidez jasotzen ditugun aldizkariak bota egiten ditugu».

Krisi ekonomikoa nabaritu dute, baina ez dokumentazio zerbitzuan. Dokumentazioburu batek honela azaldu du etorkizunerako dituzten egitasmoaren berri: «Nondik oratu krisi honi? Eutsi zerbitzuari, eta, ahal den neurrian, hobetu zerbitzua. *Krisi garaian, hobetu*». Etorkizunerako, egitasmo hauek aipatu dituzte: bertako ikerlarien argitalpen zientifikoaren funtsak digitalizatu; Intraneten erabilera bultzatu, enpresako informazioaren kudeaketarako tresna den aldetik; gaur egungo webgunea aztertu; web plataforma bat osatu, zerbitzuak barneko bezeroei eta fundazioko kideei eskaintzeko; eta web 2.0ko tresnak sustatu enpresa barruan.

Gainontzeko zerbitzuekin alderatuta, honako elementu hauek azpimarratuko genituzke: dokumentuen eta informazioaren tratamendu sakona eta azkarra; hedapenaren garrantzia, alerten, zaintza teknologikoaren eta ekoizpen sektorialen bitartez; digitalizazioaren zabalkuntza; Intraneten garapena, komunikazio tresna modura, eta dokumentazio zerbitzuen parte-hartzea enpresako webgunearen garapen eta mantentze lanetan.

Mendeko zentroak: Unibertsitateak

Hiru unibertsitatetatik jaso ditugu datuak: Mondragon Unibertsitatetik, Euskal Herriko Unibertsitatetik eta Paueko eta Aturri Aldeko Herrietako Unibertsitatetik; lehenengoa pribatua da, eta gainontzeko biak, publikoak. EHUtik bi erantzun jaso ditugu, Informatika Fakultateko eta Eskola Politeknikoko liburutegietatik (2008tik fisikoki bat eginda daude) eta Liburutegi Zentraleko Dokumentazio eta Erreferentzia Zerbitzutik. Ez dugu lortu, ordea, EHUko Liburutegiko zuzendaritzatik erantzunik jasotzea, eta ezin izan dugu liburutegi handi eta konplexu horren nondik norakoen berri jakin. Web orriak dioenaren arabera, Leioako Liburutegi Zentraletik aparte, beste 24 liburutegi daude fakultate eta eskoletan.

Unibertsitateko liburutegiek ikasleei, irakasleei eta ikertzaileei ematen diete zerbitzua, baita bertako gainontzeko langileei ere (23). Horrez gain, Pauekoak kanpokoei ordainduz gero erregistratzeko aukera ematen die. Oso ordutegi zabalak dute. Horretan Mondragon Unibertsitateko liburutegia apartekoa da, 07:45etik 00:00era ematen baitu zerbitzua; zabaltasun horrek, ordea, zentroak dituen baliabideekin kontraste egiten du, bost lagunek lan egiten baitute bertan, hiru zentroan banatuta, eta hamar bekadunen laguntzarekin.

Liburutegiek zerbitzu asko eskaintzen dituzte: orientabide eta informazio zerbitzua, informazio bibliografikoa, bilaketa dokumentala eta bilaketa bibliografiko pertsonalizatuak, dokumentuak eskuratzeko zerbitzuak, dokumentuen mailegua, dokumentuen erreserba, biblioteken arteko mailegua, material bibliografikoaren erosketa eta kudeaketa, funtsen kontsulta zuzena — liburu, aldizkari, tesi eta erreferentziako lanak —, katalogoak sarean, liburutegi digitalak, aldizkari elektronikoak, sarean dauden datu baseak, alertak eta informazioaren gaikako hedapena, errepro-

grafia eta erreproduzio digitala. Erabiltzaileak baliabide elektronikoetan eta, oro har, informazio konpetentzietan trebatzeko aukera ematen dute.

Paueko eta Aturri Aldeko Herrietako Unibertsitateak bi zentro ditu, Baionan eta Angelun, eta horietan 40.000 liburuki, 300 aldizkari paperean eta hainbat baliabide elektroniko dituzte. Dokumentu gehienak frantsesez daude; ikerketa lanak, berriz, ingelesez; euskarazko funtsak ere badituzte (hizkuntza, literatua, historia...).

EHUko Informatika Fakultateko eta Eskola Politeknikoko liburutegiak 46.400 liburu (gaztelaniaz, ingelesez eta euskaraz), aldizkako 238 argitalpen, arau tekniko (UNE, ISO...), tesi, karrera osteko proiektu, bideo, DVD eta CD-ROM, aldizkari elektroniko (Elsevier, Wiley, Springer, IEEE...) eta datu baseen sarbidea eskaintzen du.

Mondragon Unibertsitateko Biblioteka ikerkuntza eta ikaskuntzarako baliabide zentroa da, baita irakaskuntza eta unibertsitate osoaren funtzionamendu, kudeaketa eta jarduerarako baliabide zentro bat ere. Bibliotekaren eginkizuna informazio baliabideak eskuratzea eta hedatzea da, unibertsitateko ikasle eta irakasleen prozesuak eta ezagutza lantzeko beharrezkoak diren informazioa eta baliabide teknologikoak eskainita. Ekipamendu handiak ditu, guztira 2.356 metro koadroren hiru campusetan banatuta: haririk gabeko sarea, 450 lan eremu, bakarkako eta taldeko lanetarako 19 gela, ordenagailu gela bat, 125 ordenagailu, eskanerrak, fotokopiagailuak, inprimagailuak eta ikus-entzunezko ekipoak. Mailegu zerbitzuan ordenagailu eramangarriak, argazki kamerak, bideokamerak, haririk gabeko txartelak, audio grabagailuak eta abar dituzte; gainera, ikasleek beren prestakuntzarako behar dituzten aplikazio informatiko guztietara sar daitezke bibliotekako ekipo informatikoetatik. Katalogoa eta datu baseak eskaintzen ditu, Dspace biltegia (E-Print), baita liburu (Springer, e-libro, Referex...), aldizkari (Springer, Scientific.net...) eta aurkibide elektronikoak ere (mila aldizkari baino gehiagoren aurkibideak) (24). Liburutegiak zenbait ikastaro ematen ditu: bibliotekaren aurkezpena eta informazio bibliografikoari buruzko sarrera saioak; ikastaro espezializatuak, informazio baliabide konkretuei buruzkoak eta eskaintza unibertsitarioaren barnean araututa dauden ikastaro espezifikokoak.

Mendekoak: Hezkuntza zentroak

Zentro publiko zein pribatuak dira; guztietan ikasleak eta irakasleak dira erabiltzaile nagusiak, baina ez bakarrak, jende guztiarentzat zabalik baitaude. 1980an sortu zen lehena, eta 2000. urtean azkena. Langile kopurua batetik lura doa. Lau langile dituen, eskolako izateaz gain, liburutegi publikoa ere bada, eta liburutegi publikoen sarean dago.

Aldaera: Nafarroako Antzerki Eskola

1993an sortu zen zentroa, eta bi lagunek lan egiten dute bertan. Nahiz eta pribatua izan, pertsona guztientzat zabalik dago. Gordetzen dituzten funtsak hauek dira: liburuak, antzerki testuak, bideoak, CD-ROMak, musika (hainbat formatutan), aldizkariak, eskolako artxiboak, lehiaketak, ekoizpenak, mintegiak, kartelak eta pro-

gramazioak. Nafarroako arte eszenikoaren inguruko datu baseak osatu dituzte; profesional, talde, ekoizpen eta emanaldiei buruzkoa, hain zuzen. Bibliotekak zerbitzu hauek eskaintzen ditu: kontsulta, aholkularitza eta mailegua. Kultur enpresa, ekoiztetxe, *casting* enpresa eta abarrei aholkularitza eskaintzen diete. Eskolak aldizkari bat ekoizten du, udako ikastaroak egiten ditu eta beste arloetako ikasle eta ikertzaileei Nafarroako antzerkiaren inguruko informazioa zabaltzen diete. Nafarroako antzerkiaren historiaren inguruko ikerketa lanak egiten eta argitaratzen dituzte.

Aldaera: Francisco Escudero Kontserbatorioa (Donostia)

Liburutegia 80ko hamarkadan sortu zen, liburuzainaren plaza sortu zenean. Lehenagotik, kontserbatorioak funtsak biltzen zituen (partiturak, liburuak eta diskoak), irakaskuntzaren jarduerak sortutakoak, eta irakasleen eta musikarien dohaintzei esker lortutakoak. Zerbitzuak irakasle, ikasle eta erabiltzaile orori eskaintzen dizkiete. Liburutegiak eskuratze askeko funtsak ditu, erreferentziako lanak, liburuak, partiturak eta CDak; azken horietan 20.000 titulu dituzte. Horien katalogoa Absys-en dago. Horrez gain, bi artxibo musikal garrantzitsu daude, banda eta orkestrarena, kontserbatorioiko bertako eta Donostia hiriko orkestra eta banden musika jarduerak sortuak. Orkestrarena, inbentariatua eta katalogatua dago, eta bandarena, inbentariatua soilik. Bertsiok, konponketak eta eskuz idatzitako musika dituzte, eta Knosys programarekin osatutako datu basean 6.000 titulu daude. Horri guztiari, kontserbatorioiko artxibo historiko-administratiboa gehitu behar zaio.

Aldaera: Udalaren Musika eta Dantza Eskolako liburutegia (Donostia)

Eskolako liburutegia da, baina, aldi berean, publikoa da. Lau lagunek lan egiten dute bertan. Bilduman, liburuak, aldizkariak, musika dokumentuak zenbait formatutan (CD, DVD, CD-ROM), partiturak eta artxiboak (talde partiturak) daude. Pianoak ere badaude partiturak probatzeko. Zerbitzu hauek eskaintzen ditu: aretoan bertan materiala kontsultatzeko aukera (erreferentziako lanak, liburuak, partiturak, fonoteka), OPAC katalogoa sarean, jendearentzako arreta eta orientazioa, informazioa eta erreferentzia saila (bekak, deialdiak, kontzertuak, ikastaroak, lehiaketak), desideratak eta iradokizunak, eta erreprografia. Honako atal hauek daude: mikroteka-Internet, fonoteka eta hemeroteka. Jarduera osagarriak egiten dituzte: gidak, liburuxkak, ipuinaren ordua, interes guneak... Horrez gain, Interneteko gidak egiten dituzte, musika baliabideekin; eta Udalaren Txistu Bandako artxiboko saila osatzen dute (dokumentazioa, prentsa, artxiboak eta partiturak).

Mendeko zentroak: Elkargoak

Hiru zentro ikertu ditugu: bi barruko zerbitzuak dira, eta besteak sarbide librea du, baina mailegua kideek bakarrik egin dezakete. Ordutegi zabala dute. Biblioteka batek pertsona bakarra du lanean, eta beste biek, bi; horietako batean langile biek

administrari kategoria dute. Historia luzeko zentroak dira; zaharrena 1838koa da eta beste bat 30eko hamarkadakoa; hala ere, zerbitzu profesionala nahiko berria da: «Biblioteka erabilera eta zainketa egokirik gabe egon da urteetan. 2000tik aurrera ezarri zen bibliotekako zerbitzua; orduan amaitu zen funtsen katalogazioa, eta erabilera eta garapenerako hainbat arau ezarri ziren».

Hona hemen eskaintzen dituzten zerbitzuak: informazio eta erreferentzia zerbitzua, monografia eta aldizkako argitalpenen kontsulta, mailegua, arau teknikoen eta argitalpenen salmenta, dokumentazio zerbitzua (legeak, arautegia), bibliografia, eman duten ikastaroen dokumentazioa, argitalpenen erosketa eta kudeaketa, datu base juridikoen kontsulta aretoan bertan, erreprografia zerbitzua eta katalogoaren kontsulta Interneten bitartez. Zentro batek argitaratu gabeko dokumentazio juridikoaren sarbidea eskaintzen du CENDOJ Dokumentazio Judizialeko Zentroaren bitartez, eta etorkizunerako, EAEko epaitegien jurisprudentzia guztia jaso eta erabiltzaileen esku jartzeko asmoa du.

Esan dugunez, ibilbide luzeko zerbitzuak dira, hemen aurkitu dugu ordea, dokumentazio zerbitzuaren etorkizunari buruz iritzi ezkor bakarra: «Liburutegiko etorkizuna zalantzazkoa da. Ez dago aurrekontu zehatzik arlo honetarako, eta elkargoko arduradunek ez dute interes handirik ageri zerbitzu beharrezko hau bultzatzeko eta zabaltzeko; horregatik, ematen du zerbitzua gutxitzen joango dela, eta gero eta esparru eta garrantzi gutxiago emango zaiola». Zerbitzuaren arduradunak garbi adierazi du elkargoko buruzagitzan ez duela aurkitzen bibliotekak behar duen sostengua.

Mendeko zentroak: Erakunde eta elkarteak

Zentro espezializatuak dira, erakunde baten beharrak betetzeko sortuak. Erakundearen arabera, publikoak zein pribatuak izan daitezke. Gure ikerketan parte hartu dutenek hiru alorretan lan egiten dute: kulturean, gizarte gaietan eta ingurugiroan. Batzuk barruko zerbitzuak dira, beste batzuek ordea, sarrera zabala eskaintzen diete adituei edota interesatuei. Langile kopurua batetik zazpita doa; elkarte batzuetan, ordea, dokumentazio zerbitzuan ez dago jardunaldi osoko profesionalik. Zaharrena Euskaltzaindiarena da, 1919koa.

Oso zentro desberdinak ditugu, dokumentazio zerbitzuak, liburutegi espezializatuak, informazio zentroak, artxiboak... Batzuk erakunde handien barruan daude, edota ikerketarekin lotura zuzena dute; adibidez, INRA, SIIS (Eguia-Careaga Fundazioa) —non dokumentu berrien %73 formatu elektronikoa jasotzen diren— edota IISJ Soziologia Juridikoko Oñatiko Nazioarteko Institutua —non katalogazioa eta analisi dokumentala ingelesez egiten diren—.

Beste batzuk, berriz, gizarte mugimenduek sortutakoak dira. IAN-ANI Arabako Natur Institutuko idazkariak honela erantzun zion gure galdetegiari: «Ez dakit gure bilduma apal honek zenbaterainoko interesa duen zuentzat. Gure institutuak, beste elkarte askok bezalatsu, diru iturri txikiekin jokatzen du: une honetan apenas erosten dugun material berririk. Den-dena boluntarioek egiten dugu:

horregatik ez dago lanposturik. Hala ere, Gasteizen leku gutxitan egongo da naturari buruz horrenbeste dokumentu eta erreferentzia liburu baturik».

Liburuak, aldizkariak, artikulak, erreferentzia lanak, literatura grisa, artxibo dokumentuak, erakundearen dokumentuak, txostenak, irudiak, soinu dokumentuak, multimediak, kartelak, eta abar jasotzen dituzte. Dokumentuak hizkuntza anitzetan daude, eta lan hizkuntza eta sortutako ekoizpenak gaztelaniaz, euskaraz (Euskaltzaindia, HABE, Euskadiko Orkestra Sinfonikoaren Artxiboa), frantsesez edota ingelesez daude (IISJn, kasu). Eskaintzen dituzten zuzeneko aukerak hauek izaten dira: informazio eskariei erantzun, kontsulta aretoak, katalogorako sarbidea, liburuaren eta bideoen mailegua, zerrenda bibliografikoak, zentro berriei aholkularitza eman, dokumentuen kopiagintza eta erreprografia. Ondorengo zerbitzu hauek ere ematen dituzte sarri: buletin espezializatuak, dokumentu berrien bibliografiak, aurkibideen buletinak, legeen inguruko alertak, datu baseen sarrera, mailegua, bilaketa pertsonalizatuak, ikerketa zerbitzuak, bibliografien osaketa, informazioaren gaikako hedapena, atari tematikoak, argitalpenak, baliabide digitalak eta ikerketarako estekak.

Datu base dokumentalek, legeen datu baseek (legediaren hustuketa egin eta gero), albisteek edota prentsakoek osatzen dituzte, baita biografia, direktorio (zerbitzu eta gizarte zentroak, adibidez), Interneteko baliabideen gida, buletin elektronikoa, web orri, eta abarrek. Zentro batzuk sare bateko kide dira (emakumeena, Unescorena, Europako Dokumentazio Zentroena eta CEDIC Arena), eta, hala, informazioa eta baliabideak partekatzen dituzte. Hainbatetan, dokumentazioaz gain, beste zerbitzu batzuk ere ematen dira: formakuntza (baita praktiketako ikasleena ere); ikastaroak, webgunearen bidez; hitzaldiak; ikerkuntza; dokumentuen osaketa (taula estatistikoak eta mapak); memoriak; liburuxkak; argitalpenak...

Etorkizunerako, hauek dira dituzten erronkak: dokumentuak digitalizatu, katalogoak edota dokumentuak Interneten jarri, Interneteko guneak eta baliabideak kontrolatu eta webguneak berriro web 2.0ren ildotik. Zentro batek immigrazioaren inguruko thesaurus bat osatzeko asmoa duela ere adierazi du.

EMAITZAK

Aztertu ditugun dokumentazio zerbitzuak bi kategoria handitan bana daitezke: independenteak direnak, hau da, artxibo eta liburutegi handiak, eta beste erakunde baten barruan daudenak. Erakunde handiago horiek administrazioak, instituzioak, unibertsitateak, hezkuntza zentroak, ospitaleak, elkarteak, lanbide elkargoak, fundazioak, museoak, enpresak eta komunikabideak dira. Jabegoari dagokionez, aztertutako zentroen %55 publikoak dira; %43, pribatuak, eta %3k jabego mistoa dute. Era berean, zentroen %61 barne zerbitzuak dira, eta barruko erabiltzaileen beharrak betetzen dituzte soilik, kanpoko erabiltzaileerik onartu gabe. Gaiari begiratuta, oro har, kulturaren, gizarte gaien, industria eta teknologiararen, albisteen, osasunaren edota orotarikoaren inguruan lan egiten dute.

Zentroen sorrerari dagokionez, 1986tik 1990era bitartekoak izan ziren urterik emankorrenak; hain zuzen ere, informatizazioaren hasiera eta datu base dokumentalen loraldiaren sasoiak izan zenean. Horien ondoan, ikerketan ehun urte baino gehiago dituzten hiru liburutegik hartu dute parte. Bestalde, martxan jarritako zentro berriena 2006koa da.

Zentroen izendapena aldatu egiten da batetik bestera: *Dokumentazioa* da gehien errepikatzen den terminoa —47 agerraldi—; ondoren, *Liburutegia* dator —27 agerraldi—, eta, gero, askoz gutxiagotan, *Artxiboa* —13 agerraldi—. Terminok beren artean ere konbinatzen dira, eta *Dokumentazioa* eta *Biblioteka* 11 aldiz agertzen dira elkarren ondoan. Izendatzeko moduak askotarikoak dira, eta funtzioak ere sarritan horrela dira. Hala, dokumentazio, artxibo eta biblioteken funtzio tradizionalaz gain, ikertu ditugun zentroetan dokumentazio eta informazio baliabideen hezkuntzako funtzioak eta Wegunearen mantentzea ere ageri dira.

Zerbitzu gehienak neurri txikikoak dira: %29k langile bakarra dute, eta %71k, lau langile baino gutxiago; %19 zentrok lau eta hamar langile artean dauzkate enplegatuta, eta %10ek soilik dituzte hamar langile baino gehiago. Langileen artean emakumeak dira nagusi.

Dokumentazio zentroko langileek hezkuntza maila altua dute, %75ek unibertsitate ikasketak dituzte, horietatik %54 lizentziadunak dira, batik bat Historia eta Filologiako karreretan. Graduatu-ondokoak, %10 dira; horietatik %56, Bibliotekonomian eta Dokumentazioan; diplomadunak %11 dira, eta horietatik %36k diploma espezifikoa dute.

Ikusten denez, ikasketa espezifikokoak dituzten langileak oso gutxi dira; zalantzarik gabe, horren arrazoia da Euskal Herrian ez dagoela Bibliotekonomia eta Dokumentazio alorreko ikasketarik, ezta Kantauri aldean beste inon ere (25). Gabezia horrek ageriago uzten ditu lurraldeko unibertsitate publikoen eta pribatuen garrantzia eta beste edozein alorretako ikasketek duten garapen maila. Tradizionalki, jendea lanean bertan prestatu da, ikastaro laburren bitartez osatuz prestakuntza. Baina informazioko sektoreen konplexutasunak, dinamismoak eta ondorio sozioekonomikoek garbi uzten dute beste hezkuntza eredu batzuen beharra, ikerketarako bidea ematen dutenak eta unibertsitate hezkuntza espezifikoen bitartez bakarrik aurrera eramanez daitezkeenak (26). Gainera, Dokumentazioaren alorreko gabeziak eragin txarra dute Euskal Herrian egiten den ikerketaren garapen, ikusgarritasun eta balorazioan, eta hori arlo zientifiko guztietan, ez baitira garatzen beste hainbat tokitan dauden egitasmo aitzindariak adibidez, tesien digitalizazioan, biltegi digitaletan, ikerketa bibliometrikoetan, ikerketako produktzioaren promozioetan, webguneen garapenean...

Aztertu ditugun dokumentazio zentroetan, oro har, funtzioak, kategoriak eta langileen formazioa bat datoz, eta ez daude, adibidez, udal liburutegietan izaten diren desberdintasunak (27) (28). Hala ere, hainbat desoreka aurkitu ditugu unibertsitateko zentro batek eskaintzen dituen zerbitzuetan eta haren ordutegian eta duen langile kopuruan. Zentroak bekadun askoren laguntzarekin estaltzen ditu desoreka horiek. Hiru ospitale eta elkargo profesional batean ere atzeman dugu desoreka, ematen duten zerbitzuaren eta bertan lan egiten duen langile bakarraren kategoria profesionalaren artean, langileak administrazioko laguntzaile kategoria baitu.

Zentroek teknologia maila altua dute, barne erabiltzaileei datu baseak eta zerbitzuak eskaintzen zaizkie Intraneten bitartez, eta publiko zabalari edota erabiltzaile espezializatuari Interneten bidez eskaintzen zaizkio informazioa eta zerbitzuak. Dokumentazio zerbitzuen etorkizuneko egitasmo gehienak bide horretatik doaz: webguneen hobekuntza, kontsulta sistema intuitiboak, web 2.0 tresnen erabilera eta dokumentazio historikoa sarean ezartzea.

Administrazioen bultzadarekin osatutako katalogo kolektiboez aparte, beste kooperazio ekimen interesgarri bakan batzuk ere aurkitu ditugu, besteak beste, zentroyen sareak eta baliabideen aprobetxamendua osasun, ingurugiro, emakume, gazte, arte garaikide eta beste zenbait alorretan.

Ekimen horiek alde batera utzita, esan dezakegu sektorean erregulazioa falta dela, baita lidergoa aurrera eramango duten erakunde enblematikoak ere. Gabezia horri lotuta daude koordinazio falta, baliabiden sakabanaketa eta baita, sektorearen ikusgarritasuna eta eragin soziala eskasa izatea ere. Lan tresnetan normalizazio linguistikoa ere egiteke dago.

ONDORIOAK

Dokumentazioa sektore klabea da hezkuntza, ikerketa, berrikuntza, garapen kultural eta gizarte kohesiorako. Asko aldatzen ari den alorra da, osagai teknologiko handia du, eta ekonomiaren motorra izan daiteke. Euskal Herrian, baina, eremu hori inguruko beste tokietan baino gutxiago eta motelago garatzen ari da, eta horrek mugatu egiten du beste hainbat sektoretako hazkundea. Dokumentazioak aurrera egingo badu, botere publikoen lidergo ekimenak behar dira, ekimen legegileak eta arautzaileak, Biblioteka Nazionalaren sorrera eta sustapena, eta ekimen publikoak eta pribatuak bat egiteko gaitasuna. Behar-beharrezkoa da, halaber, unibertsitateko ikasketa espezifikoa ezartzea lurraldean, hezkuntza beharrei modu sistematikoan erantzuteko, eremu horretan ikerketak bultzatzeko aterabide bakarra baita.

BIBLIOGRAFIA

- (1) Agirreazaldeg, T.; Ronco, M.; Camacho, I. «Euskal Herriko komunikabideetako dokumentazio-zerbitzuen analisia». In Agirreazaldeg, T. (ed.); Camacho, I. (ed.). *Dokumentazio zerbitzuak Euskal Herriko komunikabideetan*. Leioa: UPV/EHU, 2009. 19-40 orr.
- (2) Agirreazaldeg, T.; Ronco, M.; Camacho, I. «Gida: Euskal Herriko komunikabideetako dokumentazio-zerbitzuak». In *Dokumentazio zerbitzuak Euskal Herriko komunikabideetan*. Leioa: UPV/EHU, 2009. 273-326 orr.
- (3) Agirreazaldeg, T.; Ronco, M.; Camacho, I. «El desarrollo de la documentación en los medios de comunicación del País Vasco» [linean]. In *Jornadas Españolas de Documentación* (11. 2009. Zaragoza). *Fesabid 2009 infogestión*. [Madrid]: Fesabid, 2009. 175-182 orr. <http://www.fesabid.org/zaragoza2009/> [Kontsulta: 2009/07/12]

- (4) Aranes, J.I.; Landa, C.; Plano, C.; Celaá, M. «Una propuesta de investigación: Informe sobre la Documentación en la Comunidad Autónoma del País Vasco (una aproximación 1993-1995)». In *Jornadas españolas de documentación automatizada* (5. Cáceres, 1996). *Sistemas de información*. Cáceres: Fesabid, 1996. 951-960 orr.
- (5) Aranes, J.I.; Landa, C.; Plano, C.; Celaá, M. «Informe sobre la Documentación en la Comunidad Autónoma del País Vasco». *Aldeezkaria*, 1996, ekaina, 3. zenb. <<http://www.aldee.org/>> [Kontsulta: 2000/05/06].
- (6) *Censo de archivos del País Vasco*. San Sebastián: Eusko Ikaskuntza-Sociedad de Estudios Vascos, 1986-1988.
- (7) *Censo de bibliotecas del País Vasco*. San Sebastián: Eusko Ikaskuntza-Sociedad de Estudios Vascos, 1989.
- (8) Informazio hauek M.^a Luisa Hermida Corchón, Direktorioaren arduradunak eman zizkigun 2009ko irailaren 23an izan genuen elkarrizketa telefonikoan.
- (9) Arana Palacios, J.; Olaso Val, A. «Legislación sobre bibliotecas: un estudio comparativo». *TK*, 1996, ekaina, 1. zenb., 43-53 orr.
- (10) Grupo de Biblioteca Vasca. «La Biblioteca Vasca». *Sancho El Sabio. Revista de Cultura e Investigación Vasca*, 1992, 2, 333-343. orr.
- (11) *Informazioa eta dokumentazioa Euskadin [baliabide elektronikoa]: profil profesionalaren analisisa*. [Vitoria-Gasteiz]: ALDEE, 2005. 1 CD-ROM. Sarean ere eskuragarri <http://www.aldee.org/cdi/>.
- (12) Luzuriaga, G. «Balance de la situación bibliotecaria vasca y perspectivas de futuro». *RIEV. Revista Internacional de los Estudios Vascos*, 2004, 49. liburukia, 2. zenb., 481-519 orr.
- (13) *Euskal Biblioteka Nazionala behar du*. Joana Albret Bibliotekonomia Mintegia. Uztaritze, 1998ko ekainaren 5a <http://eibar.org/blogak/joana-albret/archive/1998/06/05/21/>
- (14) Cerdón García, J.A. *El registro de la memoria: bibliografías nacionales y depósito legal*. Gijón: Trea, 1997.
- (15) Fuentes Romero, J.J. *Las bibliotecas nacionales: un estado de la cuestión*. Gijón: Trea, 2003. 590 orr.
- (16) Juanena, A. «Biblioteka Nazionala Euskal Herriarentzat». *Jakin*, 142. zenb., 2004 mai-tza-ekaina.
- (17) Juanena, A. «Liburutegi Nazionalaren premiaz». Joana Albret Bibliotekonomia Mintegiko bloga, 1998ko urriaren 5ean. <http://eibar.org/blogak/joana-albret/>.
- (18) Lanbide elkargo bat osatzeko egitasmoa agertu berri dute Nafarroan.
- (19) Euskadiko irakurketa publikoko sareko liburutegien katalogoa http://www.katalogoak.euskadi.net/cgi-bin_q81a/abnetclp/O9368/ID6d9c1d3a?ACC=101
Liburutegien direktorioa http://www.kultura.ejgv.euskadi.net/r46-4879/eu/contenidos/informacion/directorio_bibliotecas/eu_direct/alava.html
- (20) Nafarroako liburutegi publikoen zerrenda http://www.navarra.es/home_es/Catalogo+de+servicios/Todos+los+servicios/popup/BibliotecasPublicas.htm
Nafarroako liburutegi publikoen katalogoa <http://www.navarra.es/AppsExt/opac/abnetcl.exe/O7018/ID68290623?ACC=101>
- (21) Katalogoa eta errepertorioa <http://ccfr.bnf.fr/portailccfr/servlet/LoginServlet>
- (22) Nekane Zurutuza Euskadiko Orkestra Sinfonikoko liburutegien hitzak.
- (23) Informatika Fakultateko eta Eskola Politeknikoko liburutegiak EHUko langile eta ikasleei sarbide osoa eskaintzen die, hau da, 50.000 lagunek dute erabiltzeko eskubi-dea, eta zentroetako bertako erabiltzaileak 3.000 dira.

- (24) Mondragon Unibertsitateko liburutegia Dialnet-eko kidea da. Errioxako Unibertsitateko bibliotekak sortutako aurkibide elektronikoen datu basea da Dialnet.
- (25) Díez Carrera, C. «La formación de los archiveros, bibliotecarios y documentalistas en España desde sus orígenes hasta la actualidad: balance y perspectiva». *Boletín de ANABAD*, 1. zenb., 2003ko urtarrila-martxoa, 123-138 orr.
- (26) Uriarte Gonzalo-Bilbao, E. «Informazioaren eta dokumentazioaren eremuko profesionalak Euskadin. Lanbidearen testuingurua eta duela gutxiko historia». In *Informazioa eta dokumentazioa Euskadin [baliabide elektronikoa]: profil profesionalaren analisisa* [Vitoria-Gasteiz]: ALDEE, 2005. 1 CD-ROM. Helbide honetan ere eskuragarri <http://www.aldee.org/cdi/>
- (27) Kerexeta Zulaika, R.; Murgia Arrese, I. «Los profesionales vascos de archivos, bibliotecas y centros de documentación». In *Informazioa eta dokumentazioa Euskadin [baliabide elektronikoa]: profil profesionalaren analisisa*. [Vitoria-Gasteiz]: ALDEE, 2005. 1 CD-ROM. Sarean ere eskuragarri <http://www.aldee.org/cdi/>
- (28) «La situación de las bibliotecas públicas en Navarra», *TK*, 18. zenb., 2006, abendua, 10. orr. Editorial. Helbide honetan ere eskura daiteke: <http://www.asnabi.com/revista-tk/revista-tk-18>

III

Bizkaiko Foru Liburutegia eta aro digitala

Ana Ares Martínez

Funts Bibliografikoen Kudeaketa eta Foru Liburutegiko Sustatzeko Atalburua
Jefa de Sección Fondos Bibliográficos y Promoción de la Biblioteca Foral

Laburpena: 100 urte baino gehiagoko liburutegi honen ibilbidearen laburpen txiki bat. Liburutegia, xx. mendeko azken laurdenean aro digitalak ekarri dituen egitura eta abagune aldaketei aurre eginez, bere gestio, funts eta zerbitzuak informazio- eta ezagutza-gizartearen eskarira moldatuz doa.

Resumen: Biblioteca Foral de Bizkaia en el periodo digital. Breve resumen del camino recorrido por esta Biblioteca, de más de cien años, que en el último cuarto del siglo xx, y haciendo frente a los cambios estructurales y coyunturales que ha traído la era digital, moldea su gestión, fondos y servicios a las necesidades de la Sociedad de la Información y del Conocimiento.

0. SARRERA

Urteak pasatu ahala liburutegien funtzino eta rolak aldatuz joan dira, gizartea aldatzen doan antzera, baina aldaketarik jaso ez dauna da berezko dauan funtzinoa: erabiltzaileei informatzeko ardura.

Egun, Informazio Gizarte edo Informazio eta Jakite Gizartean, informazioak zeregin garrantzitsu bat dauka, ze giza jarduera eta aberaste generazioa, produktzio, manipulazio eta informazio erabileratik datoz. Interneteko zerbitzu eta produktuen garapenak aukera barri bat aurkezten deustoa sareari, jakitearen ekonomian oinarri hartuta.

Liburutegi publikoen zerbitzuen oinarrietan (2002) agertzen dan moduan, liburutegi publikoa da informazio eta jakite gizartearen atea.

Informazio gizartean dagozan ezbardintasunak eragin dabe informazioa jasotzeko ezbardintasuna, bai informazio heltzeko bidean, bai informazioari emon nahi jakon esanahian. Liburutegi publikoa oinarritzko eta beharrezko zerbitzua da, gizartearen oreka eta aukera bardinen arloan eginkizun garrantzitsua betetzen baitau. Beraz, liburutegi publikoak beharrezkoa dau sare elektronikoetan parte hartzea, informazio politikei eta teknologian aukera barriei moldatuz.

Liburutegi publikoari buruzko sarreratxua egin ondorik Bizkaiko Foru Liburutegia aurkeztuko deustuet.

I. BIZKAIKO FORU LIBURUTEGIA

Bere funts fundazionala, Don Fidel de Sagarminagaren liburutegia da, ia 12.000 alegaz. Sagarminagaren liburutegia Bizkaiko Foru Aldundiaren eskuetan jausi zan, Sagarminaga hiltzean beraren senideek eskaini eutsielako. Aldundiak liburutegirik ez ebanez, funtsak, korporazino horrek eban egoitza baten jarri ziran, Bilboko Plaza Barrian.

1900eko uztailaren 31n Aldundiaren jauregi barria inauguratu zan, eta eraikuntza honen azkeneko pisuan jarri zan Sagarminagaren liburutegia. Beti be, azkeneko pisura heltzeko arazoak zirala eta, beheko pisura aldatu zan. 1907an amaitutako erreformakin areto barriak barruratu ebazan. Agiritegi eta liburutegia antolatzeko arduraduna Carmelo Echegarai jauna izan zan.

Edozelan be, urteen joan-etorriak Jauregia hastea ekarri eben, halandaze, liburutegiak betetzen eban tokia beharrezkoa zan beste behar batzuetarako, horregaitik 1921ean proiektu baten beharra sortu zan, eta jauregiaren atzean egoan eraikuntza, Aldundiarena zana, liburutegi, agiritegi eta moldiztegi probintziale-rako erabiltzea erabagi zan. 1921eko abenduaren 19an baietza emon eutsien Juan Carlos Guerra, arkitektoaren proiektuari, eta 1929ko abuztuaren 22an zabaldu zan. Garai haretan 2 sekzinotan: Baskongada eta Orokorra, banatutako 18.000 ale ebazan. Banaketa hori 1907. urteko lehen arautegian agertzen da, azalduz holan egin zala, idazle euskaldun batenak ziralako funts fundazionalak. 1909an Agiritegi-Liburutegiko lehen arduradunak, Dario de Areitiok, Baskongada sekzinoko katalogo bat argitaratu eban.

Eta aitatu beharra dago eraikuntza honek enkoadernatzeko departamentu bat eta argazki laborategi bat ebazala.

1965ean Liburutegian ikasteko gela bat jarri zan, geroago, 1988an, Foru Moldiztegiko lokaletan finkatu zana. Urte horretantxe banatu ziran Liburutegia eta Agiritegia, azken hau egun dagoan kalean jarriz; Maria Diaz de Haroko. Liburutegiko funts eta zerbitzuak hazten joiazanez, espazio birmoldaketa bat egin behar izan zan.

Hala eta guztiz be, urteak igaro ahala funtsak gehituz joan ziran eta, zerbitzuak gizarteko beharretara moldatu behar ziranez, 2003. urtean hasiera emon jakon erreforma eta hedapen proiektu bati. Lehiaketa horren irabazleak IMB taldeko arkitektoak izan ziran, eta euren lanen emoitza da egun daukagun Bizkaiko Foru Liburutegia. Eta eraikuntza honek ahalbidetu dau Informazino eta Jakite gizartearen nahi barriei aurre egin ahal izatea.

Nahiz eta leku aldaketa pila izan, ez dau aldaketarik jaso gizartearen zerbitzupan jartzeko espirituan, eta horregaitik erronka barriak, eta ikerkuntza eta konsebazino liburutegi batek dauazan behar barriei aurre egin deutse, egun zerbitzu honeek eskainiz:

- Lurralde informazinoa.
- Aldizkari ofizialak.
- Euroliburutegia.

- Ikasketa gela.
- Irakurketa eta kontsulta-gelak.
- Erreferentzi zerbitzua.
- Itsuentzako eta begietako urritasuna dabenentzako zerbitzua.
- Liburutegiko dokumentazino zerbitzua.
- Bibliografia eta heraldika.
- Kartografia eta grabatuak.
- Hemeroteka.
- Ikertzaileentzako gunea.
- Interneterako eta teknologia barriak.
- Liburutegien arteko mailegua.
- Argitalpenak.
- Wifi.

Eta web orrialdearen eraldaketagaz Erabiltzaileentzako Prestakuntza Zerbitzua jarriko da martxan.

Aurtengo programazino kulturala pasa dan urtekoa baino zabalagoa izan denez, horreetariko batzuen barri emongo dot jarraian:

- Irakurketa txokoa, zeinegaz irakurketa zaleak bildu nahi dira, euren sentsazino, pentsamentu eta esperientzien elkartrukea bideratzeko.
- Bizkaidatz II. saria.
- Ikastaroak.
- Erakusketak.
- Dokumentazino eta bibliotekonomian adituak diranen hitzaldiak.

Horreri buruzko informazino gehiago lortu nahi izan ezker, web orrialde honetara jo daitezue: <http://www.bizkaia.net>

Ez da erraza 100 urteko historiaren laburpena egitea, eta horrexegaitik aitamenez berezi bat egin gura neuke Liburutegiko funtsai buruz, batez be, eurek daben garrantziagaitik. Adibide legez:

- Inkunable bilduma, Funts zaharrak, Grabatuak eta mapak.
- Luis Luciano de Bonaparteren euskal eskuizkribuak.
- Liburutegi eta funts partikular batzuk bere dagoz: Marques de Villarías-en Liburutegi eta agiritegia (1953), Gregorio Balpardaren Liburutegia (1954), Ampuero (1966), Landeta, Arrizubieta... Irigoyen funtsak, Augustin Zubikarai, Jon Bilbao eta abar.

Horreri eta euskal funts eta hemerotekari esker, Euskadiko ikerkuntza liburutegi garrantzitsuenetariko bat bihurtu da.

II. BIZKAIKO FORU LIBURUTEGIAREN DIGITALIZAZINO PROZESUA

Bizkaiko Foru Liburutegiak argi dauanez bere ondareak dauan garrantziaz, erabiltzailei sarrera errazteko, 1988. urtean bere katalogoaren automatizazioari hasiera emon eutsan.

Horretarako Dobis Libis programa aukeratu zan, grabaketa prozesu masibo bati hasiera emonez. Horregaz paper fisikoen oinarrizko datuak katalogora gehitu ziran (idazlea, izenburua, edizinoa, urtea), geroago sarrera printzipalak garbituz.

Hamabi urte geroago, 2000. urtean, automatizazio prozesuaren planteamentu barria egin zan, eta Innopac-Millennium, Innovativa Interface enpresarena, aukeratu zan, hainbat ikerketa eta ebaluazino egin oostean. 2001eko urtarrilean hasi zan, sistema barriaren programazinoa, katalogo bakar bat definituz (katalogo orokorra), eta 7 *scope* barneratuz (Sekzino orokorra, Baskongada, Erreserba patrimoniala, Kobie, Europar Batasuna, Bizkaiko legezko gordailua, Liburutegi digitala). Katalogoa kontsulta daikegu, euskeraz, gaztelaniaz, ingelesez eta frantsesez.

Bizkaiko Foru Liburutegiak digitalizazio politika bat aurkeztu eban, kontserbazino karakterra lantzeaz gainera, berak eban patrimonio bibliografiko aberatsa, edozein erabiltzaileri aurkezteko ahalmenagaz.

Kontserbazino eta funts hedapena oinarritzat hartuz, digitalizazio prozesua bi puntu desbardinatetik garatu zan:

- Alde batetik 35 mm-ko mikrofilmetik hasiz, funts erreserba ziurtatzeko, hamar urtetan zehar, «liburu irekia» funtsa mikrofilmatu zan, beraren bi aurpegietatik, enkoadernatuak egozalako. Digitalizatzean mikrofilm fotograma bakotxa bi iruditan itzultzen da, esku ebaketaren bitartez. TIFF IV taldeak 300 kb/irudi (hirurehun kilobite irudiko) betetzen dau (biltegitratze behararen efektuen mende), eta 200 puntu, erresoluzino hazbete bakotxeko, bi koloretako (baltza eta zuria) irudia emonez.

Digitalizazinoa CD-ROMean prestatu zan ASCII «txt» fitxategi batez. CD bakotxak azpidirektorio egitura bat sortu eban. Azpidirektorio bat kode bakotxeko edo monografiaren erreferentzia topografiko bakotxeko, eta hareetara biltzen dira irudi guztiak. Azpidirektorioen izenburua jarri zan monografiaren erreferentzia topografikoagaz, baita erreferentzia egiten deutson orrialdeagaz (erreferentzia monografikoa \ alearen orrialdea.tif). CD-ROM horretan ASCII «txt» fitxategia sartu zan, erregistro bat irudi bakotxeko jarriz eta ondorengo datuak jasoz: \erreferentzia monografikoa \ irudi fitxategiaren izena luzapen tif barik \ dagoan CD-ROMaren etiketa (adibidea: R30\0000001\cd 000015).

Funts hemerografikoetan agertu daitekezan ezbardintasun teknikoetan, batez besteko okupazinoa aitatu behar dogu, honek 600 kb/irudi (biltegitratze efektuen mende) daukaz, TIFF IV talderako eta 300 puntuko erresoluzinoa hazbete bakotxeko.

Beste alde batetik, CD bakotxaren azpidirektorio egituran, badago direktorio bat, egunkariaren erreferentzia bakotxerako. Izenburua, egunkariaren erreferentzia, urte, hil eta orrialde zehatzagaz batera jarri zan (\egunkari erreferentzia \ urte

hil erreferentzia \ alearen orrialdea. tif) eta ASCII «txt» fitxategia \ egunkari erreferentzia \ urte hil erreferentzia \ irudi fitxategiaren izena tif luzapen barik \ dagoan CD-ROMaren etiketa.

- Beste alde batetik originaletik hasiz, Foru Liburutegiko mapa eta grabatu bilduma, originaletatik digitalizatzen ari gara eskuz, eskaner lauak erabiliz, eurak daben tamaina ezbardinakaitik eta euren garrantziagaitik, edonolako apurketa saihesteko asmoz.

Digitalizazioak kolorea eta JPG formatua daukaz, 300 ppp-ko erresoluzinoa-gaz. Digitalizatu ondoren, CD-ROMetan gordetzen dira, erabiltzaileak kontsulta daiezan, eta, Interneten arinago sartzeko, erresoluzino baxuagoko kopia bat dago. Dokumentuak honelaxe deitzen dira: \erreferentzia topografikoa \0000001.jpg.

Beraz, politika honeri eta programa informatikoak eskaintzen dauzan aukera barriari esker, martxan jarri zan Liburutegi Digitala.

III. LIBURUTEGI DIGITALA

Bizkaiko Foru Liburutegiaren liburutegi digitala *scope* edo sekzino jakin baten gaineko katalogoa da, eta kontsultatu daiteke, bai haren katalogo taldetik, bai zuzenean orrialde nagusitik. Hori ere lehen aitaturiko lau hizkuntzetan kontsultatu daiteke: euskeraz, ingelesez, frantsesez eta gaztelaniaz. Eta kontsulta errazteko asmoagaz, tutoriala baimentzen dauan laguntza dinamikoko pantaila dauka.

Lehenago aitatu bezalaxe, irudiak digitalizatu ondorik, Bizkaiko Foru Aldundiaren web zerbitzariak ikusgarri dauz, Innovative zerbitzaria kargatu barik. Halandaze, irudiak zerbitzaritik kanpo dagoan disko baten gordetzen dira, espazio beharra egon ezker, oztoporik ez aurkitzeko. Irudiak CD-ROM edo DVDetan jartzen dira, eta windows esploratzailea eta digitalizatutakoa direktorio birtual bategaz egiten dira, liburu bakotxetik eskegita.

Kontsultak katalogotik egiten dira, digitalizatutako erregistro bibliografikoetan oinarrituz, bai katalogotik, bai liburutegi digitaleko *scopetik*. Horrenbestez, inprimaturiko signatura topografiko bategazlotura sortzean, Foru Liburutegiaren Inno-pac Millenium katalogoan irudiak sartzen dira. Horrela, katalogoa kontsultatzen dan bitartean, barneraturiko eta digitalizaturiko funtsai zuzeneko sarrera ematen jake.

Jarraian Liburutegi Digitalaren funtsak aitatuko dodaz:

- Inkunable bilduma, hau da, 1501. urtea baino lehenagoko inprimakiak.
- Foru Liburutegiko Erreserba Obrak, hau da, XVI, XVII eta XVIII. gizaldietako edizino inprimakiak.
- Eta egun, funts horreek ezeze, funts kartografiko eta grabatuak gehitzen ari jakoz.

Gaur egun Erreserba eta Kartografi Funtsa (jasotako azkena izan ezik, digitalizazio prozesuan dagoalako) kontsultatu daiteke, Liburutegi Digitaleko katalogo orokor eta *scopeen* bitartez; hori 12.000 ale eta 129 mapaz osoturik dago.

Liburutegi Digitalean kontsultatu daitekeen obrarik zaharrena, 1472 inguruko inkunabile dana, Johanez Gerson-ena, hain zuzen be, eta izenburu honen mendean: «De Simonia; De probatione Spirituum; De eruditione confessorum; De remediis contra recidivum peccandi».

Liburutegi digitalaren obra gehienak euskerazkoak, latinezkoak eta gaztelaniazkoak dira.

Nahiz eta oraindik gure Liburutegi digitalaren parte izan ez (ez dagoalako oraindik katalogoan sartuta), liburutegiko ekipoetan kontsultatu daitezke euskal hemerrotekatik 6 izenburu (*El Nervión, Euskalduna, Euskal Erria, Excelsior, Excelsius, Gaceta del Norte*) eta Luis Luciano Bonaparteren euskal eskuizkribuak.

Gaur egun, Bizkaiko Foru Aldundiaren laguntza informatikodun enpresagaz, hau da, Lantik-egaz, batera, proiektu ezbardinetan lanean ari da gure liburutegia. Adibidez, hainbat Liburutegi Digitalen parte-hartzean, antxinako prentsarako bilaketa motoreak, metadatuaren inplementazioa... eta horrek guztiak ehun urte honetan izan dauan jokaera bardinagaz, hau da, berak dauan ondare bibliografiko aberatsaren eta liburutegi zerbitzuen hurbilketa, edozein tokitako erabiltzailei eskainitako zerbitzu birtualak eta tokian tokikoak.

IV

Euskomedia informazio sistema

Iñaki Arrieta Baro

Euskomedia Fundazioko proiektuen arduraduna
Responsable de proyectos de la fundación Euskomedia

Laburpena: Euskomediak euskal kultura eta zientziaren erreferente bat izan nahi du sarean. Horretarako hainbat dokumentazio bilgune sortu ditu: euskomedia.org (Añamendi entziklopedia, *Euskonews* aldizkaria, liburutegi digitala, dokumentazio eta argazki fondoak biltzen dituena), 3digitala.com irudi bankua, eta 2008an sortutako *Hedatuz*, Euskal Herrian egindako ikerketaren hedapena bultzatzeko OAI gordailua.

Gako hitzak: Liburutegi digitalak, euskal kultura, e-prints.

Resumen: Sistema de información Euskomedia. Euskomedia quiere ser un referente en la red de la ciencia y la cultura vascas. Para ello ha creado varios centros de documentación: euskomedia.org (la enciclopedia Añamendi, la revista *Euskonews*, revistas, fondos en los que se recogen fotos y documentación), el banco de imágenes 3digitala.com, y, extendiéndose en el 2008, el depósito OAI, para impulsar la expansión de las investigaciones hechas en el País Vasco.

Palabras clave: Bibliotecas digitales, cultura vasca, e-prints.

1. EUSKOMEDIAREN SORRERAZ

Euskomedia Fundazioa 2002an sortu zen Eusko Ikaskuntzaren ekimenez. Bere aurrekaria Mediateka izeneko zerbitzua izan zen, 1996tik Elren webgunea kudeatzen zuena.

Euskomedia euskal kultura eta zientziarekin zerikusia duten materialak digitalizatu eta sareratzeko sortu zen, bai Eusko Ikaskuntzak bai beste eragile kulturalak ekoiztutako dokumentazioa hain zuzen ere.

Helburu hori betetzeko ezinbestekoa izan da bai digitalizazio bai katalogazio lanak aurrera eramatea. Aldi berean, dokumentazio hori guztia gizarteratzeko asmoz hainbat webgune sortu ditu Euskomediak.

Komunikazio honen helburua da webgune horiek aztertzea, jasotzen duten materiala deskribatuz eta haien arteko harremanak azalduz. Bide batez, kudeatzen dituen fondoak hobeto ezagutarazteko Euskomediak etorkizun gerturako dituen asmo eta proiektuen berri ematen da.

2. EUSKOMEDIAREN ORAINAZ

Gaur egun bere sorrera helburuak betetzen jarraitzen du Euskomedia Fundazioak, hau da, euskal kultura eta zientziarekin zerikusia duten materialak digitalizatu eta sareratzea.

Horretarako hiru webgune nagusi kudeatzen ditu: euskomedia.org —gune nagusia eta informazio sistemaren bihotza—, 3digitala.com —irudi bankua— eta *Hedatu*z —OAI gordailua—.

Webgune hauetako bakoitzak bere zeregin propioa du, baina guztiak elkarlotuta daude eta zenbait ezaugarri partekatzen dituzte:

- Kalitate handiko digitalizazioa.
- Sareratzea baino lehen egindako dokumentazioaren tratamendu sakona, katalogazio zabal baten bitartez.

2.1. euskomedia.org

Gaur egun Euskomediak kudeatzen duen webgune nagusia euskomedia.org euskal kulturarako informazio sistema da.

Sistema hau hainbat dokumentu fondoek osatzen duten: Auñamendi entziklopedia, Euskonews aldizkaria, liburutegi digitala —Eusko Ikaskuntzak eta beste erakundeek argitaratutako liburu eta aldizkariak osatutakoa—, dokumentazio fondoak, eta abar. Banan bana aztertuko dugu bakoitza.

- **Bernardo Estornés Lasa - Auñamendi Entziklopedia:** «*Auñamendi*» *Euskal Herriko Entziklopedia Irudiztatu Orokorra* delakoaren bertsio digitalizatua. 150.000 artikulua baino gehiago jasotzen dituen entziklopedia historiko hau —1960ko hamarkadan hasi zen bere argitaralpena— eguneratze prozesuan dago 2001ean sarean jarri zenetik. Prozesu honek aurrera darrai gaur egun ere. Laster, gainera, aukeratutako zenbait artikulua euskaraz ere irakurtzeko aukera izango da.

URL: <http://www.euskomedia.org/aunamendi>

- **Liburu eta aldizkariak:** Fondo honetan Eusko Ikaskuntza eta bestelako eragile kulturalen bildumetan argitaratutako artikulua eta monografiak biltzen dira. Horietatik hamabi mila baino gehiago bere osotasunean daude eskuragarri. Hau da euskomedia.org-ek biltzen dituen fondoetatik liburutegi digitalik tradizionalena.

URL: <http://www.euskomedia.org/analitica>

- **Euskonews:** 1998az geroztik argitaratzen den euskal kultura eta zientziaren gaineko astekari elektronikoa. Bere webgune propioa izateaz gain —<http://www.euskonews.com>—, aldizkaria euskomedia.org informazio sisteman aurkitu daiteke. Honek bilaketa sinpleak nahiz aurreratuak egiteko aukera ematen dio Euskonewseko irakurleari.

URL: <http://www.euskomedia.org/euskonews>

The screenshot shows the Euskomedia website interface. At the top left, there is a logo for 'euskomedia fundazioa' and a vertical sidebar with 'EUSKAL' and 'FRANZES' labels. The main header reads 'Euskomedia Fundazioa, Euskal Kultura online'. Below this, there is a 'multimedia galeria' section featuring a crucifix image and the text 'ZUMAIA, Gipuzkoako kostaldeko herria' and 'ANÓNIMO'. To the right, a '2009eko apirilak 15' section lists 'Efemerideak' with dates and names like 'RUZ OPITZ, Maria Joséren jaiotza'. Below the gallery, there are several content tiles: 'Leizarraga Testament Berria', '3digitala Irudi bilduma berria', 'Monografikoa Juan Garmendia Larrañaga', 'Mapa Auñamendiko eduki guztiak eskura', 'Gipuzkoa 1915eko irudiak', 'RIEV Lehenengo mendeurrena', 'ADEVE Naturaren ahotsa', and 'Kooperatibak Arizmendiarieta Fonda'. At the bottom, there is a search bar with tabs for 'SINPLEA', 'ZEHATZA', 'AUDIOA', and 'IRUDIAK', and a search button labeled 'bilatu fondooan'.

1. irudia

euskomedia.org: hasiera orria

— **Agiri Fondoak:** Manuel de Irujo, Angel de Apraiz, Manuel Lekuona, José María Arizmendiarieta eta Ameriketako Euskal Erakundearen Federazioaren agiri fondoak. Pertsona eta erakunde hauen agiri fondoen katalogazioa topatuko dugu bertan, guztira hamar mila dokumentu baino gehiago biltzen direlarik. Zenbait kasutan —adibidez, Arizmendiarieta fondoko gutun hau: <http://www.euskomedia.org/fondo/15997>— dokumentu digitalizatua ere eskuratu daiteke.

URL: <http://www.euskomedia.org/fondo>

— **Euskal Kultura prentsan:** 1900 eta 1975 bitarteko Euskal Herri osoko egunkarietan bildutako euskal kulturaren inguruko prentsa berriak. Gainera, Eusko Ikaskuntzak 1936rarte jasotzen zuen prentsa txostena ere biltzen du, kasu honetan berria digitalizaturik jaso daitekelarik.

URL: <http://www.euskomedia.org/cultura>

— **Euskal kantutegia:** Bost mila partiturez osatutako kantutegia. Kantu bakoitzeko deskribapen zabal bat —izenburua, jasotze lekua eta data, letra, eta abar—, partitura, eta musika MIDI formatuan jasotzen ditu.

URL: <http://www.euskomedia.org/cancionero>

- **Multimedia galeria:** Hainbat argazki fondo —Eusko Ikaskuntza, Auñamendi, eta zenbait udal—, audioak eta bideoak biltzen ditu. Azken hauek oraindik gutxi dira baina ekoizpena geroz eta zabalagoa da.
URL: <http://www.euskomedia.org/galeria>
- **Jaien egutegia:** Euskal herrietako jai eta ospakizunen erregistroa.
URL: <http://www.euskomedia.org/calendario>
- **Euskal Bibliografia:** gaika sailkatutako hainbat bibliografia jasotzen ditu fondo honek. Bestek beste Erdi Aroko artea, Errenazimenduko artea, Arte Barrokoa, Historia Garaikidea, Erromanizazioa Euskal Herrian, Arrantza, Antropologia, eta abar.
URL: <http://www.euskomedia.org/bibliografia>

Argi geratzen da beraz, fondo hauetatik zenbait erreferentzialak direla, dokumentuaren kokapenaren berri baino ematen ez dutelarik. Honen adibide, *Euskal Bibliografia* fondoa litzateke. Beste kasu gehienetan, erreferentziaz gain dokumentua bere osotasunean jasotzeko aukera dago, bai fondoaren osotasunean —*Auñamendi* edo *Euskonews* kasu—, bai fondoaren zati batzuetan —*Liburu eta Aldizkariak*, *Agiri fondoak*—.

Ikus dezakegunez hainbat motatako dokumentu eta fondo ez osatutako liburutegi digital bat da [euskomedia.org](http://www.euskomedia.org), bere zentzurik zabalenean gainera, testuez gain irudia, bideoa, eta audioa ere jasotzen baititu.

Sistemara gehitzen den dokumentu bakoitza modu oso sakonean katalogatzen da. Modu honetan, erabiltzaileak informazioa hainbat daturen arabera bilatzeko aukera dauka: egilea, izenburua, laburpena, baita testu osoa ere dokumentua eskuagarri baldin badago.

2.2. 3digitala.com

Euskomedia Fundazioaren irudi bankua da 3digitala.com. 2009ko apirillean guztira 28.825 irudi biltzen ditu, hainbat egile eta gairen ingurukoak baina beti ere euskal kulturari lotutakoak: arkitektura, paisaia, jaiak, etnografia, eta abar.

Bere alderdi komertziala duen zerbitzu bat da 3digitala.com. Izan ere, bertan erakusgai dauden argazkiak erabili ahal izateko lizentzia bat ordaindu beharra dago.

Baina argazki horien kultur balioa ere azpimarratu nahi du Euskomediak. Horretarako argazki bakoitzaren deskribapena egiten da, alderdi kulturala azpimarratuz. Horrez gain irudiaren testuinguru kulturala ulertzen lagundu nahi da, Euskonews eta Auñamendi Entziklopedian harremana duten artikuluetara loturak eginez.

2.3. Hedatuz

Hedatuz —<http://hedatuz.euskomedia.org>— Euskomediak 2008an sortutako e-prints gordailu bat da. Nazioartean geroz eta gehiago dira mota honetako liburutegi digitalak, gure artean oraindik indar handirik ez badute ere.



Enskomedia Fundazioa

izena eman | pasahitza berreskuratu

IRUDIARIK
HIRRE KONTUAN
NOIA EROSI
HORTZUK GARRA

Klik ta klik Euskal Herrian barrena



Baionako Bonnat museoko lehenengo pisua (Lapardi).
[000422/4] Javier Juanes (Kudeatutako eskubideek)



Sahats Artxiboa
Ezagutu Nafarroako artsa eta inguruak



Kepa Etchandy
Iparraldeko sekulako argazkiak Kepa Etchandy-ren eskutik



argazkiak
28.825

ausazko irudia
Hauen arabera bila daitezke: gako-hitza, kontzeptua, irudien zenbakia, argazkilaria...

[bilakata zehatza](#)

SAILAK

- Artea
- Ekonomia
- Etnografia eta folklorea
- Pertsoniak
- Natura
- Paisajeak
- Material grafikoa
- Numismatika
- Ikurra
- sailik gabe
- Bibliografia materiala

BESTE AUKERAK

- > TOPONIMOAK
- > EGILEAK
- > SAILAK

ARGAZKILARIAK

Argazkilaria al zara?
Profesionala edo zaletua izanik Euskal Herriko aparteko irudiak al dituzu?

Zuk ere esku har dezakezu

eskatu informazioa

BUZOTIA

Bidal itzazu zure zalantzak, ideiak, iradokizunak, ...

HARREMANA | BALDINTZAK
info@3digitala.com
LEGEZKO OIBARRA | PRIBATASUN POLITIKA




2. irudia

3digitala.com: hasiera orria

E-prints gordailuen helburua erakunde baten barruan egindako ikerketa zientifikoaren emaitzak edo gai jakin bat jorratzen duten dokumentu akademikoak —artikuluak, tesiak, eta abar— biltzea da.

Dokumentu hauen hedapena errazteko OAI-PMH metadatuaren banaketarako protokolo berezia garatu zen. OAI-PMH protokoloa betetzen duten gordailuen kudeaketarako programen artean, *Hedatu*-ek EPrints softwarea —<http://www.eprints.org/>— erabiltzen du. Honek *Hedatu*-en dauden dokumentuak protokolo bera darabilten bilatzaileek oso modu errazean topatu eta indexatu ahal izatea ahalbidetzen du.

3. irudia

Hedatuz: hasiera orria

Orain arte *Hedatuz* gordailuan Eusko Ikaskuntzak argitaratutako aldizkarietako artikuluak biltzen dira. Epe ertainera, Euskomediak duen helburua beste eragile batzuek aurrera eramandako ikerketaren emaitzak jasotzea da.

3. EUSKOMEDIAREN ETORKIZUNAZ

2009 eraldaketa urtea izango da euskomedia.org webgunearentzako, goitik behera aldatuko delarik.

Alde batetik, diseinua berrituko da. Oraingoak urte batzuk baditu eta aldaketen beharrea dago. Webgunearen berritze hau ez da soilik estetikoak izango, egituraketa ere aldatuko da. Hasierako orriak erabiltzaileari informazio gehiago ematea bilatuko da, sistema osoan aurki daitekeen informazioaren adierazgarri bilakatuz.

Diseinu berriak, bestetik, informazio biltegien integrazioa bilatuko du. Orainartean euskomedia.org-en garrantzi handia eman zaio dokumentu batek sistema osoan duen kokapenari —zein dokumentu fondoren atal den hain zuzen ere—. Baina erabiltzaile gehienek nahi dutena informazio jakin bat da, informazio horren iturriari garrantzi handiegirik ematen ez dietelarik. Hori dela eta hemendik aurrera ez zaio horrenbesteko pisurik emango dokumentuen kokapenaren informazioari —eskuragarri jarraituko duen arren—, sistemari osotasuna ematea bilatuko delarik.

Integrazio hori bultzatzeko, batetik, nabigazioa erraztea bilatuko da. Orainartean, behin euskomedia.org-eko dokumentu jakin batetara iristsita, ez da erraza

gertatzen berarekin lotuta dauden beste batzuetara joatea, bilaketa berri baten birtartez ez bada. Hemendik aurrera, horrelako harremanak erabiltzailearen esku jarriko dira. Horretarako beraien edukiaren arabera elkarlotuta dauden dokumentuen artean lotura zuzenak erabiliko dira. Lotura horiek fondo bereko nahiz ezberdinetako dokumentuen artean egitea posible izango da.

Aldi berean, dokumentu bilguneen artean ere harremanak sendotuko dira. Adibidez, Auñamendiko artikulua eta bilaketetan harremana duten 3digitala.com-eko argazkiak erakutsiko dira.

Hurrengo aldaketa garrantzitsua sistemaren bilaketa azpiegitura berritzea izango da. Laster euskomedia.org-en bilaketak egiterakoan *Google Search Appliance* izeneko teknologia erabiliko da. Google-ek garatutako sistemak bilatzaile honen azkartasun eta erraztasuna emango dizkio erabiltzaileari euskomedia.org-en bilaketak egiterakoan.

Honek abantaila asko eskainiko ditu:

- **Ezagutzen duen bilaketa modu bat**, GSAk Google-n funtzionamendu bera baitu. Erabiltzaile adituak bilaketa aurreratua erabiltzen jarraitu ahal izango du baina noizbehinkakoak nahikoa izango du hitz bakar bat idaztea, Interneten bilaketak egitean ohi duen moduan.
- **Azkartasuna**, bilaketetarako zerbitzari berezi batek ematen dion arintasunarekin. Honekin euskomedia.org-ek duen arazorik handienari, webgunearen arintasun faltari hain zuzen ere, aurre egitea posible izango da.
- **Dokumentuen arteko harremanak modu errazean ikustea**. Lehenago adierazi dugun moduan, orain artean sistema osatzen duten dokumentuen harremanak nahiko ezkutuan daude. Hemendik aurrera, GSAren laguntzaz eta katalogazio sakon bat oinarri hartuta, elkar lotuta dauden dokumentuen artean nabigatzea erraztuko da.

Bi aldaketa hauekin, diseinua/egitura eta bilaketa, euskomedia.org guztiz berrituko da 2009an. Honela erabiltzaileei zerbitzu hobea ematea lortuko dugulakoan gaude.

Informazio sistemaren berrikuntza hauez gain, Euskomediak dokumentazio zientifikoaren hedapenerako linea estrategikoarekin aurrera jarraitzeko asmo sendoa dauka. *Hedatuz* gordailura argitalpen berriak gehitzeaz gain, Fundazioa doktorego tesien bilgune izango den gordailu bat irekitzea aztertzen ari da. Bi zerbitzu osagarri hauek Euskal Herrian egiten den ikerketa zientifikoaren emaitzen hedapena hobetzea ekarriko lukete, nazioartean erabiltzen ari diren sistema berak sustatuz.

Euskal kulturaren informazio sistema osatu honekin eta etengabeko berrikuntza oinarri hartuta, Euskomedia arlo honetan sarean erreferente izatearen helburura gerturatzen ari garelakoan gaude.

2. atala

Enpresako dokumentazio
zerbitzuak aro digitalean

2.^a parte

*Servicios de documentación
de empresa en la era digital*

V

Fagor sozietate kooperatiboaren Dokumentazio Zerbitzua: 20 urte erne eta tinko

Juan Carlos Guridi

Fagor Kooperatibako Dokumentazio Zentroko Arduraduna
Responsable del Centro de Documentación de Fagor Cooperativa

Laburpena: Artikulu honetan azaldu dugu Fagor sozietate kooperatiboaren dokumentazio zentroa nola bilakatu den, zer zerbitzu dauzkan eta nork erabiltzen dituen, eta lanerako zer tresna baliatzen ditugun.

Resumen: El Centro de Documentación de Fagor Sociedad Cooperativa: Veinte años al pie del cañón. En esta ponencia se señala la evolución del Centro de Documentación de Fagor S. Coop, contemplando los servicios que presta, qué usuarios lo utilizan y las herramientas que maneja.

SARRERA

Fagor Dokumentazio Zerbitzua 1987an sortu genuen, Fagor taldeko kooperatibei honako gai hauei buruzko informazioa emateko asmoarekin: merkatua, lehiakideak, bezeroak, ingurune ekonomikoa eta enpresen gestioa, besteak beste.

Egitasmo horretan Fagor sozietate kooperatiboaren Zuzendaritzaren bultzada ezinbestekoa izan zen.

BILAKAERA

Hamarkadaz hamarkada hauexek izan dira mugarri nagusiak:

1980

Fagor Dokumentazio Zerbitzua sortu zenean, Fagor taldean 11 kooperatiba zeuden, hainbat industri sektoretakoak: etxetresna elektrikoak, automobilgintza, makina erreminta, elektronikak, e.a.

Zer motatako Dokumentazio Zerbitzua nahi genuen jakiteko, kooperatiben organigramaren lehen lerroko pertsonak elkarrizketatzea erabaki genuen. Xedea zen haien informazio beharren profila lortzea. Jasotako erantzun/inkestak ikusirik, erabaki genuen bi produktu dokumental sortzea:

- Boletín de Noticias - Notizien Buletina.
- Boletín de Referencias - Erreferentzien Buletina.

Gero, hamarkada amaieran, honako hau sortu genuen:

- Resumen de Prensa - Prentsa Laburpena.

Kontsulta zehatzak hartzen ere hasi ginen: edozein galderari erantzuteko zerbitzua zen. Era guztietako gaiei buruzko kontsultak egiten zizkiguten: merkatua, enpresak gestionatzeko teknika berriak, teknologia, zuzenbidea, e.a.

Informazioa erabiltzeko eta gordetzeko baliatu genituen datu-base dokumentala (JLB-doc), testuen prozesadorea (DW3), fotokopiagailua (paperik gabe lan egitea ezinezkoa zen urte haietan) eta faxa.

Informazioa erabiltzen zuen, hasiera batean, lehen aipatu organigramaren lehen lerroko jendeak. Urteek aurrera egin ahala, informazioa bigarren lerroko jendeari eta gure informazioa behar zuen jende gehiagori ematen hasi ginen.

1990

Hona hemen hamarkada horretako mugari nagusiak:

- MCC sortu zen (Mondragon Corporacion Cooperativa, oraingo Mondragon). Hala, lehen eskualdeka antolatuta zeuden kooperatibak elkarte sektorialetan antolatzen hasi ziren.
- Notizien Buletinak eta Erreferentzien Buletinak orokorrak izatetik (buletin berdina sektore desberdinetako pertsona guztientzat) sektorialak izatera pasatu ziren.
- Teknologiari dagokionez, Internet erabiltzen hasi ginen eta, hamarkada amaieran, baita posta elektronikoa ere.

Horiez gain, ezin ditugu ahaztu hiru zerbitzu berri hauek:

- Jardueraren txostena.
- Dosierra.
- Zaintza teknologikoa.

Teknologia arloan, Basis datu-base dokumental berriarekin lan egiten hasi ginen. Kontsulta zehatzen eta erabiltzaileen kopuruak gora egin zuen.

Jauzi handia egin genuen: Fagor Taldeko kooperatibei informazioa ematetik MCCko industria arloko dozenaka kooperatibei informazioa ematera pasa ginen.

2000

Hamarkada hasieratik gaur egun arteko mugari garrantzitsuenak hauek dira:

- Gure web orria jarri dugu martxan.
- Prentsa laburpena formatu elektronikoan daukagu.
- Erabiltzaile kopuruak gora egin du.
- Internet denontzat eskuragarri dago.
- Bulegoak paperik gabekoak dira, hein handi batean.

ZERBITZUAK

Zerbitzu garrantzitsuenak Notizien Buletina, Prentsa Laburpena eta Erreferentzien Buletina dira.

Notizien Buletina

Buletinaren helburua da kooperatiben inguruko gai hauei buruzko informazioa ematea: lehia, merkatua, bezeroak, produktuak eta testuinguru ekonomikoa. Denborarekin, testuinguru ekonomikoari buruzko informazioa kendu egin dugu.

Notizien adibideak

**VAILLANT Y HONDA
DESARROLLARÁN
UN SISTEMA DE
COGENERACIÓN**

Honda Motor Co. and German heating equipment maker Vaillant will jointly develop a compact residential cogeneration system for sale in the European market, Honda said Tuesday.

Honda will be in charge of developing a unit combining a natural gas-powered engine and an electricity generator, while Vaillant will develop a heating and hot water supply unit. The two units will be put together to make the cogeneration system.

The system will be first launched in Germany in 2011, according to Honda. It can reduce energy costs and carbon dioxide emissions by about 25 pct, Honda said.

In Japan, Honda launched residential cogeneration systems in 2003. More than 80,000 households are using Honda products in the nation, according to the firm.

In the United States, the company has been selling cogeneration systems jointly with a local company, called Climate Energy, since 2007.

JIII PRESS, 10 MARZO 2009

**ARCELIK: LAS VENTAS DE LOS 5 PRINCIPALES
APARATOS DE LÍNEA BLANCA CAYERON
UN 3,9% EN 2008**

VENTAS DE LOS 5 PRINCIPALES APARATOS
DE LÍNEA BLANCA

Sales – 5 Main Goods						
Units '000	Q4 2008	Q4 2007	% Change	2008	2007	% Change
5 MAIN PRODUCTS-TOTAL	2.368	2.588	-8,5	9.977	10.384	-3,9
Refrigerator	900	931	-3,3	3.942	4.177	-5,6
Washing Machine	649	793	-18,2	2.717	2.969	-8,5
Dishwasher	297	332	-10,7	1.420	1.362	4,2
Oven	417	432	-3,5	1.614	1.626	-0,7
Dryer	106	101	5,3	285	251	13,5

ARCELIK, MARZO 2009

Informazio mota hauek jorratzen ditugu:

Enpresari buruzkoa: emaitzak (salmentak, etekinak, ratioak...), salerosketak, akordioak, ezarpenak, inbertsioak, enpresen indarguneak eta ahulguneak, ikuspegiak, aurreikuspenak, lortutako kontratuak, itxitako lantegiak, lan murrizketak, esportatutako produktuak, inportazioa (nondik norakoak).

Sektoreari buruzkoa: esportazioa, inportazioa, salmentak, bilakaera, e.a.

Informazio iturriak, besteak beste, hauek dira: munduko sektore guztietako al-dizkariak, elkar-teak, egunkari ekonomikoak eta orotarikoak, Internet eta nazioar-teko datu-baseak.

Buletina bi astero eta hilero egiten dugu, informazio kopuruaren eta erabiltzai-leen premien arabera.

Informazio gehiena gazteleraz, ingelesez eta frantsesez dago; italieraz, portu-gesez eta beste hizkuntza batzuetan gutxiago.

Prensa Laburpena

Dokumentu horretan prentsan agertzen den informazioa laburtzen dugu: gure kooperatibei buruzko albisteak, egoera ekonomikoaren adierazleak, langabezia, be-rankortasuna, krisialdia, gurean eragina izan dezakeen beste herrialde batzuetako

Prensa Laburpenean agertutako berri bat

El Gobierno golpea a las renovables con un nuevo real decreto

Lidia Velasco. Sevilla

El Gobierno continúa ahogando a las energías renovables a golpe de decreto. El último, el 6/2009, que entra en vigor hoy, sacude especialmente a la termosolar, ya que obliga a los promotores que actualmente están construyendo plantas a presentar de nuevo toda la documentación en los próximos treinta días, al objeto de inscribirse en el registro de preasignación.

De este modo, limita a 500 megavatios (MW) la potencia instalada hasta 2010, mien-tras que en el anterior decreto

se contemplaba que, cuando se pusiera en marcha el 85% de este volumen, se daría una prórroga a quien estuviera desarrollando proyectos, lo que elevaría ahora mismo la cifra hasta los 900 MW.

Por ello, fuentes del sector entienden esta nueva norma como un freno a las inversiones, al igual que ocurrió con la energía fotovoltaica, aunque, por otro lado, consideran que algunas restricciones son buenas porque alejan la especulación.



La termosolar debe presentar un aval de 50 millones.

A partir de hoy, las empresas que estén levantando centrales tendrán que enviar toda la documentación (han de acreditar que poseen el 50% de los recursos y un acuerdo de compra de equipos), a falta de un

aval de 50 millones (100 euros/kilovatio), que deberán presentar al mes siguiente.

El real decreto también da luz verde al bono social, que reduce la factura de la luz a personas con bajo poder adquisitivo. Se configura como una protección del derecho al suministro de electricidad y su financiación será compartida por las compañías titulares de instalaciones de generación del sistema eléctrico.

Expansión, 8 mayo 2009

informazio ekonomikoa, sektoreko albisteak eta lehiakideenak, enpresa eta produktuetan eragina duten lege berriak, e.a.

Informazio iturriak dira harpidetuta ditugun egunkariak eta Internet.

Maiztasuna: egunero.

Erreferentzien Buletina

Zerbitzu horretan testuinguru ekonomikoari, enpresen gestioari, sektore industrialari, lehiari eta bezeroari buruzko informazioa lantzen dugu.

Arlo horietako hainbat aldizkaritan gaude harpidetuta. Horrez gain, dokumentuak Interneten biltzen ditugu eta erakunde pribatuek, administrazio bulegoek eta beste informazio iturri batzuk ere ematen dizkigute.

Erabiltzaileei erreferentzien buletinak bidaltzen dizkiegu, gero nahi dituzten dokumentuak eska ditzaten, erreferentzia aipatuta. Guk, ahal dugun neurrian, dokumentu horiek helarazten dizkiegu.

Erreferentzia pare bat

Guía-País. Brasil

Icex, 2009, marzo: 143 páginas; Oficina Económica y Comercial de España en Brasilia

Guía país de Brasil. Se aborda la estructura político-administrativa y la estructura económica; las diferentes legislaciones —mercantil, fiscal, laboral...—, el régimen de comercio exterior y los costes de establecimiento.

Recoge la superficie, demografía y muestra información de carácter práctico —formalidades de entrada y salida, comunicaciones, moneda, direcciones de utilidad, etc.—. Incluye anexos —indicadores económicos, PIB, balanza de pagos, etc.—.

Rieter Group. Annual Report 2008

2009: 96 páginas

Memoria de Rieter Group correspondiente a 2008.

Bestelakoak

Dossierra

Notizien Buletineko informazioa bildu eta dossierra egiten dugu. Jasotzen ditugu, hain zuzen, buletinean azken sei hilabeteetan agertutako albisteak, sektore bakoitzeko enpresen eta herrialdeen arabera antolatuta.

Jardueraren txostena

Txostena urtean behin egiten dugu. Erabiltzaileek aukeratutako enpresen bilakaera aztertzen dugu. Enpresen erretratu egiteko informazio iturri hauek baliatzen ditugu: enpresen urteko kontuak, «annual report», memoriak eta Notizien eta Erreferentzien Buletinak.

Kontsulta zehatzak

Zerbitzu horretan era eta gai guztietako galderei erantzuten diegu. Azken urteetan kontsulta gutxiago jaso ditugu, Interneten informazio asko dagoelako, neurri handi batean.

Zaintza teknologikoa

Orain dela gutxi arte zaintza teknologikoaren arloan literatura grisa zabaltzen genuen. Literatura grisa da ohiko argitalpen komertzialetatik kanpo zabaltzen diren dokumentuak: tesiak, «conference proceedings», aldizkari teknikoak, e.a. Dokumentu horiek datu-base teknologikoetan (Inspec, Compendex, e.a.) eta beste hedabide batzuetan bilatu eta erabiltzaileei bidaltzen genizkien. Orain lan hori Jabetza Industrialeko Sailak egiten du.

VI

Retos y respuestas de la gestión de la información en la Corporación Tecnológica Tecnalia

Edorta Aranguena, Pilar Ruiz e Izaskun Urrutia

Responsables de los centros de documentación AZTI Tecnalia,
ROBOTIKER Tecnalia y LABEIN Tecnalia respectivamente
AZTI Tecnalia, ROBOTIKER Tecnalia eta LABEIN Tecnaliako
dokumentazio zerbitzuetako arduradunak, hurrenez hurren

Resumen: La corporación tecnológica Tecnalia es la mayor empresa tecnológica del estado y está integrada por ocho centros tecnológicos vascos de distinto origen y orientación sectorial. En esta ponencia se describen las soluciones que tres de sus centros de documentación han desarrollado en relación con necesidades clásicas y comunes a las entidades a las que dan servicio: el diseño a medida de un sistema de gestión documental, la selección y optimización de recursos de información de calidad, y las acciones de formación en competencias informacionales. Pero los profesionales de la documentación de los centros tecnológicos tienen que afrontar además los retos que implican el nuevo escenario informacional, y responder a las complejas necesidades de unos usuarios que son —por definición— trabajadores del conocimiento.

Laburpena: Tecnalia teknologia korporazioko informazio kudeaketaren eronkak eta erantzunak. Tecnalia teknologia korporazioa Espainiako estatuko teknologia enpresa nagusia da, eta jatorri eta orientazio sektorial desberdineko zortzi euskal teknologi zentrok osatzen dute. Txosten honetan, bertako hiru dokumentazio zentrok zerbitzua eskaintzen dieten entitateen ohiko beharrei erantzuteko garatu dituzten soluzioak deskribatzen dira: dokumentazioaren gestio sistema diseinatzea bakoitzaren neurrira, kalitatezko informazio baliabideak hautatzea eta hobetzea eta informazio gaitasunetan trebakuntza eskaintzea. Baina teknologia zentroetako dokumentazioaren profesionalak aurre egin beharko diete informazioaren alorrean sortutako erronka berriei eta erabiltzaileen, definizioz ezagutzaren langile baitira, beharizan korapilatsuei erantzun.

LOS CENTROS TECNOLÓGICOS Y EL SISTEMA VASCO DE CIENCIA Y TECNOLOGÍA

Los Centros Tecnológicos son entidades privadas sin ánimo de lucro que forman parte del sistema de Ciencia y Tecnología, cuya misión es la de apoyar e impulsar todos los procesos de innovación y desarrollo tecnológico (I+DT), a fin de que el entorno empresarial alcance cotas cada vez más altas de competitividad industrial.

La financiación de los centros tecnológicos proviene fundamentalmente de los proyectos de I+D+I, transferencia tecnológica, asistencia técnica y formación firmados con las empresas. En un porcentaje importante, estos centros también se financian mediante subvenciones que reciben de los programas de I+D e innovación tecnológica y de los acuerdos firmados con sus respectivas comunidades autónomas.

En España existen unos 100 centros tecnológicos¹, participados tanto por empresas públicas como privadas. De éstos, 67 son fundaciones privadas sin ánimo de lucro adheridas a FEDIT, Federación Española de Entidades de Innovación y Tecnología. Más de la mitad se concentran en el País Vasco (20) y en la Comunidad Valenciana (15), que son las regiones donde más se ha desarrollado este tipo de organizaciones.

En la CAV, la política tecnológica ha estado tradicionalmente vinculada al Departamento de Industria y Comercio, poniendo el énfasis en la investigación aplicada. Por ésta y otras razones —escaso protagonismo de la universidad, ausencia de centros del CSIC o de organismos públicos de investigación en la CAV— los Centros Tecnológicos han tenido un gran protagonismo en el sistema vasco de ciencia y tecnología.

En la década de los ochenta, el Gobierno Vasco impulsó las infraestructuras de centros tecnológicos tutelados y apoyó a la I+D de las empresas. Se constituyó la red de centros tecnológicos vascos inicialmente con cinco centros; CEIT, perteneciente a la ETSII de la Universidad de Navarra en San Sebastián, IKERLAN, del grupo Mondragón, INASMET, perteneciente a la Asociación guipuzcoana de empresas de la fundición, LABEIN, de la ETSII de la Universidad del País Vasco y TEKNIKER, relacionado con la Escuela Politécnica de Armería de Eibar. Con ello se pusieron los cimientos de la denominada Red Vasca de Ciencia, Tecnología e Innovación.

En la actualidad, los Centros Tecnológicos vascos se agrupan en dos grandes entidades. La Alianza IK4 está formada por los centros CEIT, CIDETEC, GAIKER, IDEKO, IKERLAN, TEKNIKER y VICOMTech; y TECNALIA Corporación Tecnológica está formada por ocho centros: INASMET, LABEIN, ROBOTIKER, AZTI, ESI, NEIKER, FATRONIK y CIDEMCO.

Según un estudio elaborado por la consultora Deloitte en 2008, los centros han aportado un total de 914 millones de euros al Producto Interior Bruto (PIB) de Euzkadi en los últimos diez años, periodo en el que invirtieron más de 230 millones. En la actualidad trabajan más de 2.500 personas en los centros existentes en el País Vasco, de los que más del 10% son doctores. A esas cifras hay que añadir los más de 240 puestos de trabajo anuales que se generan a través de las nuevas empresas de base tecnológica.

¹ Registro de entidades de innovación y ciencia 2008.

TECNALIA CORPORACIÓN TECNOLÓGICA

TECNALIA Corporación Tecnológica nació en 2001 como un acuerdo entre los centros INASMET, LABEIN y ROBOTIKER, a los que se sumaron posteriormente AZTI, ESI, NEIKER, FATRONIK y CIDEMCO. En enero de 2008, seis de estos centros se han comprometido a iniciar un proceso de fusión que supondrá el nacimiento del mayor centro tecnológico de estado español, y uno de los mayores de Europa.



Los miembros que conforman Tecnalia, que cuenta con una plantilla de más de 1.300 investigadores y una cifra de negocio de 111 millones de euros en el 2007, están presentes en 18 sectores económicos, dominando 36 áreas tecnológicas, más de 350 aplicaciones de esas tecnologías y 10 tecnologías estratégicas. Tecnalia asiste anualmente a más de 3.000 empresas con las que colabora en múltiples proyectos.

Los centros tecnológicos de Tecnalia comparten un modelo operativo común basado en Unidades de Negocio de carácter sectorial. Actualmente, existen 20 Unidades de Negocio conformadas por equipos multidisciplinares, en sectores estratégicos:



En 2007 el 83 % de los ingresos de Tecnalia proceden de proyectos de investigación y desarrollo; el 50% correspondió a proyectos bajo contrato, el 29% a financiación pública competitiva y el 21% a financiación pública no competitiva.

Tecnalia participa activamente en los programas de I+D+i de la Unión Europea y lo demuestra siendo la segunda entidad del Estado de carácter privado en retornos europeos del VII Programa Marco. A lo largo de 2008 ha participado en 172 proyectos, liderando 29 de ellos.

LOS CENTROS DE DOCUMENTACIÓN DE TECNALIA

Los recursos documentales e informativos son una parte esencial de la infraestructura para la investigación, y en la mayoría de los centros tecnológicos de Tecnalia la existencia de una unidad o servicio de información y documentación es algo contemplado desde su misma constitución.

Bajo la denominación de Biblioteca, Centro o Servicio de Documentación, forman parte de los llamados servicios generales, dependiendo de la dirección científica, de tecnología o de marketing según los centros. Se trata de un tipo de servicio descentralizado, que atiende a cada uno de los centros tecnológicos. En

cuanto a recursos humanos, están compuestos por una o dos personas, según los centros, titulados superiores con formación específica en biblioteconomía y documentación.

Desempeñan las siguientes tareas:

- Adquisición y suministro de documentos.
- Tratamiento documental de las adquisiciones.
- Gestión de suscripciones y asociaciones.
- Creación y mantenimiento de bases de datos internas.
- Acceso a bases de datos y servicios de información online.
- Búsquedas de información (servicio pregunta-respuesta).
- Alertas informativa.
- Seguimiento de normas y legislación.
- Seguimiento y alertas de ayudas y subvenciones.
- Catalogación y conservación de las publicaciones y documentación de cursos internos.
- Archivo de prensa.
- Edición de boletines y guías.
- Servicio de Archivo.

A estos servicios, comunes en cualquier centro de documentación de empresa, se añaden otros más específicos, relacionados con la actividad específica de Tecnalia:

- Vigilancia tecnológica e inteligencia competitiva.
- Gestión del conocimiento.
- Creación, diseño y gestión de contenidos tanto en la Intranet como en la web corporativa.
- Participación en proyectos de I+D o en servicios técnicos.
- Transferencias de resultados de investigación a través de las OTRIs.

El trabajo de las unidades de documentación se desarrolla fundamentalmente en el ámbito de la investigación tecnológica, y el perfil y las necesidades de los usuarios a los que se da servicio tienen unas características propias, distintas a las de entorno académico.

La búsqueda intensiva de información tiene un peso muy importante en el desarrollo de los proyectos de investigación que se llevan a cabo en los centros. Parte de esa información o documentación es adquirida a través de las relaciones de los investigadores con la comunidad científica, por la participación en congresos, la pertenencia a sociedades profesionales y, por supuesto, por la pertenencia a grupos de investigación. Pero otra parte muy importante ha de obtenerse en las fuentes de información más formales: libros, manuales, artículos, informes...

El seguimiento de las necesidades específicas del colectivo de investigadores es crítico para que los centros de documentación sean capaces de desarrollar los servicios adecuados, en el formato adecuado. De nuestra capacidad de respuesta y

de anticipación a esas necesidades depende en gran medida conseguir que la organización nos considere como un servicio de valor para la consecución de sus objetivos estratégicos.

En este trabajo expondremos tres experiencias desarrolladas por los centros de documentación de Azti, Labein y Robotiker en esta línea, relacionadas con la identificación de los mejores recursos de información disponibles, el aprovechamiento óptimo de estos recursos por parte del usuario mediante la formación, y el diseño de un sistema avanzado de gestión del conocimiento.

LA EXPERIENCIA DE AZTI-TECNALIA: LA SELECCIÓN Y OPTIMIZACIÓN DE RECURSOS DE INFORMACIÓN DE CALIDAD PARA UN COLECTIVO DE INVESTIGADORES

En 1981 se puso en marcha el Servicio de Investigación Oceanográfica (SIO) en San Sebastián. Tres años más tarde, la Diputación Foral de Bizkaia autoriza la constitución del Arrantzuarekiko Zientzia eta Teknika Ikaskundea, A.B. (A.Z.T.I.) - Instituto para la Ciencia y Tecnología Pesquera, S.A. en Sukarrieta (Bizkaia).

En 1988, la Administración de la Comunidad Autónoma Vasca adquiere la Sociedad de la Diputación Foral e integra ambas instituciones, dando origen a la empresa pública AZTI-SIO.

Integrado en la Corporación TECNALIA en 2003, AZTI-Tecnalia es un Centro Tecnológico con forma jurídica de Fundación privada sin ánimo de lucro cuyo objetivo es contribuir al desarrollo social y económico de los sectores pesquero y alimentario, mediante la Investigación e Innovación Tecnológica.

Cuenta con un equipo humano de más de 220 profesionales, desarrollando su actividad de investigación alimentaria y marina en 3 sedes: Sukarrieta y Derio (Bizkaia) y Pasaia (Gipuzkoa). Cuenta con aproximadamente 12.000 m² de espacio para laboratorios, planta piloto y otras instalaciones necesarias para el desarrollo de su actividad.

La constitución del Servicio de Documentación de AZTI data de 1988 siendo su objetivo principal «la selección y difusión del información en el campo de la ciencia y tecnología pesquera». Teniendo como visión convertirse en un centro de referencia y divulgación para el sector pesquero.

Tradicionalmente las tareas fundamentales del Servicio de Documentación de AZTI-Tecnalia no han diferido mucho de las habituales en otros Centros de Documentación: selección, adquisición, análisis documental, difusión. Toda esta serie de actividades nos habían hecho focalizar nuestra atención en el documento, convirtiéndose en eje principal de nuestro cometido, creyendo erróneamente que únicamente el conocimiento experto de las fuentes de información, podría permitirnos alcanzar el objetivo primordial: satisfacer las demandas informativas de nuestros investigadores.

Pero la experiencia nos ha demostrado que para alcanzar ese objetivo, o por lo menos acercarnos lo máximo a él, la piedra angular ha de ser el usuario, cono-

cer lo que quiere, interpretar adecuadamente sus necesidades de información, ya que únicamente de este modo podremos llegar a recursos de información pertinentes que resuelvan sus demandas de información.

Por lo tanto, la creación del fondo documental de nuestro Centro de Documentación se basa, y tiene su origen, en la necesidad o demanda de información del usuario. Después de su adecuada identificación e interpretación, nos permite seleccionar recursos de información idóneos que satisfagan la necesidad informativa, aplicando nuestro conocimiento especializado en fuentes de información, proveedores de estas fuentes etc.

Por ello, consideramos que el éxito de nuestro Servicio de Documentación se basa fundamentalmente en la suma de dos conocimientos indisolubles; conocimiento del usuario (necesidad de información) + conocimiento experto de las fuentes de información, la suma de ambos generará un producto de información que resuelva la necesidad de información (objetivo final).



Para conocer al usuario, identificando y entendiendo sus necesidades, hemos de potenciar y desarrollar el contacto con nuestros usuarios, estableciendo cauces o vías de comunicación adecuados, valiéndonos de todas las posibilidades que nos ofrecen actualmente las tecnologías de la información y comunicación, como por ejemplo: el correo electrónico, Intranet (Web Servicio de Documentación), programas como el Skype, etc.

Pero teniendo muy claro que no podemos relegar el contacto personal, directo. Es fundamental acercarse al usuario, hablar con él, no sólo para conocer sus demandas concretas de información, sino también porque un encuentro con nuestros usuarios implica un intercambio de información, de conocimientos, que redundará en beneficio de la actividad de ambos.

Una experiencia en este sentido del Servicio de Documentación de AZTI, la cual ha tenido una buena acogida por parte de los investigadores del centro, es la participación activa en los proyectos de investigación, en lo concerniente a la búsqueda bibliográfica y captura de información relevante para el proyecto. Esta integración permite por un lado, identificar de un modo más preciso las necesidades informativas derivadas de los proyectos y por otro, también propicia un claro acercamiento a nuestros usuarios, al vernos como colaboradores, o participantes de su equipo de trabajo.

Las necesidades de información de los usuarios que identificamos en nuestro SD podríamos tipificarlas en dos grandes grupos, atendiendo a su origen y procedencia:

1. **Directas y específicas:** Procedentes de los investigadores que tienen demandas de información concretas.

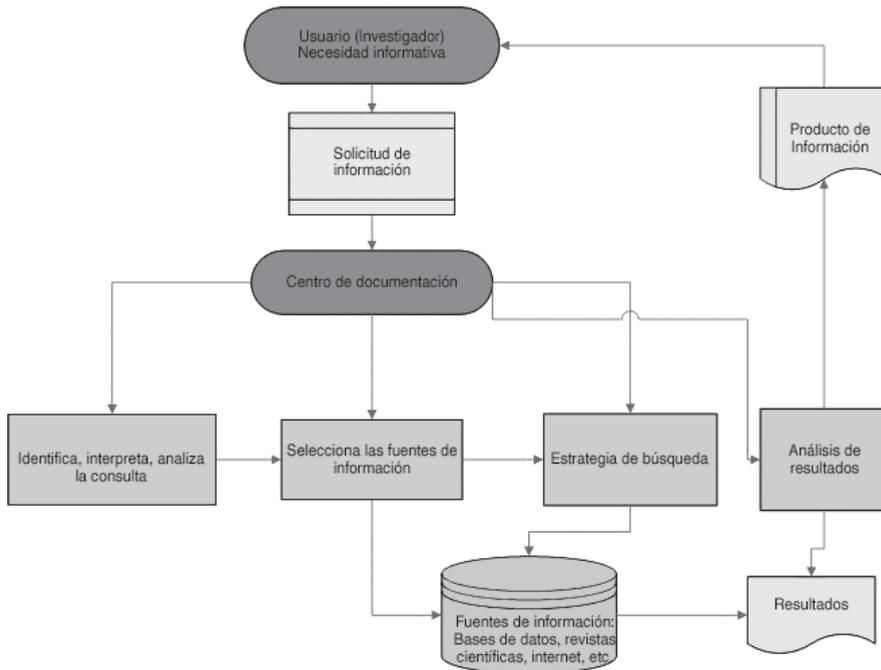


Diagrama de Flujo de una necesidad de información directa

2. **Indirectas y generales:** Proactividad, no esperamos a recibir una consulta puntual, sino que nos anticipamos a esa necesidad.

- a) **Seguimiento, captación y difusión de diferentes tipos de información**, en los que hemos detectado previamente que existía una demanda de información general.

Los parámetros que hemos de tener en cuenta para diseñar una estrategia adecuada de un seguimiento de información son los siguientes: perfil de usuarios potenciales (consumidores de la información); tipo de información que requieren; identificación de fuentes a revisar; establecimiento de la periodicidad del seguimiento; canales de difusión; responsables de la misma.

Tipo información	Responsable seguimiento	Fuentes	Frecuencia	Canal de difusión	Usuario final
Legislación técnica	Servicio de Documentación	Boletines oficiales (BOE, Doce, etc.)	Diaria	Intranet	Unidades de investigación
Subvenciones del Estado, autonómicas y provinciales	Servicio de Documentación	Boletines oficiales (BOE, BOPV, etc.)	Diaria	Intranet, Web corporativa	Unidades de investigación, empresas e instituciones
Concursos	Área comercial y Servicio de Documentación	Maninvest (Empresa especializada)	Diaria	Email	Área comercial y unidades de investigación
Eventos externos (congresos, jornadas, etc.)	Servicio de Documentación, Marketing	Internet, revistas especializadas	Semanal	Intranet, Portales especializados (Alimentatec), boletines sectoriales	Unidades de investigación, empresas, investigadores externos
Prensa	Servicio de Documentación	Prensa online, empresa especializada	Diaria	Web corporativa	AZTI-Tecnalia, sectores, sectores de actuación
Normas (UNE, ISO, etc.)	Servicio de Documentación	Aenor, ISO, ENAC	Semestral	Intranet	Laboratorios, área de calidad

Tabla de seguimiento y difusión de información

b) **Vigilancia tecnológica.**

El Servicio de Documentación de AZTI participa en labores de vigilancia tecnológica, especialmente en la Unidad de Investigación Alimentaria, detectando, analizando y difundiendo información científico-tecnológica, con el objetivo de que este conocimiento nos permita aprovechar oportunidades, y poder hacer frente a las amenazas del sector, facilitando la toma de decisiones estratégicas.

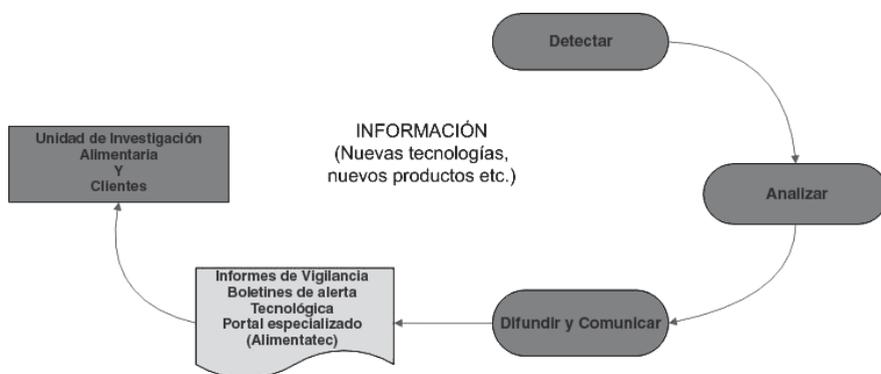


Diagrama de Vigilancia Tecnológica

- c) **Creación de productos de información específicos.** En este caso somos creadores de contenidos, previamente identificamos necesidades de información, las cuales no pueden ser cubiertas con productos existentes en el mercado, por lo tanto, es nuestro Servicio de Documentación el que soluciona ese vacío de información elaborando un productos de información. Ejemplo: Base de datos de legislación pesquera.

Conclusión

En el entorno de la explosión de los contenidos digitales y de la sobre-exposición a la información, el Servicio de Documentación gana valor como experto en la selección de los recursos de información más adecuados para su centro. El conocimiento de las necesidades de información de nuestros investigadores es el factor clave a la hora de seleccionar recursos de información pertinente, y este conocimiento se adquiere a través de una actitud pro-activa, basada en una estrecha relación con el usuario. Es el Servicio de Documentación quien ha de tomar la iniciativa, utilizando todos los cauces de comunicación a su alcance, para que a través de los mismos pueda identificar de manera clara y precisa las demandas de información de su organización.

LA EXPERIENCIA DE LABEIN: FORMACIÓN EN CAPACITACIÓN INFORMACIONAL DE LOS USUARIOS

En 1955, por iniciativa de la Escuela de Ingenieros, se crean en Bilbao los Laboratorios de Ensayos e Investigaciones Industriales Leandro José Torrontegui, origen del actual centro tecnológico LABEIN, hoy parte de TECNALIA Corporación Tecnológica.

LABEIN es una fundación privada, sin ánimo de lucro, con seis unidades de orientación sectorial: Construcción, Energía, Medio Ambiente Urbano e Industrial, Sistemas de Innovación, Automoción y Siderurgia. Cuenta con 3 sedes en Bizkaia, y una plantilla de 347 personas (a 31/12/07).

En 1979, el Reglamento del centro tecnológico mencionaba expresamente entre su funciones la «*Gestión de un Centro de Documentación científica y técnica, con fondo bibliográfico propio e intercambio con otros Centros de esta naturaleza, tanto nacionales como extranjeros, incluso a través de sistemas de información automatizada*».

De hecho, como afirmamos en la introducción de este trabajo, contar con un servicio de documentación o una biblioteca especializada ha sido algo consustancial a la existencia de los propios centros tecnológicos. Pero la naturaleza de su trabajo ha conocido una importante evolución en estos últimos 20 años.

En los años 90 las principales tareas —las de mayor valor añadido— de los centros de documentación eran las relacionadas con el acceso al documento primario, las suscripciones a revistas, el acceso a fuentes de información especializada y las búsquedas de información en bases de datos.

Respecto a las búsquedas de información en bases de datos de ciencia y tecnología, se trataba de un trabajo muy especializado, que recaía casi de manera exclusiva en los documentalistas. Recordemos que cada distribuidor o host (DIALOG, DATA-STAR, ESA, ORBIT, DIMDI) tenía su propio protocolo de comunicación y lenguaje interrogación, (CCL, DIALOG, UNIX...). Se trataba de recursos muy costosos y complejos, que tarifaban por conexión, tiempo y visualización de resultados; a los que se accedía por contrato y con un usuario único... y que sólo ofrecían información referencial.

Dos décadas después, la explosión de los contenidos digitales y la progresiva orientación de los servicios de información online hacia el usuario final, han «destronado» a los documentalistas y han provocado que el acceso a la documentación y la información se encuentren mucho más dispersos en la organización. El centro de documentación no es ya el centro esencial de las actividades de documentación en las organizaciones, y ya no son sólo los documentalistas quienes gestionan estos contenidos.

Internet es hoy nuestra principal fuente de información para todo. Hemos experimentado una drástica reducción de las demandas de información estructurada (bases de datos), y consecuentemente una reducción del uso de los recursos documentales tradicionales (biblioteca, revistas impresas y/o electrónicas...)

Dentro de las organizaciones en las que trabajamos, se constata un cambio cultural propiciado por Internet: los investigadores son mucho más autónomos y tienen mayor libertad para localizar la información directamente. Si embargo, esa mayor independencia ha propiciado situaciones como:

- Incapacidad de recuperar la información recopilada en el sistema de información de la organización.
- Desconocimiento de las capacidades de los sistemas de gestión documental corporativos.
- Infratilización de los portales y las bases de datos internas.

Por otra parte, frente a a las bases de datos estructuradas, Internet ofrece acceso a una ingente cantidad de información, pero carece hoy por hoy de herramientas de interrogación que aseguren la relevancia y la calidad de las búsquedas y se dedica mucho más tiempo a buscar linformación.

Por todo ello, a los centros de documentación se les plantea un doble reto:

- Asegurar que los investigadores aprovechen al máximo los recursos de información disponibles en los centros de documentación.
- Ayudarles a mejorar la calidad y la exhaustividad de sus búsquedas, disminuyendo el tiempo empleado en las mismas.

Se dice que vivimos en la era de la economía de la atención, y esta realidad afecta también a los objetivos del Centro de Documentación. Ante la avalancha de información que reciben nuestros usuarios ya sólo a través del correo electrónico, ¿cómo conseguir que nos presten atención?

La solución adoptada por el Centro de Documentación de LABEIN ha sido poner en marcha una actividad típica de las bibliotecas universitarias, un programa de formación de usuarios, con el objetivo de mejorar sus «competencias informacionales».

La Competencia Informacional —lo que en inglés se conoce como *Information Literacy*— es un concepto nacido en el ambiente académico y educativo anglosajón a finales de los 90. Es bien conocido en nuestra Universidad, pero es aun novedoso en el ámbito de empresa.

En castellano se utilizan diversas denominaciones, «cultura informacional», «habilidades informativas», «alfabetización informacional», «competencia en el manejo de la información»... para un mismo concepto:

Conjunto de procedimientos, conceptos y valores necesarios para resolver problemas que impliquen la búsqueda, selección, organización, análisis y comunicación de la información.

En otras palabras, se trata de desarrollar un tipo de competencias de carácter horizontal vinculadas al uso efectivo de la información.

Nuestro programa de capacitación informacional abarca tanto contenidos como herramientas vinculadas a las nuevas tecnologías de la red, como por ejemplo las aplicaciones Web 2.0. Para impartir las acciones formativas se ha optado por el formato presencial de «mini píldoras formativas». Son sesiones cortas (máximo dos horas) que no interrumpen en exceso el ritmo de trabajo de una jornada, con una orientación muy práctica y mínimamente teórica.

La formación se acompaña de una documentación en forma de guía rápida a la que el usuario, una vez ya en su mesa de trabajo, puede recurrir cuando lo necesite. Estas guías están disponibles también en la Intranet y están permanentemente actualizadas.

La formación se ofrece a toda la plantilla, con un calendario anual en el que se repiten los diferentes módulos. Los módulos se imparten también a medida, fuera de calendario, para grupos de trabajo que lo demanden.

<p>Buscar y entender la información sobre patentes</p>	<ul style="list-style-type: none"> – La propiedad industrial y su importancia como fuente de información técnica y de mercado. – Cómo se busca la información sobre patentes. – Cuáles son las principales bases de datos de patentes gratuitas en Internet.
<p>Estrategias avanzadas de búsqueda en Internet</p>	<ul style="list-style-type: none"> – Profundizar en las posibilidades y limitaciones de Internet como fuente de información. – Mejorar hábitos y prácticas de búsqueda en Internet. – Referenciar los resultados científicamente.
<p>Para los fans de Google</p>	<ul style="list-style-type: none"> – Ventajas y limitaciones de Google frente a otros buscadores – Cómo mejorar los resultados de nuestras búsquedas utilizando a fondo sus posibilidades. – Aplicaciones especiales: IGoogle, Google Reader, Alertas, Docs, Books, Academics...
<p>Herramientas 2.0 para la Vigilancia Tecnológica</p>	<ul style="list-style-type: none"> – Herramientas gratuitas en Internet aplicables al trabajo, los proyectos y las actividades de vigilancia. – RSS, blogs, wikis, Google Docs, páginas de inicio, marcadores sociales. – Las redes sociales y su aplicación en I+D.
<p>WOK: Web of Knowledge</p>	<ul style="list-style-type: none"> – Conocer el contenido y las características de las bases de datos y servicios de información disponibles en la Web of Knowledge. – Mejorar las estrategias de búsqueda en la WOK. – Personalización de acceso y configuración de alertas.
<p>ScienceDirect</p>	<ul style="list-style-type: none"> – Conocer el contenido y las características de las bases de datos y servicios de información disponibles en Science Direct. – Mejorar las estrategias de búsqueda. – Personalización de acceso, configuración de alertas.

Conclusión

La experiencia ha tenido una excelente acogida en términos de participación, y para el Centro de Documentación ha supuesto abrir un canal de comunicación muy activo con los usuarios, mucho más efectivo que otras iniciativas realizadas anteriormente (charlas de presentación, campañas de difusión de servicios a través del correo electrónico, encuestas de satisfacción, etc.).

A través de estos mini-cursos tenemos la oportunidad de conocer mejor a nuestro cliente interno: sus hábitos de búsqueda de información, las dificultades más habituales con las que se encuentra, sus carencias. Pero además, las sesiones

de formación se han convertido en foros de intercambio de experiencias y conocimiento entre los propios investigadores.

Esta iniciativa ha contribuido a orientar mejor nuestras capacidades para ayudarles en los problemas relativos a la saturación de información, el desconocimiento de recursos y fuentes de información —tanto internas como externas— y en la utilización productiva de tecnologías de la información que están en constante evolución y cambio.

LA EXPERIENCIA DE ROBOTIKER-TECNALIA: UN MODELO DE GESTIÓN AVANZADA DEL CONOCIMIENTO

Integrado en la Corporación Tecnológica TECNALIA, ROBOTIKER-Tecnalia es un centro tecnológico especializado en Tecnologías de la Información y Comunicaciones (TICs), y en Desarrollo de Producto, que tiene como objetivo contribuir activamente al desarrollo sostenible de la sociedad a través de la investigación y la transferencia tecnológica.

Es una fundación privada, sin ánimo de lucro, en la vanguardia de la investigación, el desarrollo tecnológico y la innovación (I+D+i). Actúa en sus mercados de referencia a través de cinco Unidades de mercado carácter Sectorial: Telecom, Energía, Automoción, Infotech e Innova. Estas cinco Unidades operan desde dos sedes: Zamudio (Bizkaia) y Madrid.

El Centro Tecnológico es un proveedor integral de I+D+i bajo contrato (Investigación + Desarrollo + innovación), que dispone de una completa gama de servicios y productos que van desde la vigilancia y la prospectiva hasta la creación de empresas de base tecnológica. En ese amplio abanico de modalidades de colaboración con las empresas, destacan las de desarrollo de proyectos de I+D y los servicios de consultoría tecnológica.

Cuenta con una plantilla de 203 personas.

COPLA: la Intranet del conocimiento

En el año 2000, ante la creciente necesidad de coordinar y gestionar la generación, organización y distribución de la información, se planteó en ROBOTIKER-Tecnalia la creación de una «Intranet del conocimiento». El proyecto se denominó COPLA, acrónimo de gestión del CONocimiento e imPLAntación.

Partía de un análisis previo de los procesos clave, con sus carencias en cuanto a la gestión del conocimiento. Ha sido resultado de la ejecución interna de la metodología de ROBOTIKER-Tecnalia para la implantación de los procesos de gestión del conocimiento.

El objetivo de COPLA era conectar la oferta y la demanda de conocimiento interno. En concreto, se trataba de dar respuesta y soporte a las siguientes actividades:

- Vigilancia tecnológica interna (creatividad).

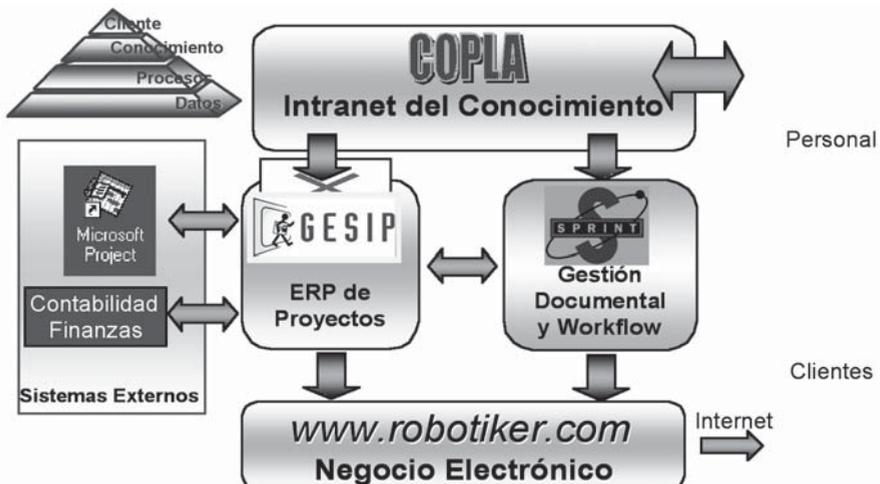
- Vigilancia tecnológica externa.
- Mapa de conocimiento.
- Gestión de relación con clientes.
- Organización interna.

Para ello se trabajó en la integración de dos productos ya existentes y también desarrollados internamente: GESIP y SPRINT.

- **GESIP:** Es la herramienta que permite planificar y controlar los proyectos en curso y gestionar los clientes y recursos de la propia organización. A partir de estos datos, genera informes de gestión o cuadros de control con los que monitorizar tanto la actividad comercial como la de proyectos y la económica.
- **SPRINT:** Es una Intranet, una herramienta, diseñada para gestionar la información y los flujos de trabajo (*workflows*) del centro, es decir, su Conocimiento. Su objetivo es la integración y coordinación del trinomio clave: personas, tareas e información. Para ello debe:
 - Gestionar todos los documentos importantes para la organización.
 - Organizar y automatizar la actividad diaria en función de *workflows*.
 - Capturar e integrar datos de otras bases de datos relacionales estándar.
 - Integrar el lanzamiento de aplicaciones desde un único interface común para todos.

COPLA permite gestionar todos los documentos, conectar con aplicaciones y datos de cualquier otro sistema de información de la organización y además, ordena y automatiza los procesos de trabajo ya sistematizados que involucran a las personas.

El gran logro de COPLA es que integra perfectamente a SPRINT y GESIP en un único mecanismo de comunicación como se puede observar en el siguiente gráfico:



A través de la Intranet COPLA se puede acceder a la información clave sobre el centro tecnológico —misión, visión, valores, cultura, procedimientos, formas de trabajo, etc.—, al mapa del conocimiento y a las herramientas corporativas. Así mismo, contiene la información sobre las actividades que realizan tanto los departamentos técnicos —ofertas, proyectos, cursos, ponencias, seminarios, artículos, etc.— como los departamentos de apoyo: Centro de Documentación, RRHH, Calidad, etc.

En SPRINT, todos los servicios del Centro de Documentación están automatizados. Fueron diseñados y programados a medida, a partir de un estudio previo de necesidades. Se compone de Biblioteca, Hemeroteca, Archivo, Base de datos de Normas, Patentes y Recursos especializados. Mucha de la información susceptible de ser buscada está en formato digital, por lo que el usuario no tiene necesidad de acercarse al espacio físico que ocupa el Centro de Documentación.



Tras nueve años utilizando COPLA podemos constatar que sus principales características son:

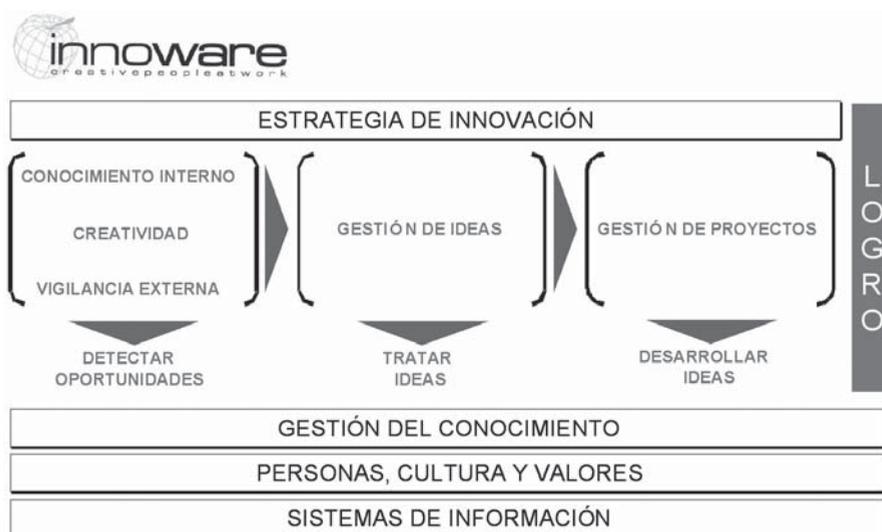
- Es una aplicación concebida para facilitar el acceso de cada persona al conocimiento que requiere en cada momento concreto para realizar su labor.
- Contribuye al aumento y desarrollo del conocimiento.
- Mejora la productividad y la competitividad al aprovechar los conocimientos adquiridos y trabajos realizados para proyectos posteriores.

INNOWARE e INNOSOFT: Modelo y herramientas para la Gestión de la Innovación

El estudio y la experiencia de COPLA fueron tan satisfactorios que, en el año 2005, se decidió ir un paso más allá y abarcar la gestión de la ya tan exaltada y alabada innovación. Para ello se desarrolló, a partir de COPLA, una metodología y un producto comercial: Innoware e Innosoft.

INNOWARE es un modelo de gestión que abarca todas las facetas del proceso innovador, desde la planificación tecnológica hasta el desarrollo de los nuevos productos y/o servicios, pasando por la vigilancia y prospectiva tecnológica y la gestión del conocimiento. INNOWARE se soporta sobre INNOSOFT tecnología propia de ROBOTIKER-Tecnalia para hacer más fácil el proceso innovador.

Innoware marca una estrategia de innovación que se plasma visualmente a través del siguiente gráfico:



ACK 2009: Access to Corporate Knowledge

2009, el año europeo de la Innovación y la Creatividad, marca un nuevo reto en el ámbito de la gestión del conocimiento para ROBOTIKER-Tecnalia, debido a que:

- Está inmerso en un proceso de fusión con otros centros Tecnalia.
- Existen unidades multisede: investigadores de una misma unidad se encuentren dispersos en oficinas ubicadas incluso en diferentes provincias.

- Cuenta con una importante experiencia previa en tecnologías semánticas, gracias a los proyectos europeos: WORKSAFE, EUROWORKSAFE y AQUARING y proyectos a nivel nacional y local.
- La metodología Innoware le permite aplicar la estrategia de innovación a los sistemas de información.

Tecnalia esta organizada en Unidades de Negocio —veintidós hasta el momento—, que son a su vez unidades estratégicas de información. Lógicamente, existen usuarios que empiezan a demandar un replanteamiento de la gestión documental, pero no sólo del centro, sino de todo Tecnalia. También lamentan que no sean capaces de conocer, con las herramientas actuales, qué hacen o hicieron sus compañeros, ni quién es experto en qué, lo que consideran esencial para agilizar su trabajo.

Más difícil todavía, cómo saber: ¿cuántos documentos pueden descargarse durante la elaboración de una propuesta o de un documento de proyecto?, ¿dónde los guardan?, ¿cómo recuperarlos sin necesidad de utilizar nuestra buena memoria?, ¿cuántas veces nos hemos bajado más de una vez el mismo documento?, ¿cuál era aquella Web tan buena?, ¿quién ha desarrollado ya tal funcionalidad utilizando tal plataforma?... Es evidente y urgente la necesidad de un contacto y acceso más directo al origen de la información, al conocimiento interno.

Como anécdota decir, que nadie parecía atreverse a mencionarlo a su director de unidad, pero el concepto «red social Tecnalia» como posible medio de comunicación interno estaba en todas sus mentes y en los corrillos de la máquina de café.

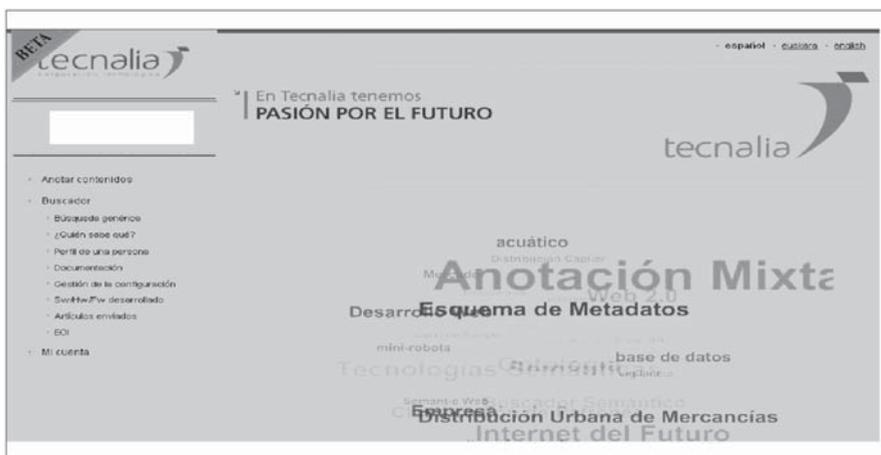
Una novedosa herramienta de extracción de conocimiento corporativo, haciendo uso de las tecnologías semánticas, podía ser la solución, si se complementase con nuestra metodología y experiencia en la gestión del conocimiento e innovación.

A principios de año comenzó un proyecto interno dentro de la Unidad de Infotech, ACK, una herramienta de acceso a la información corporativa, que como hemos mencionado es resultado de las diferentes situaciones a las que la plantilla de ROBOTIKER-Tecnalia se enfrenta todos los días, y que han alcanzado mayor envergadura tras la creación de Tecnalia y el actual proceso de fusión.

Existe un prototipo, a modo de Intranet, que ya esta utilizándose dentro de la propia Unidad de Infotech.

Algunas de sus características son las siguientes:

- Su uso es tan fácil e intuitivo como el de un buscador web.
- Ofrece un único punto de acceso a toda la información, manteniendo la operativa actual de la empresa.
- Permite crear «perfiles de conocimiento», obtener la respuesta a «¿quién sabe qué?».
- Facilita herramientas en aquellos casos en los que no se disponga de ellas, por ejemplo, para clasificar y recuperar documentos, presentaciones realizadas, urls, contactos que actualmente cada integrante del personal hace de forma individual.



Para una mejor comprensión del funcionamiento de esta herramienta, vamos a detenernos brevemente en el concepto y terminología de las tecnologías semánticas de las que se nutre.

Un buscador semántico se diferencia de un buscador convencional en que realiza sus búsquedas sobre anotaciones semánticas, es decir, aquellas anotaciones realizadas utilizando una ontología del dominio.

Las anotaciones nos permiten agregar recursos heterogéneos: documentación digital, procesos, e incluso personas, independientemente de su ubicación, formato e idioma utilizado para su creación.

La ontología define conocimiento consensuado sobre un dominio, lo que permite dotar de significado a los recursos mediante la anotación. Dicho significado junto con el conocimiento en el dominio —la ontología— son los recursos utilizados por el buscador semántico para obtener resultados.

La herramienta desarrollada por ROBOTIKER-Tecnia es independiente del modelo de metadatos utilizado para la anotación y de la ontología definida. Además, es multilingüe.

Los resultados obtenidos a través del buscador semántico se componen, por un lado, de los recursos en sí mismos y, por otro, de ayuda al refinamiento o recorrido por dominio relacionado. Nos permite afinar tanto, que el «ruido documental» es apenas imperceptible.

Conclusión

Es nuestra esperanza que pronto sea una herramienta aplicable a todo Tecnia, y que llegue a ser, más que una Intranet, un Centro de Documentación online. Pero concretamente nos gustaría poder hablar de un «Centro de Conocimiento» que, más allá de convertirse en un gran repositorio de información, sea capaz de gestio-

nar y tratar el conocimiento interno, las ideas, las personas (los perfiles de conocimiento), los valores, la creatividad... intangibles que gracias a la metodología de Innaware y a las posibilidades de ACK puede llegar a ser una realidad en un futuro no muy lejano, ¿quizás para 2010?



BIBLIOGRAFÍA

- Wilson, T. D. (2000). «Recent trends in user studies: action research and qualitative methods». *Information Research* 5 (3). [artículo en línea]. < <http://informationR.net/ir/5-3/paper76.html> > [Fecha de consulta: 11/03/09].
- Calva González, J. J. (1994). *Las necesidades de información: la importancia de su estudio*. Biblioteca Universitaria IX, (3) [artículo en línea]. < <http://www.dgbiblio.unam.mx/servicios/dgb/publicdgb/bole/fulltext/volIX3/necinfo.html> > [Fecha de consulta 13/3/2009].
- Ortoll, E. (2003). *Gestión del conocimiento y competencia informacional en el puesto de trabajo* [artículo en línea]. UOC < <http://www.uoc.edu/dt/20343/index.html> > [Fecha de consulta: 27/03/09].
- Macoustra, J. (2003). «Information Literacy in a Corporate Environment». *FreePint Newsletter* 132 [artículo en línea] < <http://www.freepint.com/issues/060303.htm#feature> > [Fecha de consulta: 27/03/09].
- González, M. (2004). Web semántica «Tecnologías de la futura Web» [Documentación Jornada Técnica] Zamudio: Tecnia-Infotech, 2004.
- Martínez, A. *Gestión de flujos de trabajos + gestión documental avanzada = SPRINT* (2000). [artículo en línea]. < <http://revista.robotiker.com/revista/articulo.do?method=detalle&id=3> > [Fecha de consulta: 27/03/09].

VII

La evolución del concepto de gestión documental en las organizaciones potenciada por la digitalización de los documentos y las herramientas de colaboración. El caso práctico de Bahía de Bizkaia Gas

Ainhoa Arpide Etxano

Responsable del departamento de compras y gestión documental
Erosketen eta dokumentazio kudeaketaren saileko arduraduna

Resumen: Las oportunidades que ofrece la digitalización de los archivos han posibilitado un cambio de mentalidad en los usuarios, así como un cambio de planteamiento en la gestión de la documentación dentro de las organizaciones. Del archivo estático de referencia, se va evolucionando a un archivo vivo, colaborativo, y que reproduce la transversalidad de las necesidades de la organización, que si bien en estos momentos todavía no conlleva la desaparición del archivo tradicional en papel, puede implicar incluso dicha posibilidad a futuro. El archivo no es el punto final de la gestión documental sino que convive con la actividad de la organización en tiempo real.

Laburpena: Dokumentuen digitalizazioak eta lankidetzak tresnek eraginda, erakundeetan dokumentazio kudeaketaren kontzeptuan izandako bilakaera: BBG-ren kasu praktikoa. Agiritegien digitalizazioak eskaintzen dituen aukerak erabiltzaileen pentsamoldearen aldaketa ahalbidetu du, baita, erakundearen barruan, agiriak kudeatzeko moduaren aldaketa ere. Erreferentziazko agiritegi estatistikotik agiritegi bizi eta lankidetzazko batera aldatuz doa, eta erakundearen beharrezko transbertsalitatearen isla da. Horrek oraingo ez dakar paperezko ohiko agiritegiaren desagertzea, baina etorkizunean baliteke hala gertatzea. Agiritegia ez da kudeaketa dokumentalaren azken gunea, baizik eta erakundearen jarduerarekin bat eginda aritzen da, denbora errealean.

QUÉ ES BAHÍA DE BIZKAIA GAS

Bahía de Bizkaia Gas, S.L. (BBG), constituida por cuatro socios (EVE, BP, IBERDROLA y REPSOL), tiene por objeto la explotación comercial de una planta de recepción, almacenamiento y regasificación de gas natural licuado. Para ello dispone de dos tanques de almacenamiento de 150.000 m³ cada uno y una capacidad de regasificación de 800.000 m³ (n)/h.

Este gas natural, que llega a BBG mediante barcos metaneros, se destina al consumo doméstico, comercial e industrial, incluyendo en este último el suministro a ciclos combinados de generación de energía eléctrica, como es el caso de la planta de ciclo combinado limítrofe, de Bahía de Bizkaia Electricidad (BBE).

EL COMIENZO

Hacia 2001 comienza la construcción de la planta de recepción, almacenamiento y regasificación de BBG. En estas fechas, la oficina de la propiedad comienza a recibir un gran volumen de documentación técnica correspondiente al proyecto de construcción de la planta, y es entonces cuando surge la necesidad de que un documentalista gestione dicha documentación.

Hasta ese momento la documentación que se genera es responsabilidad del área que la produce y gestiona. El criterio de organización documental depende por tanto del responsable de esa documentación, al igual que su gestión y custodia.

Existe un pequeño archivo común de aquella documentación considerada «importante» pero no hay un concepto de archivo en un sentido más global. El hecho de existir varias personas que gestionan la documentación de acuerdo con criterios individuales, supone la existencia de multicriterios o incluso de falta de criterio, y ello implica que es difícil establecer trazabilidad en la documentación manejada.

En 2001 se contrata a un documentalista para que controle la llegada de la documentación técnica del proyecto. En este momento existe y va creciendo un gran volumen de documentos que hay que recepcionar y controlar, tanto en cuanto a su mantenimiento como a su ubicación física. Para gestionar esto se genera una base de datos en Access, que ayuda a responder las consultas de los usuarios y a controlar la ubicación física de la documentación.

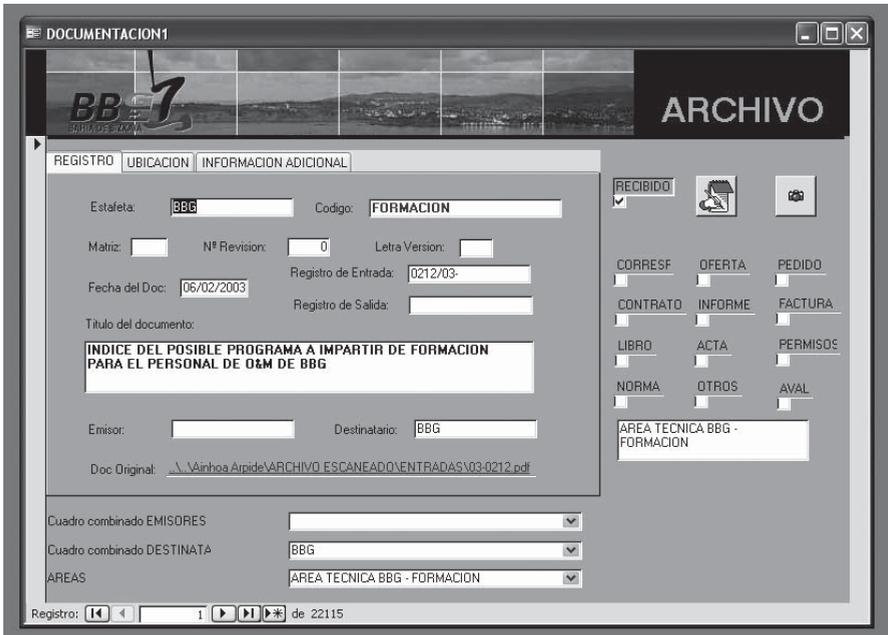
PRIMERA EVOLUCIÓN

Se decide que toda la documentación que se registra en los libros de entradas y salidas quede en Archivo, y que las personas trabajen con copias de dicha documentación. De esta manera se centraliza la gestión documental en un punto común a toda la organización.

El segundo paso es que la documentación administrativa que entra o sale de la organización comienza a escanearse y se introduce un enlace en la base de datos documental. De esta forma los usuarios no necesitan recurrir al papel más que en ocasiones puntuales. Sin embargo, la documentación sólo es accesible desde el puesto de Archivo. El usuario solicita el documento y éste se le envía por correo electrónico o se le hace una copia, dependiendo de la necesidad.

Nuevas necesidades

El 5 de diciembre de 2003 la planta comienza su explotación comercial y se reestructura como empresa.



La gestión documental también sufre su propia evolución:

- Cambio de fase de proyecto a fase de explotación.
- Creación de los nuevos departamentos.
- Generación de nuevos tipos de documentación.

Surge la posibilidad de implantar un nuevo gestor documental que va a dar nuevas posibilidades de gestión y un enfoque diferente. Se detectan cambios que provocan una novedosa concepción de la gestión documental y se abordan nuevos retos:

- Digitalización de toda la documentación desde registro.
- Posibilidad de consulta desde todos los PC de la planta.
- Sistema de política de accesos.
- Determinación de las nuevas tipologías documentales.
- Identificación de los nuevos expedientes.

PRIMERA CONSULTORÍA Y NUEVO GESTOR DOCUMENTAL

A mediados del 2004, se pone en marcha la primera consultoría documental dentro de la organización. Mediante entrevistas con los usuarios, se analiza el tipo de documentación que manejan y generan. Se identifican las series documentales existentes y las nuevas que comienzan a surgir en la actividad diaria de la planta,

se identifican nuevos expedientes; y con todo ello se cumplimentan unos formularios que nos ayudarán a hacer el mapa documental.

Una vez hecho este análisis se busca un nuevo gestor documental. Para la implantación se necesitará:

- Una aplicación comercial.
- Un sistema de captura de documentos en papel mediante escáner.
- La participación de todas las personas de la organización.
- Un cambio de mentalidad en la gestión de documentos.

El nuevo gestor permitirá:

- Un control exhaustivo de los documentos de la organización en todo momento.
- La incorporación de los documentos electrónicos al archivo.
- Potentes opciones de búsqueda.
- La consulta de los documentos en pantalla.
- La eliminación de copias.
- El control de versiones de los documentos.
- La eliminación de documentos en papel que no tengan valor legal.

Los objetivos a conseguir con esta nueva herramienta son:

- Recuperación de los documentos por sus perfiles, metadatos y contenidos.
- Integridad de los documentos. Sólo pueden ser modificados por los que tienen derechos, bajo determinadas condiciones y dejando rastro.
- Asignación de derechos y roles a nivel de personas.
- Controles de acceso a librerías, carpetas, documentos y partes de documentos.
- Trazabilidad. Todas las actuaciones sobre los documentos dejan una traza que se puede consultar.
- Control de versiones.
- Integración con otras herramientas: *workflow*, ERP, etc.
- Gestión de documentos complejos.

El proceso implica materialmente una serie de cambios:

- Un repositorio de documentación de gran potencia.
- Un escáner ágil.
- Un ordenador en cada puesto de trabajo.
- Trabajo en red.
- Reducción al mínimo de la documentación en papel.
- Ganancia de espacio físico.
- Reducción de gasto
- Impide el deterioro de documentos originales que sólo se presentan/utilizan en casos de necesidad.
- Integración total con las herramientas ofimáticas.
- Si se incluye un sistema web, posibilita el acceso remoto a la documentación corporativa.

Por otro lado, todo cambio en una organización conlleva un proceso de Gestión del cambio con las personas, en este caso los trabajadores que también ven modificada su forma de trabajo:

- Evita la saturación de papel en los despachos.
- Acceso rápido a la documentación en cualquier momento.
- Cada persona tiene un rol definido dentro del proceso de *workflow* y es «puerta» del trabajo de los demás.
- Es necesaria una persona que controle la documentación externa y la introduzca en el sistema.
- Permite la confidencialidad de la documentación a diferentes niveles.

Con toda esta modificación aún quedan retos por conseguir:

- Que cualquier usuario sea capaz de encontrar cualquier documento, sin conocer la estructura de la herramienta. Nos encontrábamos con que a los usuarios les era difícil localizar documentación ajena a sus departamentos.
- Que la entrada y salida de documentos del sistema no dependa de una única persona que alimenta la herramienta. Durante este periodo inicial si la persona de archivo se encontraba ausente, la documentación se quedaba pendiente de ser introducida en el sistema.
- Que sea posible el acceso desde el exterior. Tampoco se utiliza como una Intranet, a pesar de que técnicamente hubiera podido ser así.

NUEVO ANÁLISIS Y NUEVO GESTOR DOCUMENTAL

Durante el año 2007 BBG se vuelve a plantear la necesidad de dar un impulso a la gestión documental de la organización. La conciencia de los usuarios ha evolucionado, la documentación forma parte de BBG y el servicio que se demanda ha de ser más ágil que lo que en ese momento se ofrece. Por otro lado, las posibilidades técnicas también han evolucionado y es momento de plantearse otra vuelta de tuerca.

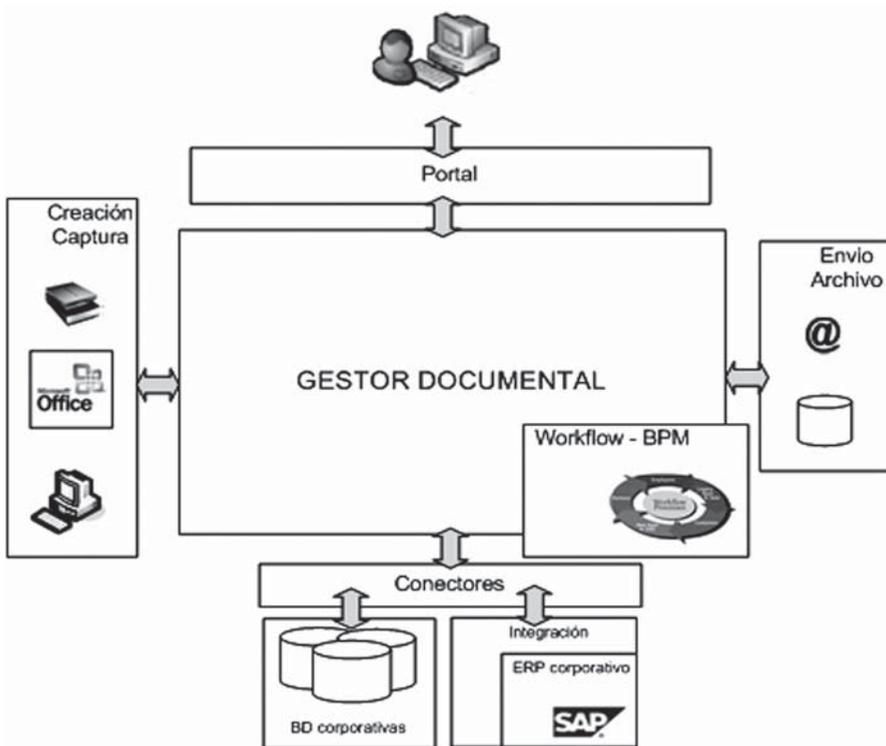
En este momento nos vamos a centrar en buscar:

- Una plataforma común que permita una gestión global y eficiente de la información.
- Digitalizar y mecanizar el proceso de registro de los documentos.
- Clasificar la información a partir del uso de tipologías documentales que permitirán jerarquizar y organizar el contenido.
- Mejorar la seguridad aplicada a la información. Control de acceso a la documentación basada en niveles de seguridad más concretos.
- Disponer de una plataforma que permita habilitar una Intranet o portal común.
- Disponer de una plataforma con posibilidades de integración con otras aplicaciones, que permita el acceso rápido y fácil a la documentación.

- Facilitar el acceso a la información y a las personas que los usuarios necesitan para la toma de decisiones desde la plataforma, fomentando la comunicación.

El nuevo planteamiento en este caso es más ambicioso y global que los anteriores. Se comienza a manejar no sólo el término documentación, sino también el de información. Se buscan:

- Funcionalidades de Intranet y servicio como Portal a los empleados de BBG.
- Gestión de los procesos de negocio documentales.
- Gestión global de la información documental.
- Integración con ERP Corporativo (SAP) y otras aplicaciones.



PROCESO DE MEJORA

Durante el análisis y la consultoría previos a la implantación del nuevo gestor se detectan determinados puntos que se pueden mejorar.

- El mapa documental se ajustaba a la documentación en formato papel. Era preciso diseñar un modelo de acceso rápido a los documentos electrónicos.

- Se propuso la taxonomía estructural de la empresa, que refleja las áreas de trabajo.
- La mejora fundamental del mapa existente se centraba en la necesidad de clasificar la documentación mediante:
 - Definición de tipologías documentales:
 - Agrupación de documentación en función de su naturaleza.
 - La documentación abandona el formato papel.
 - Creación de metadatos:
 - Son propiedades asociadas a un elemento de información.
 - Los metadatos permiten establecer consultas sobre la información.

El objetivo era que cualquier usuario fuera capaz de encontrar cualquier tipo de documentación en la herramienta, sin conocer la estructura en la que se encontrase ubicado el documento.

Por un lado, la taxonomía representa las funciones de las diferentes áreas y, por otro, los documentos son identificados con metadatos que ayudan a su rápida localización.

Una de las grandes evoluciones, aunque a primera vista parezca muy sencilla, es la decisión de que toda la documentación que pasa por el registro de documentos sea introducida en el sistema en ese mismo momento. El documento que entra en el sistema es guardado por un lado en el registro, bien de entrada o bien de salida, y queda así una primera copia del mismo, fiel reflejo del documento en el momento en que pasa por el registro. Por otro lado, y simultáneamente, este documento es redirigido a su ubicación en el sistema, en función del departamento que ha de trabajar con él o de su pertenencia a uno u otro proceso dentro de la organización. Para ello se automatiza el proceso de registro:

- Se automatizan los procesos encargados de registrar todas las entradas/salidas de documentación de BBG:
 - Se definen las tipologías documentales, los perfiles de usuario y los roles implicados en el proceso de registro de la documentación.
 - Se automatizan los medios de entrada de documentación al Gestor documental (correo electrónico, digitalización, importación, fax, etc.).
 - Se definen los medios de distribución de la información.
- Los documentos almacenados en el Gestor documental mediante el registro de entrada y salida toman entidad de tipo documental perteneciente a:
 - Una biblioteca.
 - Un expediente.
 - O un proceso formando entidades de expedientes virtuales.
- Se aplica la definición previa de los medios de distribución de la información.

Entrada de un documento externo.



Una vez situada en su departamento, la documentación queda pendiente de ser catalogada por el responsable de archivo, aunque también se da la posibilidad de catalogarla al responsable del departamento al que pertenece dicha documentación. De esta forma, su responsable tiene la posibilidad de ubicar el documento donde crea más conveniente; y a la hora de encontrar la documentación ésta será localizada más rápidamente.

En el departamento

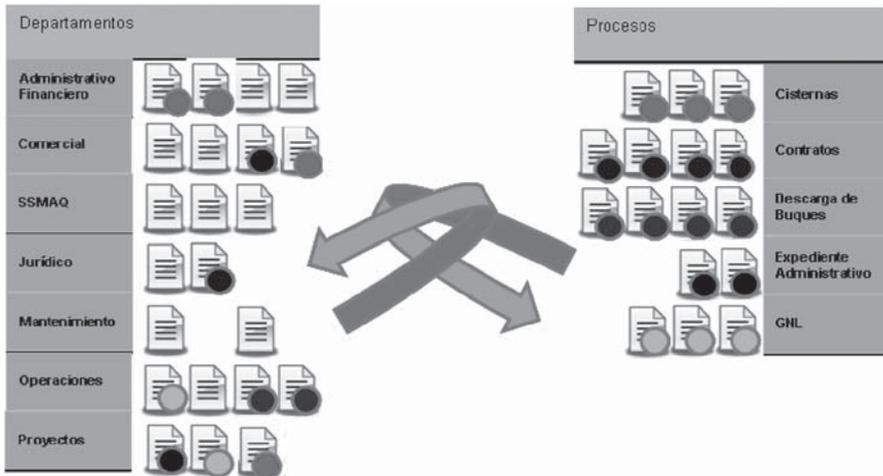


Una vez catalogado y asignadas todas sus propiedades y relaciones, será accesible en función de los permisos de cada usuario del sistema.

Los documentos dejan de ser un ente en sí mismo para formar parte de un proceso de la empresa. En el caso de los procesos transversales a varios departamentos, como puede ser un expediente de compra o un contrato, esta herramienta facilita la posibilidad de agrupar documentos que aún estando ubicados en diferentes

departamentos forman un conjunto común en función del proceso al que pertenecen. Éste será el gran avance conseguido en este punto.

Procesos. Permite relacionar un documento almacenado en un departamento con un proceso determinado. De este modo, cada usuario gestiona la información que le es afín a su departamento y, a la vez, permite que el resto de usuarios puedan acceder a conocer toda la información relacionada con un proceso, sin necesidad de navegar por cada uno de los departamentos



BBG7
BARRIO DE BETALES

Acciones del sitio

DEPARTAMENTOS PROCESOS DOCUMENTACIÓN ANTIGUA BÚSQUEDAS COLABORACIÓN GESTIONES ADMINISTRATIVAS

Avisos

Cisternas

Contratos

Descarga de Buques

Expediente Administrativo

Expediente Reclamación

Lista "Avisos Telefónicos".

Comun

Informe Mensual

Solicitud de ATR

Asunto	Emtido por	Firmado por	Fecha
Suministros			
Solicitudes de materia de Unión o moorario			02/04/2009
Incorporación de Becarios			23/03/2009
Redistribución de tareas			12/03/2009
Códigos de acceso a fotocopiadora			06/03/2009
Gestión de mobiliario, fotocopiadoras e impresoras.			13/02/2009
Consultas al Dpto de RRIH			13/02/2009
Procedimiento solicitud de vacaciones			13/02/2009
Comidas solicitadas a Punta Lucero			27/10/2008
Suspensión suministro de agua embotellada			05/09/2008
Cumplimiento Ley Orgánica de Protección de Datos			04/09/2008
NUEVOS EQUIPOS DE IMPRESIÓN			04/04/2008

Ausencias

Filtro Ordenar por: Vacío Agrupar por: Vacío

No hay elementos que mostrar.

Web Mail BBG

¿Quién está en la oficina?

Dossier de Prensa

Reserva de Salas

Sistema Integrado

Fotos Mozah

Solicitudes a informática

Título	Estado
Comprobar que puede acceder al Sharepoint desde el exterior	No iniciada
Acceso de marivi externo al MOSS	No iniciada

Los procesos dibujan en muchos casos el flujo del trabajo de la organización. Por ejemplo:

- Una solicitud de COMPRA se genera en un departamento, y por tanto genera un pedido.
- Existen además varias ofertas para efectuar el pedido.
- El departamento administrativo financiero genera, en el formato correspondiente, el pedido que será enviado al exterior.
- Se recibe el acuse de pedido por parte del proveedor.
- El albarán o resultado del servicio lo recibe mantenimiento.
- La factura llega a contabilidad.

No toda la documentación la genera el mismo departamento, pero sí es interesante que toda ella vaya referenciada al pedido, de manera que sea consultable con éste.

En el caso de un contrato, la documentación la comparte el departamento que solicita el servicio y lo gestiona en inicio, el departamento jurídico que participa en la gestión del documento contractual, el departamento de compras que ejecuta el pedido de compra para el servicio del contrato... Pero toda la documentación se agrupa en un punto común, aunque en el sistema esté ubicada en cada uno de los departamentos que la manejan.

Otra de las ventajas adquiridas es, por ejemplo, la integración del sistema con las herramientas ofimáticas habituales. Los documentos ofimáticos, los correos, pueden guardarse y son relacionables con cualquiera de los procesos, y entran a formar parte de la política de accesos en el momento en que son introducidos en el sistema desde cualquier puesto autorizado para ello.

Las posibilidades de búsqueda han sufrido también una fuerte evolución. La búsqueda avanzada es potente y permite buscar por cualquiera de los metadatos del documento. Igualmente se puede buscar por la ubicación del mismo o se puede navegar por el sistema.

COMO CONCLUSIÓN

Los logros conseguidos hasta ahora son:

- Que los documentos no formen parte sólo de un área sino que se integran en un todo que es la organización.
- Que cualquier persona autorizada pueda ver el expediente completo de un tema, aunque en su área sólo se haya trabajado parte del mismo.
- No hace falta conocer el diseño del resto de las áreas, ya que los procesos están diferenciados de la estructura de cada departamento.
- La documentación forma parte del sistema documental desde el momento de su creación o su entrada en la organización.

Funciones que vienen a dar valor añadido al sistema:

- El acceso desde el exterior mediante una dirección IP posibilita a las personas autorizadas el poder disponer de toda la información de la organización, independientemente de su ubicación física dentro o fuera de la organización.
- Los comunicados internos de la empresa se gestionan mediante un tablón general virtual. Teniendo en cuenta que un alto porcentaje del personal de nuestra organización trabaja a turnos y no posee una dirección de correo electrónico individual, esto posibilita su acceso a la información fuera de horarios de oficina.
- Área colaborativa de trabajo. Se ha generado en la herramienta un espacio en el que determinados usuarios pueden poner en común los documentos en los que están trabajando y, mediante un control de versiones y de autoría, realizar modificaciones con número de revisión, lo que les facilita el trabajo y les evita colapsar los correos electrónicos con pesados documentos que van y vienen.

Otros servicios informativos de utilidad son:

- La agenda de quién está y quién no está en la oficina.
- La reserva de salas de la organización para diferentes reuniones.
- Publicación de los dossieres de prensa diaria sobre temas energéticos, que se actualizan diariamente y son de utilidad para todos.
- Las solicitudes de material ofimático o las de apoyo informático que, en lugar de hacerse por correo electrónico o por teléfono, se realizan a través de la herramienta siendo así mejor su trazabilidad y seguimiento.

3. atala

Internet sarea artxibatzen

3.^a parte

Archivando la red Internet

VIII

Euskal hauteskundeetako webak Ondarenet-en

Francisca Pulgar Vernalte eta Sonia Marcos Maciá

Liburutegi Zerbitzuko arduraduna Eusko Jaurlaritzan
eta teknikari dokumentalista ODEI enpresan

Responsable del Servicio de Bibliotecas del Gobierno Vasco
y técnico documentalista de la empresa ODEI

Laburpena: 2009an Eusko Legebiltzarrerako hauteskundeetan parte hartu zuten alderdi politikoek web-ak berreskuratu, zaindu eta hedatzeko proiektuaren aurkezpena da komunikazio honen helburua. Halaber, «2009ko Eusko Legebiltzarrerako Hauteskundeak» bilduma osatuko duten baliabide elektronikoen hautaketa eta berreskuraketa deskribatzen da. Baliabide hauek edozein ikertzailek edota interesaturik dagoen edonork Ondarenet-en kontsultatu ahal izango ditu. Ondarenet Euskadiko ondare digitalaren fitxategia da, eta fitxategi elektronikoa hau Eusko Jaurlaritzaren Kultura Sailak jarri du martxan, formatu digitalean sortutako baliabideak berreskuratzeko, zaintzeko eta hedatzeko asmoz.

Hitz gakoak: Gordetegi digitalak; Liburutegi digitalak; Web fitxategiak, Zaintz.

Resumen: Las webs de las Elecciones vascas en Ondarenet. El objetivo de esta comunicación es presentar el proyecto de recuperación, preservación y difusión desde Ondarenet de las páginas webs de los partidos políticos que han participado en las elecciones al Parlamento Vasco en el 2009. Asimismo, se describe la selección y captura de los recursos electrónicos que conformarán la colección de «Elecciones al Parlamento Vasco 2009», y que podrán ser consultados por cualquier investigador o persona interesada desde Ondarenet, el archivo electrónico creado por el Departamento de Cultura del Gobierno Vasco para la recuperación, conservación y difusión de los recursos nacidos en formato digital.

Palabras clave: Repositorios digitales; Bibliotecas digitales; Archivos web, Preservación.

1. SARRERA

Agerikoa da gero eta hobeto ezagutzen eta menperatzen direla Internetek eskaintzen dituen aukerak, bai edukiak argitaratzeko edota hedatzeko ere. Edukiak gero eta gehiago zuzenean Interneten argitaratzen dira, hau da, paperean editatu gabe. Hori dela eta, web-a berri sozialez zein gai kulturez betetzen ari da. Baina, nahiz eta albiste hauek oso interesgarriak izan ikertzaileentzat, edozein momentutan, Interneten berehalakotasunarengatik, gehiago ez argitaratzea edota desagertzea ere gerta liteke.

Horrexegatik, Eusko Jaurlaritzaren Kultura Sailak jarri zuen martxan Ondarenet. Fitxategi honek, formatu digitalean sortutako baliabideak, aldi berean, be-

rreskuratzeko, zaintzeko eta hedatzeko aukera eskaintzen du. Artxibategi hau, 2008. urteko bukaera aldean hasi zen martxan.

Otsailaren hamahiruan, 2009ko euskal hauteskundeetarako kanpainaren hasiera izan zen. Kanpaina horren ezaugarriarik esanguratsuenetariko bat, Interneten izan duen presentzia izan da., Komunikabide guztietan aipatu da baliabide hau. Era berean, 2.0 webaren erabilera ere azpimarratu da, alderdi politikoek komunikazio tresna bihurtu baitute. Politikariek, baliabide berri hauen erabilpenarekin, kanpaina honetan eskaintzen zituzten proposamenenei ikuspen nabarmenagoa ematea lortu nahi izan dute.

Eusko Jaurlaritzaren Liburutegi Zerbitzua kanpaina honen originaltasunaz eta garrantziaz ohartu da eta ondorioz, «gertaera adierazgarriak» izeneko maila berri bat sortu du Ondarenet-en. Honela, hauteskunde kanpaina honekin harremana duten web guztiak epigrafe honen azpian bilduko dira.

2. ZER DA ONDARENET

Euskal ondare digitala berreskuratu, zaindu eta hedatzeko artxibategi elektronikoa da Ondarenet (www.euskadi.net/ondarenet). Honela, edozein artxibori dago-kion zaintza lanaz gain, eduki digitalen gordetzea eta sailkatzea baino harago doa. Izan ere, bere baitan, gorde nahi diren baliabide elektronikoak bilatu eta errekupe-ratzeko oinarritzko tresna barne hartzen du. Era berean, baliabideak hautatzeko eta biltzeko unean, badu kontsultatzeko interfaze erraza. Aukera horri esker, Euskadire-kin erlazionatutako gaietan edota euskararen hizkuntza-eremuan interesa duen edo-zeinek, sarbide erraza aurkituko du.

2.1. Arkitektura eta tresnak

Proiektua burutu ahal izateko beharrezkoak ziren tresnak hautatzean bi aukera aintzat hartu ziren: beharrezko tresnak neurrira garatuko zuen software komertzia-laz baliatzea edo International Internet Preservation Consortiumek (IIPC) proposa-tutako tresna-kita erabiltzea, enpresa espezializatu batek burututako garapen gehi-garria kontratatuz eta, betiere, EJIeko (Eusko Jaurlaritzako Informatika Elkarteko) teknikariren ikuskaritzapean. Aukera bakoitzaren alde on eta txarrak aztertu eta bi-garrena hautatu zen hainbat arrazoirengatik:

- Antzeko nazioarteko ekimen gehienek erabiltzen dituzten tresnak dira.
- Nahiko erraz instalatu eta mantentzen dira.
- Kode irekikoak direnez, «ad hoc» garapen batean askatasun osoz jardutea ahalbidetzen dute eta kostuak merkatzen dituzte.

Tool-kit honek osatzen dituen tresnak:

- **Heritrix**. Bildumako osagai digitalak, web guneak eta orriak biltzeko pro-zesua egiten duen atzipen-robot.



1. irudia

Ondarenet proiektuaren informazio-sistemako tresnak

- **NutchWAX.** Kode irekiko bilatzaile honen bitartez, Heritrix-ek bildutako elementuak bilatu eta indizatzen dira.
- **Web Curator.** Zeelanda Berriko Liburutegi Nazionalak, British Library-rekin elkarlanean diseinatutako tresna da. Honen bitartez, elementu digitalak atzitu eta biltzeko prozesuak kudeatzen dira. Erraz erabiltzeko moduko web interfazea bideratzen du, eta horren bidez, atzipenak planifikatu eta programatu ahal dira.
- **WERA.** Heritrix-ek atzitu eta WERAk indizatutako guneak bisitatzea ahalbidetzen duen aplikazioa da. Bilaketa errazak zein bereziak egin ahal dira honekin.

3. EUSKAL HAUTESKUNDEAK ONDARENENET-EN

Eusko Legebiltzarrerako hauteskundeak martxoaren lean egin ziren. Aurretik aipatu den bezala, Internetek, komunikabide, hedabide eta iritzibide gisa eskaintzen duen orokortzeari esker, gertaera hau ia minutuz minutu jarraitu ahal izan zen web orrian. Honekin, Eusko Jaurlaritzaren Liburutegi Zerbitzuak hauteskunde hauekin harremana zuten web orrialde adierazgarrienak berreskuratu, babestu eta hedatu nahi izan zituen. Hauek, «Eusko Legebiltzarrerako hauteskundeak (2009)»

izeneko «bilduma berezian» bildu ziren. Modu honetan, Ondarenet-i esker, aipatu orrialdeak babestu eta hedatu egingo dira, informazioa galtzeko aukera ekidituz. Azken finean, Ondarenet-en bidez, etorriko diren belaunaldiek ere informazio hau ikusteko aukera izango dute.

3.1. Behera-kargen plangintza eta hautaketa

Euskadik Inteneten berezko domeinurik ez duenez, edozein web edo elementu elektronikoren bilketa egiteko momentuan, berreskuratu nahi ditugun edukien alde aurretiko azterketa beharrezkoa da.

Ildo honetan, baliabide digital batzuk hautatu ziren (alderdien web-ak, komunikabideetako web-ak, eta abar). Hautatutako elementu guztiek harremana izan zuten gertaera honekin eta haien behera-kargak posible egingo dute etorkizunean hauteskunde hauen jarraipena egitea. Ondorioz, hauteskunde bilduma honek osatzen duen url-en artxiboa ikusteko, Ondarenet-en sartu beharra izango da.

3.1.1. Alderdi politikoak

Azterketa egin ondoren, euskal hauteskundeetan parte hartu zuten alderdi politikoen url-ak berreskuratzea erabaki zen. Alderdien webak zehazteko, «Hauteskundeak Euskadin» (http://www.euskadi.net/elecciones/indice_e.htm) web orriaz baliatu zen, Eusko Jaurlaritzako Herrizaingo Sailak mantentzen duena. Iturri honi esker hamahiru alderdi identifikatu ziren, eta horiek guztiek beren web orria zutela egiaztatu zen.

3.1.2. Blogak

Blogak zera dira: halako webgune pertsonalak, aldiro-aldiro egileen iritziak jasotzen dituztenak, eta irakurleei sarrera bakoitzean haien iruzkinak txertatzeko aukera ematen dietenak. Eusko Legebiltzarrerako hauteskundeetarako hautagai gehienek erabili dute web 2.0 teknologia berria, bai kanpainan zehar eta bai kanpaina ondoren ere, eta blogak, gutxi-asko aktiboak izan dituzte. Horrekin batera, hainbat herritar anonimok ere hauteskundeetan martxan oinarritutako blogak sortu dituzte.

Hauteskundeekin lotutako blog guztien lagin erakusgarri bat egiteko, hiru multzotan banatu ziren blogak: hautagaien blog propioak lehenik, Juan Jose Ibarretxe eta Patxi López hautagai nagusien aldeko blogak bigarrenik, eta, azkenik, hauteskundeetan prozesuan espezializatutako blogak.

3.1.3. Hedabideak

Bistan da gaur egun hedabideen bitartez baino ezin dela behar bezala jarraitu eguneroko gertaeren martxa eta garapena. Horregatik, hedabide nagusien URL-ak aukeratu ziren, bai Estatukoak bai autonomikoak, kapturak egiteko. Kontua da, zo-

ritzarrez, informazio-bolumen handia duten webguneetik behera-kargak egitea (esate baterako egunkari digitaletakoak) ez dela erraza, robotak dominiotik kanpo egin ditzakeelako kapturak, eta orduan prozesua baliogabetu egin behar da. Proben Zerbitzuan bolumen-karga ezberdinekin hainbat proba egin ondoren, ikusi zen barririk hiru egunkarirekin egin zitekeela arazorik gabe: *El Correo Digital*, *Deia* eta *El Diario Vasco*. EITBren webguneko «Euskal hauteskundeak 2009» bereziaren kaptura ere saiatu zen, baina ezinezkoa izan zen.

3.1.4. Fundazio politikoak

Badira halako fundazio politiko edo sozialak —gizarte-erakundeak eta bakearen aldeko taldeak, adibidez— herritarren iritziak jasotzen dituztenak, eta Interneten presentzia dutenak. Horietatik, garrantzitsuenetako bost aukeratu ziren.

Hementxe duzue kapturatutako webguneen koadroa, tipologiaren arabera ordenatuta. Hortxe azaltzen da kaptura egiteko denbora eta bakoitzaren tamaina:

Elementua	Behera-kargaren eguna	Tamaina	Iraupena (ee/hh/mm/ss)
<i>Alderdi politiko hautagaiak</i>			
Partido Popular del País Vasco (PP)	2009/02/04	467.36 MB	00:00:51:46
	2009/03/05	523.26 MB	00:01:06:11
Partido Socialista de Euskadi-Euskadiko Ezkerra (PSE-EE)	2009/02/05	3.33 GB	00:04:48:14
	2009/03/09	4.32 GB	00:04:48:46
Eusko Alkartasuna (EA)	2009/02/04	460.27 MB	00:02:35:26
	2009/03/09	530.46 MB	00:03:38:15
Euzko Alderdi Jeltzalea-Partido Nacionalista Vasco (EAJ-PNV)	2009/02/05	11.61 GB	01:04:21:11
	2009/03/05	12.7 GB	01:12:09:40
Aralar	20/01/2009	553 MB	00:21:24:00
Unión Progreso y Democracia (UPyD)	2009/02/11	469.19 MB	00:01:20:15
	2009/03/05	346.1 MB	00:01:02:04
Ezker Batua-Berdeak	ezin da		
Partido Por Un Mundo Más Justo (PUM+J)	2009/02/06	142.15 MB	00:01:12:43
	2009/02/05	144.53 MB	00:00:43:47
Partido Obrero Socialista Internacionalista	2009/02/06	31.32 MB	00:00:27:47
	2009/03/05	31.89 MB	00:01:10:21

Elementua	Behera-kargaren eguna	Tamaina	Iraupena (ee/hh/mm/ss)
Partido Carlista de Euskalherria-Euskalherriko Karlista Alderdia	2009/02/06	30.71 MB	00:00:02:45
Partido Familia y Vida	2009/02/12	33.07 MB	00:00:29:35
	2009/03/05	36.01 MB	00:00:11:29
Partido Antitaurino Contra el Maltrato Animal	2009/02/01	51.82 MB	00:00:27:42
	2009/03/05	54.54 MB	00:00:28:47
Partido Humanista	2009/02/12	6.06 MB	00:00:03:09
	2009/03/05	6.88 MB	00:00:04:11
<i>Hauteskundeei buruzko blogak</i>			
Elecciones Vascas 2009 (Blogspot)	2009/02/12	132.49 MB	00:01:59:35
	2009/03/10	239.98 MB	00:22:05:39
Hemen Ibarretxe	2009/03/10	1.3 MB	00:00:01:06
Konpondu	2009/02/25	271.12 MB	00:01:08:32
	2009/03/11	167.2 MB	00:08:24:44
Metablog Elecciones vascas	2009/02/12	132.49 MB	00:01:59:35
	2009/03/10	239.98 MB	00:22:05:39
Red para el cambio	2009/03/12	204.81 MB	00:06:40:19
<i>Gizarte-erakundeak eta bakearen aldeko taldeak</i>			
Elkarri. Movimiento social por el diálogo y el acuerdo	2009/02/13	127.87 MB	00:00:37:24
Parte hartuz	2009/02/13	526.29 MB	00:01:50:08
Coordinadora gesto por la paz de Euskal Herria	2009/02/13	27.22 MB	00:00:27:39
Gernika Gogoratuz	2009/02/12	47.29 MB	00:00:06:09
Batetik. Centro por la paz	2009/02/16	110.14 MB	00:01:50:41
<i>Hedabideak</i>			
<i>El Correo Digital</i> (Especial Elecciones Vascas 2009)	2009/02/16	121.97 MB	00:00:31:44
<i>El Diario Vasco</i> (Especial Elecciones Vascas 2009)	2009/02/16	19.74 MB	00:00:02:49
<i>Deia</i> (Especial Elecciones Vascas 2009)	2009/02/20	1.59 GB	00:05:56:17

Elementua	Behera-kargaren eguna	Tamaina	Iraupena (ee/hh/mm/ss)
<i>Hautagaiak</i>			
Antonio Basagoiti (PP)	2009/02/16	503.52 MB	00:00:47:15
Aintzane Ezenarro (Aralar)	ezin da		
Juan José Ibarretxe (EAJ-PNV)	2009/02/20	325.85 MB	00:00:14:20
Patxi López (PSE)	2009/02/19	8.39 GB	00:04:49:33
Javier Madrazo (EB-IU)	2009/02/26	3.19 GB	00:15:05:11
Gorka Maneiro (UPyD)	ezin da		
Unai Ziarreta Bilbao (EA)	ezin da		

4. ONDARENEN: ZERGATIK? ZERTARAKO?

Ondarenen-en helburuetariko bat da euskal gizarteko edota kulturako gertaera esanguratsuekin zerikusia duten webguneak eskuratzea eta informazio garrantzitsu hori gordetzea. Bestela, Interneten informazioak zein gutxi irauten duen kontuan izanda, galdu egingo litzateke edota ez ginateke ohartu ere egingo.

Ondarenen-en gordetzen den informazioaren zabalkundea eta atzipena behin eta berriro aipatu da idazlan honetan. Horregatik, alde batetik, baliabideen gaikako sailkapen taula bat sortu da. Bestalde, bilaketa errazteko interfazea diseinatu da, Internet-en erabiltzen diren bilatzaileen antzera. Honen bidez, edozeinek, berariazko prestakuntzarik gabe, Ondarenen-en sailkatuta dauden baliabideak berreskura ditzake.

Bilaketa-inprimakia

Bilaketa erraza

Testua:

Url-a:

Bilaketa aurreratua

Testua:

Formatua: -tik: -ra

Bilduma: Hurrenkera:

<ul style="list-style-type: none"> <input type="checkbox"/> Artea <input type="checkbox"/> Kultura <input type="checkbox"/> Hezkuntza eta ikerketa <input type="checkbox"/> Euskara <input type="checkbox"/> Politika eta gobernua <input type="checkbox"/> Gizartea <input checked="" type="checkbox"/> Gertaera nabarmenak 	<ul style="list-style-type: none"> <input type="checkbox"/> Zientzia eta teknologia <input type="checkbox"/> Ekonomia eta negozioak <input type="checkbox"/> Enpresa <input type="checkbox"/> Aisialdia eta kultura <input type="checkbox"/> Osasuna <input type="checkbox"/> Ezagutzaren gizartea
--	--

2. irudia

Ondarenen-eko bilaketa interfazea

Berreskurapenaren maiztasunari dagokionez, aukeratutako baliabide bakoitzeko bi behera-karga egitea erabaki da. Hala ere, hauteskunde bilduma honetan, bere iraupena kontutan hartuta, bi edo hiru behera-karga egin dira, betiere web-aren tainaren arabera.

Web bakoitza diseinatu den arkitekturari loturik dago berreskurapenaren bistaritzea. Gainera, webaren kapturak begiratzera edota orrialde bakoitzaren helbide zuzenera joatea posible da ere.

Emaitzak

Aurkitutako emaitza-kopurua: **1509**. Dokumentuen rankina **1-10**

1. Comunidad Blogs El Correo Digital » Tags » directores de orquesta
(... Correo Digital... Otros temas Portada Directorio Tags Login Contacto Reglamento Crea tu propio blog Estás en: Inicio » Tags » directores de **orquesta** Hoy , 26 Feb 2009 Divergencias Un concurso para directores de **orquesta** César Coca 1 comentario Lo habrán leído ustedes también estos días en los periódicos: la televisión británica está emitiendo un ... Comunidad Blogs El Correo Digital » Tags » directores de **orquesta** Actualidad Arte y Cultura Ciencia y tecnología Cine y Televisión Deportes Humor Internet Montaña Personales Tendencias Varios Vascos en el ...)

Bistaritzea | Atzipenen historikoa

2. Orquesta Sinfónica de Euskadi :: Euskadiko Orkestra Sinfonikoa :: Basque National Orchestra
(... del Principado de Asturias **Orquesta** Sinfónica de Castilla y León **Orquesta** y Coro Comunidad de Madrid **Orquesta** Sinfónica de Madrid **Orquesta** y Coro Nacional de España Real Filharmonía de Galicia Real **Orquesta** Sinfónica de Sevilla Orkestrak munduan **Orquesta** Filarmónica de Berlín **Orquesta** Concertgebouw de Amsterdam **Orquesta** Sinfónica de Londres **Orquesta** Filarmónica de Nueva York Royal Philharmonic Orchestra **Orquesta** Filarmónica de Londres **Orquesta** Filarmónica de Los Angeles **Orquesta** Filarmónica de Boston **Orquesta** Filarmónica de Israel **Orquesta** de París **Orquesta** de la)

Bistaritzea | Atzipenen historikoa

3. | Bilbao Orkestra Sinfonikoa |
(... Bilbao Orkestra Sinfonikoa | Idioma | Home | Mapa web | Contacto quienes somos La **Orquesta** El director Los músicos Historia Patronos / Patrocinadores Ubicación y sede Bilbao, Bizkaia, País Vasco conciertos Sinfónicos Cámara Cámara en Bizkaia ...)

3. irudia

Ondarenet-en egindako bilaketaren emaitza

Ondorio bezala, eta ondare digitala zaintzen eta babesten den neurrian, Ondarenet web-a berreskuratzeko eta zaintzeko ezinbesteko tresna dela esan daiteke. Dudarik gabe, ardura eta lan handia eskatzen duen asmo handiko proiektu baten aurrean gaude. Hala ere, bidea hasita dago eta 2009. urtean zehar, Ondarenet-eko artxiboan gero eta web gehiago aurkituko ditugu.

Horrela, nazioarteko erakunde entzutetsuek sortutako softwarea erabiliz gero, beste antzeko proiektu batzuek markatutako bidetik jo dezakegu (adibidez, Pandora, Internet Archive o Padicat proiektuak). Eta, noski, horrekin batera, *International Internet Preservation Consortiumek (IIPC)* erakundeak gomendatutako bertsió eta joera berriak erantsi.

BIBLIOGRAFIA

Cóccera, Daniel; Llueca, Ciro. «PADICAT: realitat i reptes de 3 anys d'arxiu web de Catalunya». Non: *Jornades Catalanes d'Informació i Documentació*, 2008, 163-178 orr.

- Llueca, Ciro. «Webs siempre accesibles: las bibliotecas nacionales y los depósitos digitales nacionales». Non: *BiD: textos universitaris de biblioteconomia i documentació* 2005 desembre, n. 15. Kontsulta: 2008-08-12. http://www2.ub.edu/bid/consulta_articulos.php?fichero=15lluec2.htm
- Paynter, Gordon; Joe, Susana; Lala, Vanita; Lee, Gillian. «A Year of Selective Web Archiving with the Web Curator at the National Library of New Zealand». Non: *D-Lib Magazine*, 2008 May/June, v. 14, n. 5/6. Kontsulta: 2008-08-12. <http://www.dlib.org/dlib/may08/paynter/05paynter.html>
- Plan Vasco de la Cultura*. Vitoria-Gasteiz : Eusko Jaurlaritzaren Argitalpen Zerbitzu Nagusia = Servicio Central de Publicaciones del Gobierno Vasco, 2004. ISBN 84-457-2166-6.
- Pulgar Vernalte, F. & Marcos Maciá, S. «Ondarenet: el archivo del patrimonio digital vasco». Non: *XI Jornadas de Gestión de la Información*, Madril, 2008ko azaroaren 20-21ean.
- Serra, Eugènia. «Archivando la Web catalana: iniciativas cooperativas de preservación digital en Catalunya». Non: *La Recuperación de la memoria, muchas más oportunidades que realidades: el trabajo cooperativo de archivos, bibliotecas y museos*. Universidad del País Vasco, 23-25 2006. Kontsulta: 2008-08-12. http://www.bnc.es/bc/archivando_web_catalana.pdf
- UNESCO. *Directrices para la preservación del patrimonio digital*, 2003. Kontsulta: 2008-08-09. <http://unesdoc.unesco.org/images/0013/001300/130071s.pdf>

IX

Las webs de las elecciones vascas en Ondarenet

Francisca Pulgar Vernalte y Sonia Marcos Maciá

Responsable del Servicio de Bibliotecas del Gobierno Vasco
y Técnico documentalista de la empresa ODEI
Liburutegi Zerbitzuko arduraduna Eusko Jaurlaritzan
eta Teknikari dokumentalista ODEI enpresan

Resumen: El objetivo de esta comunicación es presentar el proyecto de recuperación, preservación y difusión desde Ondarenet de las páginas webs de los partidos políticos que han participado en las elecciones al Parlamento Vasco en el 2009. Asimismo, se describe la selección y captura de los recursos electrónicos que conformarán la colección de «Elecciones al Parlamento Vasco 2009», y que podrán ser consultados por cualquier investigador o persona interesada desde Ondarenet, el archivo electrónico creado por el Departamento de Cultura del Gobierno Vasco para la recuperación, conservación y difusión de los recursos nacidos en formato digital.

Palabras clave: Repositorios digitales; Bibliotecas digitales; Archivos web, Preservación.

Laburpena: Euskal Hauteskundeetako webak Ondarenet-en, 2009an Eusko Legebiltzarrerako hauteskundeetan parte hartu zuten alderdi politikoen web-ak berreskuratu, zaindu eta hedatzeko proiektuaren aurkezpena da komunikazio honen helburua. Halaber, «2009ko Eusko Legebiltzarrerako Hauteskundeak» bilduma osatuko duten baliabide elektronikoek hautaketa eta berreskuraketa deskribatzen da. Baliabide hauek edozein ikertzailek edota interesaturik dagoen edonork Ondarenet-en kontsultatu ahal izango ditu. Ondarenet Euskadiko ondare digitalaren fitxategia da, eta fitxategi elektroniko hau Eusko Jaurlaritzaren Kultura Sailak jarri du martxan, formatu digitalean sortutako baliabideak berreskuratze, zaintzeko eta hedatzeko asmoz.

Hitz gakoak: Gordetegi digitalak; Liburutegi digitalak; Web fitxategiak, Zaintza.

1. INTRODUCCIÓN

Es evidente que cada vez se conocen y dominan mejor las posibilidades que ofrece Internet tanto para la publicación de contenidos como para la difusión de los mismos. Cada día es más frecuente la publicación de contenidos directamente en Internet, es decir, sin editar en papel, así que la web se llena de noticias sociales y culturales de interés para los investigadores pero que, a veces, condicionadas por la inmediatez de Internet, pueden dejar de publicarse y desaparecer.

Por ello, el Departamento de Cultura del Gobierno Vasco puso en marcha Ondarenet, un archivo que permite al mismo tiempo la captura, preservación y difusión de los recursos electrónicos que se crean en formato digital, y que inició su andadura a finales del 2008.

El día trece de febrero dio comienzo la campaña para las elecciones autonómicas vascas 2009. Una de las características principales de esta campaña, muy comentada en todos los medios de comunicación, fue la presencia en Internet de los contenidos de dicha campaña, y especificando aún más, el uso de la web 2.0 como herramienta de comunicación por parte de los partidos políticos. Igualmente, los políticos no podían dejar de lado esta nueva tendencia social de editar contenidos y de estar presentes en la red, y han hecho de Internet un instrumento para dar mayor visibilidad a sus propuestas en esta campaña de las elecciones vascas.

La originalidad e importancia de este evento han llevado al Servicio de Bibliotecas del Gobierno Vasco a crear en Ondarenet una nueva categoría denominada «hechos relevantes», y bajo este epígrafe se recopilarán todas las webs relacionadas con esta campaña electoral.

2. ¿QUÉ ES ONDARENET?

Ondarenet (www.euskadi.net/ondarenet) se define como un archivo electrónico que permite la captura, conservación y difusión del patrimonio digital vasco. Quiere esto decir, que además de la labor de conservación reconocida a cualquier tipo de archivo, Ondarenet va más allá del almacenamiento y ordenación de los contenidos digitales, incluyendo una herramienta esencial para la búsqueda y recuperación de los recursos electrónicos que se pretenden conservar. Es decir, los selecciona y recoge ofreciendo una sencilla interfaz de consulta que facilita su difusión, y por lo tanto, el acceso a su información a cualquier persona interesada en la investigación de temas relacionados con Euskadi o con el ámbito lingüístico del euskera.

2.1. Arquitectura y herramientas

Uno de los aspectos más interesantes de Ondarenet son las herramientas informáticas necesarias que permiten conseguir los objetivos descritos. Para llevar a cabo el proyecto se valoraron dos posibilidades: contratar un software comercial que desarrollara a medida las herramientas necesarias o utilizar el Toolkit propuesto por el International Internet Preservation Consortium (IIPC), con un desarrollo adicional llevado a cabo por una empresa especializada y supervisado en todo momento por los técnicos de EJIE (Sociedad Informática del Gobierno Vasco). Se analizaron los pros y los contras de cada elección y se optó por la segunda de las opciones por varias razones:

- Son las herramientas utilizadas por la mayoría de las iniciativas internacionales similares existentes.
- Son relativamente fáciles de instalar y mantener.
- Son de código abierto lo que permite una total libertad en un desarrollo «ad hoc» y abarata los costes.

Este Toolkit está compuesto por:

- **Heritrix**. El robot de captura que realiza el proceso de recolección de los componentes digitales, sitios y páginas web de la colección.
- **NutchWAX**. El motor de búsqueda de código abierto que permite la búsqueda e indización de los elementos de la colección recolectados por Heritrix.
- **Web Curator**. Es la herramienta, diseñada por la Biblioteca Nacional de Nueva Zelanda en colaboración con la British Library, que gestiona los procesos de captura y recolección de los elementos digitales (urls) que van a componer la colección. Proporciona un interfaz web de fácil utilización a través del cual planificar y programar las capturas.
- **WERA**. Es la aplicación que hace posible al usuario final la consulta de los sitios capturados por Heritrix e indizados por WERA. Permite realizar búsquedas tanto simples como avanzadas.



Figura 1

Herramientas del sistema de información del proyecto Ondarenet

3. LAS ELECCIONES VASCAS EN ONDARENET

El uno de marzo de 2009 se celebraron las elecciones al Parlamento Vasco. Tal y como se ha apuntado, gracias a la generalización de Internet como medio de información, difusión y opinión, este acontecimiento pudo seguirse casi minuto a minuto a través de la web. La intención del Servicio de Bibliotecas del Gobierno Vasco fue recoger en Ondarenet las webs más representativas relacionadas con la celebración de estos comicios, agrupadas todas ellas bajo una «colección especial» denominada «Elecciones al Parlamento Vasco (2009)». De este modo, y gracias a Ondarenet, estas webs se conservarán y difundirán posibilitando a las generaciones futuras el acceso a una información que, de otro modo, podría acabar desapareciendo.

3.1. Planificación y selección de las capturas

Al no existir para Euskadi un dominio propio en Internet, cualquier recopilación de webs o recursos electrónicos obliga a una investigación previa de aquellos contenidos que se desean capturar.

En esta línea, se seleccionaron una serie de recursos digitales (webs de partidos políticos, medios de comunicación, etc.) directamente relacionados con este evento y cuya captura permitirá realizar en cualquier momento un seguimiento de la campaña electoral y de sus resultados en la web vasca, es decir, una búsqueda de urls potencialmente interesantes para esta colección electoral.

3.1.1. *Partidos políticos*

Tras un análisis previo se decidió realizar la captura de urls de partidos políticos que presentaban su candidatura a las elecciones vascas. Con el fin de conocer cuáles eran, se recurrió a la página web «Elecciones en Euskadi» (<http://www.euskadi.net/elecciones>) mantenida por el Departamento de Interior del Gobierno Vasco. A través de esta fuente se identificaron trece partidos, todos los cuales poseían su correspondiente página web.

3.1.2. *Blogs*

Los blogs son páginas web dónde los autores expresan sus opiniones de una manera periódica, ofreciendo la posibilidad a sus lectores de realizar comentarios a cada una de las entradas. Los principales candidatos a las elecciones al Parlamento Vasco han aprovechado las tecnologías web 2.0., y la mayoría de ellos han mantenido blogs más o menos activos durante y después de la campaña electoral. Asimismo han aparecido blogs creados por ciudadanos anónimos dedicados exclusivamente al desarrollo de los comicios.

Con el fin de capturar una muestra representativa de blogs relacionados con las elecciones, se seleccionaron blogs de candidatos, blogs de «apoyo» a los dos principales candidatos, Juan José Ibarretxe y Patxi López, y por último dos blogs «especializados» en el desarrollo de las elecciones.

3.1.3. Medios de comunicación

Es evidente que la mejor manera de seguir los acontecimientos relacionados con cualquier evento es a través de los diferentes medios de comunicación. Por ello, se seleccionaron las urls de los principales medios de comunicación escritos tanto autonómicos como estatales, con el fin de proceder a su captura. Lamentablemente la descarga de webs con grandes volúmenes de información, como es el caso de las webs de los periódicos digitales, presenta algunas dificultades, ya que se corre el peligro de que el robot realice capturas fuera de dominio y haya que abortar la descarga. Una vez realizadas distintas pruebas de captura en el servidor de pruebas con diferentes perfiles de profundidad, se comprobó que tan sólo eran factibles las capturas de tres medios de comunicación: *El Correo Digital*, *Deia* y el *Diario Vasco*. Se intentó asimismo la captura del especial «Elecciones vascas 2009» mantenido en la web de EITB, pero su descarga fue imposible.

3.1.4. Fundaciones políticas

Existen algunas fundaciones de carácter político y/o social, tales como asociaciones cívicas y movimientos por la paz, que recogen la opinión de la ciudadanía y que también tienen presencia en Internet. Se identificaron y seleccionaron cinco de las más representativas.

Se presenta a continuación un cuadro de los sitios web capturados, agrupados por tipología, en él se explican los tiempos de captura y el tamaño de cada recurso:

Recurso	Fecha	Volumen	Tiempo descarga
<i>Partidos políticos con candidatura</i>			
Partido Popular del País Vasco (PP)	04/02/2009	467.36 MB	00:00:51:46
	05/03/2009	523.26 MB	00:01:06:11
Partido Socialista de Euskadi-Euskadiko Ezkerra (PSE-EE)	05/02/2009	3.33 GB	00:04:48:14
	09/03/2009	4.32 GB	00:04:48:46
Eusko Alkartasuna (EA)	04/02/2009	460.27 MB	00:02:35:26
	09/03/2009	530.46 MB	00:03:38:15

Recurso	Fecha	Volumen	Tiempo descarga
Euzko Alderdi Jeltzalea-Partido Nacionalista Vasco (EAJ-PNV)	05/02/2009	11.61 GB	01:04:21:11
	05/03/2009	12.7 GB	01:12:09:40
Aralar	20/01/2009	553 MB	00:21:24:00
Unión Progreso y Democracia (UPyD)	11/02/2009	469.19 MB	00:01:20:15
	05/03/2009	346.1 MB	00:01:02:04
Ezker Batua-Berdeak	imposible la descarga		
Partido Por Un Mundo Más Justo (PUM+J)	06/02/2009	142.15 MB	00:01:12:43
	05/02/2009	144.53 MB	00:00:43:47
Partido Obrero Socialista Internacionalista	06/02/2009	31.32 MB	00:00:27:47
	05/03/2009	31.89 MB	00:01:10:21
Partido Carlista de Euskalherria-Euskalherriko Karlista Alderdia	06/02/2009	30.71 MB	00:00:02:45
Partido Familia y Vida	12/02/2009	33.07 MB	00:00:29:35
	05/03/2009	36.01 MB	00:00:11:29
Partido Antitaurino Contra el Maltrato Animal	01/02/2009	51.82 MB	00:00:27:42
	05/03/2009	54.54 MB	00:00:28:47
Partido Humanista	12/02/2009	6.06 MB	00:00:03:09
	05/03/2009	6.88 MB	00:00:04:11
<i>Blogs sobre las elecciones</i>			
Elecciones Vascas 2009 (Blogspot)	12/02/2009	132.49 MB	00:01:59:35
	10/03/2009	239.98 MB	00:22:05:39
Hemen Ibarretxe	10/03/2009	1.3 MB	00:00:01:06
Konpondu	25/02/2009	271.12 MB	00:01:08:32
	11/03/2009	167.2 MB	00:08:24:44
Metablog Elecciones vascas	12/02/2009	132.49 MB	00:01:59:35
	10/03/2009	239.98 MB	00:22:05:39
Red para el cambio	12/03/2009	204.81 MB	00:06:40:19

Recurso	Fecha	Volumen	Tiempo descarga
<i>Asociaciones cívicas y movimientos por la paz</i>			
Elkarri. Movimiento social por el diálogo y el acuerdo	13/02/2009	127.87 MB	00:00:37:24
Parte hartuz	13/02/2009	526.29 MB	00:01:50:08
Coordinadora gesto por la paz de Euskal Herria	13/02/2009	27.22 MB	00:00:27:39
Gernika Gogoratzuz	12/02/2009	47.29 MB	00:00:06:09
Batetik. Centro por la paz	16/02/2009	110.14 MB	00:01:50:41
<i>Medios de comunicación</i>			
<i>El Correo Digital</i> (Especial Elecciones Vascas 2009)	16/02/2009	121.97 MB	00:00:31:44
<i>El Diario Vasco</i> (Especial Elecciones Vascas 2009)	16/02/2009	19.74 MB	00:00:02:49
<i>Deia</i> (Especial Elecciones Vascas 2009)	20/02/2009	1.59 GB	00:05:56:17
<i>Candidatos</i>			
Antonio Basagoiti (PP)	16/02/2009	503.52 MB	00:00:47:15
Aintzane Ezenarro (Aralar)	imposible la descarga		
Juan José Ibarretxe (EAJ-PNV)	20/02/2009	325.85 MB	00:00:14:20
Patxi López (PSE)	19/02/2009	8.39 GB	00:04:49:33
Javier Madrazo (EB-IU)	26/02/2009	3.19 GB	00:15:05:11
Gorka Maneiro (UPyD)	imposible la descarga		
Unai Ziarreta Bilbao (EA)	imposible la descarga		

4. POR QUÉ Y PARA QUÉ ONDARENET

Uno de los objetivos de Ondarenet es realizar capturas selectivas de webs que tengan que ver con algún acontecimiento clave relacionado con la sociedad o la cultura vasca, con el fin de conservar informaciones importantes que de otro modo, y debido a lo efímero o la saturación de la información en Internet, podrían acabar desapareciendo o pasar desapercibidas.

A lo largo de la comunicación se ha insistido en la difusión y acceso de la información que se almacena en Ondarenet. De ahí que, por un lado se haya creado

Patrimonio Digital - Buscador

Búsqueda simple

Texto:

Url:

Búsqueda avanzada

Texto:

Formato: **Todos los formatos** Desde: Hasta:

Colección: **Todas las colecciones** Orden: **Ranking**

- Arte
- Cultura Vasca
- Educación e Investigación
- Euskera
- Política y gobierno
- Sociedad
- Ciencia y Tecnología
- Economía y Negocios
- Empresa
- Ocio y Cultura
- Salud
- Sociedad de al información

Figura 2

Interfaz de búsqueda de Ondarenet

una tabla con la clasificación temática de los recursos y, por otro lado, se haya diseñado una interfaz de búsqueda sencilla, al estilo de los buscadores de Internet, que permita a cualquier persona, sin una preparación específica, la recuperación de los recursos clasificados en Ondarenet.

En cuanto a la periodicidad de las capturas, se ha establecido hacer dos capturas anuales de cada recurso seleccionado, aunque en esta colección electoral, por su propia duración, se han hecho dos o tres descargas siempre en relación con el tamaño de cada web.

La visualización de las capturas está siempre en función de la arquitectura con la que ha sido diseñada cada web, y cabe mencionar que se puede acceder tanto a las capturas periódicas llevadas a cabo sobre cada recurso como a la web actual de cada uno de ellos.

Patrimonio Digital - Resultados

Número de versiones encontradas : **1120**. Ranking de versiones **1-7**

1. http://www.euskadikoorkestra.es/giras/0506_02/orquesta.gif
(... http://www.euskadikoorkestra.es/giras/0506_02/orquesta)

Visualización | Historico de Versiones | Mas de este sitio

2. Orquesta popular vasca
(Orquesta popular vasca euskera | castellano | english recherche avancée | index web Publications Jentilbaratz. Cuadernos de Folklore Cuadernos de Sección. Folklore Orquesta popular vasca Hernández Arsuaga, Javier FICHE Publication: Donostia-San Sebastián : Eusko Ikaskuntza, 1983 Número: 1 Toponymie: Euskal Herria Matières: Música ... Español Pages: 241-246 ISSN: 1137-859X ISBN: 84-7086-093-3 prix 4, 50€. Résumé La base de la orquesta popular vasca es el txistu en sus diferentes tonos: txistu-txiki, txilibitu, txistu, silbote, txistu-bajo, silbote-bajo. La orquesta ...)

Visualización | Historico de Versiones | Mas de este sitio

Figura 4

Resultado de búsqueda en Ondarenet

A modo de conclusión podría decirse que Ondarenet es una herramienta imprescindible para capturar y preservar la web, del mismo modo que se cuida y preserva el patrimonio bibliográfico. Se trata, sin duda alguna, de un proyecto ambicioso, que requiere de un gran esfuerzo y dedicación, pero, el camino ya está marcado, y a lo largo de este año 2009 se irá incrementando el número de páginas webs archivadas en Ondarenet.

Utilizar el software creado por instituciones de prestigio en el ámbito internacional le permite seguir en la línea marcada por proyectos similares como Pandora, Internet Archive o Padicat, entre otros, y lógicamente evolucionar junto a ellos, incorporando las nuevas versiones y tendencias que se indiquen desde el International Internet Preservation Consortium (IIPC).

BIBLIOGRAFÍA

- Cócera, Daniel; Lluca, Ciro. «PADICAT: realitat i reptes de 3 anys d'arxiu web de Catalunya». En: *Jornades Catalanes d'Informació i Documentació*, 2008, pp. 163-178.
- Lluca, Ciro. «Webs siempre accesibles: las bibliotecas nacionales y los depósitos digitales nacionales». En: *BiD: textos universitaris de biblioteconomia i documentació* 2005 desembre, n. 15. Consultado en: 12-08-2008. http://www2.ub.edu/bid/consulta_articulos.php?fichero=15lluca2.htm
- Paynter, Gordon; Joe, Susana; Lala, Vanita; Lee, Gillian. «A Year of Selective Web Archiving with the Web Curator at the National Library of New Zealand». En: *D-Lib Magazine*, 2008 May/June, v. 14, n. 5/6. Consultado en: 12-08-2008. <http://www.dlib.org/dlib/may08/paynter/05paynter.html>
- Plan Vasco de la Cultura*. Vitoria-Gasteiz : Eusko Jaurlaritzaren Argitalpen Zerbitzu Nagusia = Servicio Central de Publicaciones del Gobierno Vasco, 2004. ISBN 84-457-2166-6.
- Pulgar Vernalte, F. & Marcos Maciá, S. «Ondarenet: el archivo del patrimonio digital vasco». En: *XI Jornadas de Gestión de la Información*, Madrid, 20-21 Noviembre 2008.
- Serra, Eugènia. «Archivando la Web catalana: iniciativas cooperativas de preservación digital en Catalunya». En: *La Recuperación de la memoria, muchas más oportunidades que realidades: el trabajo cooperativo de archivos, bibliotecas y museos*. Universidad del País Vasco, 23-25 2006. Consultado en: 12-08-2008. http://www.bnc.es/bc/archivando_web_catalana.pdf
- UNESCO. *Directrices para la preservación del patrimonio digital*, 2003. Consultado en: 09-08-2008. <http://unesdoc.unesco.org/images/0013/001300/130071s.pdf>

X

La experiencia catalana archivando la red: el repositorio Padicat (Patrimonio Digital de Cataluña) de la Biblioteca de Catalunya

Ciro Lluca y Daniel Cócera

Coordinador del Padicat en la Biblioteca de Catalunya
y Responsable de gestión de colecciones del Padicat
Padicat-en Koordinatzailea Kataluniako Liburutegi Nagusian
eta Padicat-eko Bildumen kudeaketaren arduraduna

Resumen: Desde 1996 administraciones de diversos países han iniciado proyectos destinados a preservar los contenidos publicados en Internet con el objetivo de preservar el acceso a esa información nacida digital. En España, la Biblioteca de Catalunya creó en 2005 el repositorio Padicat (Patrimonio Digital de Cataluña). La estrategia de preservación digital se concreta en las páginas web publicadas en lengua catalana u otras lenguas, bajo el dominio .CAT y otros dominios geográficos o temáticos, que estén relacionadas temáticamente con Cataluña. Tras cuatro años de Padicat, la BC ofrece en abierto (www.padicat.cat) más de 4.000 recursos digitales, capturados en base a la captura sistemática y por los acuerdos con 400 instituciones; así como por la recopilación sistemática de diversos acontecimientos clave para comprender la sociedad catalana, como las últimas campañas electorales en Internet. Los beneficios de un repositorio como Padicat llegan a todos los sectores de la sociedad: para la ciudadanía representa el acceso abierto y permanente a los recursos que son fruto del conocimiento y expresión de los creadores del siglo XXI. Para las instituciones, empresas, administraciones y particulares que producen páginas web supone la preservación de la propia producción y garantía de acceso con los condicionantes que en cada caso la ley regula a los contenidos y diseños que, de otro modo, desaparecerían.

Laburpena: Kataluniako esperientzia sarea artxibatzen: Kataluniako Liburutegiko Padicat errepositorioa (Kataluniako Ondare Digitala). 1996tik aurrera Interneten argitaratutako datuak gordetzeko proiektuak abiarazi dituzte zenbait estatutako administrazioek, modu digitalean sortutako informazio horren sarbidea babesteko helburuarekin. Kataluniako Bibliotekak, esaterako, 2005ean Padicat biltegia, (Kataluniako Ondare Digitala) sortu zuen. Preserbazio digitaleko estrategia katalanez edo beste hizkuntzetan argitaratutako web orrietan zehazten da, .CAT domeinuan edota beste domeinu geografiko edo tematikoetan, baina betiere tematikoki Kataluniarekin lotutakoak. Lau urteren ondoren, Padicat-en (www.padicat.cat) 4.000 baliabide digital baino gehiago eskaintzen ditu modu irekian. Kataluniako Bibliotekak atzitze sistematikoen eta 400 instituzioekin dauzkan hitzarmenen bitartez lortuak, baita Kataluniako gizartea ulertzeko giltzari diren gertaeren atzitze sistematikoen bidez ere; esate baterako, azken hauteskunde kanpainak Interneten. Padicat mo-

duko biltegi batek dituen onurak gizarteko sektore guztietara iristen dira. Alde batetik, herritarrek sarrera irekia eta iraunkorra dute jakinduriaren fruitu eta XXI. mendeko sortzaileen adierazpide diren baliabideetara. Bestetik, web orriak ekoizten dituzten instituzio, enpresa, administrazio eta partikularrei, betiere legeak kasu bakoitzean ezartzen dituen baldintzetan, euren produkzioa gordetzen eta diseinu eta edukiak eskuratzea bermatzen zaie; izan ere, bestela hori guztia galdu egingo zen.

1. INTRODUCCIÓN: ¿QUÉ ES EL PATRIMONIO DIGITAL?

Las tecnologías de la información y la comunicación han permitido que el patrimonio cultural y científico se presente en formato digital. Tal como se expresa la Unesco en sus *Directrices para la preservación del patrimonio digital*¹, los recursos que son fruto del conocimiento o la expresión de los seres humanos, ya sean de carácter cultural, educativo, científico o administrativo, o comprendan información técnica, jurídica, médica o de otro tipo, se generan cada vez más a menudo directamente en formato digital, o se convierten a este formato a partir de material analógico ya existente. Pero ya antes de la publicación de estas Directrices, desde aproximadamente el inicio de la década de los 90 (1992-1995), el auge en progresión geométrica de los servidores y páginas web de Internet, motivaron la necesidad de articulación de proyectos para la creación de repositorios con el objetivo de preservar totalmente o en parte la producción documental sustentada en las páginas web, publicada en Internet. Así, desde 1996 las administraciones de diversos países han llevado a cabo estrategias para garantizar el acceso permanente a la producción digital propia. Estas acciones están orientadas a asegurar en la medida de las posibilidades tecnológicas actuales la adecuación del ciclo documental clásico a las páginas web: la compilación, el procesamiento, la preservación y el acceso permanente a la producción bibliográfica digital. El reto no es menor y las amenazas son múltiples: por una parte, la manifiesta obsolescencia del texto legal que posibilita el depósito legal en España. Complementariamente, el crecimiento exponencial de la producción digital, sumado a la baja permanencia de los materiales publicados en Internet y, por supuesto y finalmente, el respeto a la legislación en materia de propiedad intelectual. La propia naturaleza dinámica de la red es el factor de máxima erosión de los documentos que alberga. La información muta a diario: es sustituida o simplemente desaparece. El diseño, el entorno físico, evoluciona sin dejar rastro del anterior. Los nombres de los dominios varían, haciendo que, de un día para otro, no localicemos un recurso que ayer estábamos consultando. Como arena fina ente las manos desaparece a diario una ingente producción literaria, científica, educativa, lúdica o de creación artística, una producción que debemos considerar patrimonial, más aún cuando, como se expone en las referidas directrices, la vía digital es y será una de las principales herramientas, bien de creación, bien de difusión, de los creadores del siglo XXI.

¹ *Directrices para la preservación del patrimonio digital*. Canberra: Unesco, 2003. [Consulta: 30/09/2009] <http://unesdoc.unesco.org/images/0013/001300/130071s.pdf>

Pese a estas dificultades, diversos países están realizando acciones de preservación de la producción digital más obvia: las páginas web. Las bibliotecas nacionales han sido a menudo impulsoras de estas acciones y en el caso español, la Biblioteca de Catalunya (BC) puso en marcha en 2005 el repositorio Padicat (Patrimonio Digital de Cataluña)², dedicado al archivo sistemático de la producción digital en Cataluña.

2. EL MUNDO ARCHIVA LA WEB

El archivo de la Web, como popularmente se conoce al conjunto de técnicas dirigidas para la creación de repositorios digitales como Padicat, no es hoy en día una acción consolidada en la mayoría de bibliotecas nacionales o entre otros órganos competentes en preservación patrimonial.

Un repositorio (o depósito) digital nacional es la herramienta fruto de la iniciativa dedicada a compilar, procesar y dar acceso a los recursos digitales de todo tipo creados en un territorio determinado, o sobre este territorio. Estos «archivos web» son complementarios a los depósitos institucionales (como el *E-Prints* (<http://eprints.ucm.es/>) de la Universidad Complutense de Madrid o el *DSPACE Revistes* y el *DSPACE Eprints* (<https://upcommons.upc.edu/>) de la Universitat Politècnica de Catalunya, entre muchas otras iniciativas); o temáticos (como la *Biblioteca Virtual Miguel de Cervantes* (<http://www.cervantesvirtual.com>), el portal *Tecnociencia e-revistas* (<http://www.tecnociencia.es/e-revistas>), o el portal *Temaria* (<http://temaria.net>) que coordina la Facultat de Biblioteconomia i Documentació de la Universitat de Barcelona.

Existen diversos repositorios nacionales en funcionamiento, así como extensa bibliografía que los ha detallado y analizado³. Los más conocidos son también los que dieron los primeros pasos en 1996: el sueco *Kulturarw3* (<http://www.kb.se/english/find/internet/websites/>) y el australiano *Pandora* (<http://pandora.nla.gov.au/index.html>); así como un repositorio de alcance internacional, el gigante *Internet Archive* (<http://www.archive.org>). Trece años más tarde podemos contar hasta 36 proyectos en diversas fases de implementación, siendo acciones consolidadas un tercio de esta cifra.

El análisis de estas experiencias muestra dos modelos básicos de sistemas con una tendencia generalizada hacia un modelo híbrido. El primero es el modelo integral o exhaustivo (mayoritario, y característico especialmente de los países escandinavos), que persigue la integración automática de la Web a partir de determinados criterios infraestructurales (lingüísticos, según el dominio de las páginas web, según la ubicación del servidor, etc.). El segundo modelo es el selectivo (asimilado por Australia, el Reino Unido o Japón, entre otros países), dirigido a compilar la Web en base a una política selectiva (sobre un espacio geográfico determinado, un tema de interés nacional, etc.). Estos dos modelos han dado paso, en lo que es ya

² El portal PADICAT, <http://www.padicat.cat> está operativo desde septiembre de 2006.

³ Para un panorámica global sobre estos proyectos véase: Llucca, Ciro. «Webs siempre accesibles: las bibliotecas nacionales y los depósitos digitales nacionales». En: *BiD: textos universitaris de biblioteconomia i documentació*, 2005, diciembre, n. 15. [Consulta: 30/09/2009] <http://www.ub.es/bid/15lluca2.htm>

una tendencia generalizada, a modelos híbridos (cuya caso más evidente es el de Dinamarca) que complementan la captura periódica de la Web con acciones selectivas, y en ciertos casos ampliando además esa cobertura a determinados acontecimientos de interés social (elecciones, competiciones deportivas, etc.).

Lamentablemente, el número de depósitos que, como Padicat u *Ondarenet* (<http://www.ondarenet.kultura.ejgv.euskadi.net/>), permiten acceder libremente a sus colecciones o a su fondo, es muy limitado. A menudo se trata de evitar potenciales conflictos con la vulneración de los derechos de propiedad intelectual de los recursos capturados sin autorización expresa; aunque complementariamente a esta legítima preocupación, un factor determinante radica en el hecho que no se hayan perfeccionado las interfaces de visualización de la información depositada.

3. LA BIBLIOTECA DE CATALUNYA Y EL REPOSITORIO PADICAT

A las puertas de su centenario en 2007, la BC apostó con firmeza por una serie de acciones para evolucionar hacia un modelo de biblioteca abierta, fiable y orientada al usuario. Una de las líneas estratégicas ha sido el impulso de proyectos digitales⁴, de carácter eminentemente cooperativo, para contribuir a la preservación del patrimonio catalán y aumentar la presencia de contenidos catalanes en Internet. Algunos de esos proyectos son *ARCA* (Archivo de Revistas Catalanas Antiguas), *RACO* (Revistas Catalanas de Acceso Abierto), *CLACA* (Clásicos Catalanes), el proyecto *Google Books* de digitalización de parte del fondo de la Biblioteca libre de copyright, o el propio Padicat.

De hecho, tal como se expresa en las leyes catalanas de bibliotecas de 1981 y 1993⁵, la BC tiene por misión recopilar, conservar y difundir la producción bibliográfica catalana y la relacionada con el ámbito lingüístico catalán, y velar por la conservación y la difusión del patrimonio bibliográfico. Entendemos que este patrimonio bibliográfico incluye también la producción bibliográfica digital, la que se publica en Internet.

Así, tomando la misión ya descrita e incluyendo en ella la producción digital, establecemos el objetivo genérico de Padicat: diseñar y producir un sistema que permita a la BC compilar, procesar y dar acceso permanente a la producción digital catalana, confeccionando de esta manera la bibliografía digital de Cataluña.

⁴ La estrategia y los proyectos, a excepción del Google Books que fue un acuerdo posterior, fueron presentados en: Lamarca, D.; Serra, E. «L'estratègia de la Biblioteca de Catalunya en projectes digitals». En: *Ítem*, 2005, setembre-desembre, n. 41, pp. 41-43.

⁵ Artículo 7.1 de la «Llei de biblioteques de Catalunya, de 24 d'abril de 1981». En: *Diari Oficial de la Generalitat de Catalunya*, 1981, 29 abril, n. 123. Dicho artículo se refrendó en la «Llei 4/1993 del sistema bibliotecari de Catalunya, de 18 de març de 1993». En: *Diari Oficial de la Generalitat de Catalunya*, 1993, 29 març, n. 1727.

Padicat es un depósito digital pionero en España y cuenta con la colaboración del Centre de Supercomputació de Catalunya (Cesca) y de la Generalitat de Catalunya, mediante su *Secretaria de Telecomunicacions i Societat de la Informació*. Tiene un presupuesto aproximado cercano al millón de euros. Está basado en el modelo híbrido (modelo danés), que se vertebra alrededor de tres ejes: la compilación masiva de recursos digitales; la selección de recursos digitales representativos de la Internet catalana; la creación de centros temáticos de interés entorno a acontecimientos de la vida pública catalana.

El día 11 de septiembre de 2006 se inauguró la web de Padicat (www.padicat.cat) en una versión trilingüe que hoy, corregida y aumentada, se mantiene. Desde el primer día, como filosofía del repositorio, se ha dado acceso abierto vía Internet a toda la colección disponible. Primero, con un motor de búsqueda a texto completo. En una segunda fase, con la creación de centros de interés: paquetes temáticos de interés público como puedan ser las campañas electorales en Internet, que se analizan más profundamente en líneas venideras. Finalmente, se han completado las opciones anteriores con la opción de búsqueda a través de URL y, sobre todo, de un directorio temático, dedicado al público que prefiere la navegación como fórmula de visita del fondo que forma la colección de Padicat. Más allá del éxito de estos trabajos, a día de hoy debemos afirmar que los procesos de posicionamiento de los resultados en respuesta a las búsquedas por texto completo, y muy especialmente la presentación de los recursos digitales capturados, están aún lejos de poderse considerar como óptimos.

Pese a todo, el depósito Padicat ha ayudado a posicionar a la Biblioteca de Catalunya en una situación de liderazgo en lo referente a preservación digital de páginas web. A nivel internacional, el repositorio forma parte desde febrero de 2007 de la principal red de trabajo en preservación digital, el International Internet Preservation Consortium (IIPC), y ha sido distinguido en enero de 2008 por la Library of Congress, responsable de comunicación de este consorcio, como ejemplo de archivo web por sus acciones en la campaña electoral de las elecciones municipales 2007.

Por otro lado, la BC ha asistido en estos tres años a un centenar de actos profesionales para explicar la iniciativa, proyectando una imagen de liderazgo en preservación del patrimonio digital, y ha tenido un impacto permanente en medios de comunicación especializados, y también generalistas, gracias a la emisión periódica de comunicados de prensa y otras fórmulas informativas.

Esta vocación de liderazgo es la que ha motivado la estrategia de la BC, basada en el aprendizaje constante tomando como espejo a los líderes mundiales en preservación digital, las entidades internacionalmente pioneras, como el referido *Internet Archive*, las bibliotecas nacionales escandinavas, los grupos de trabajo de estos organismos, etc. Hoy sabemos que tanto la distancia física como la lengua de contacto, en todo caso, no permiten el aprovechamiento de las sinergias (proyectos idénticos con objetivos similares alrededor del mundo) en la medida de lo deseable. Las listas de distribución y las reuniones esporádicas no suplen cualitativamente, aún, las posibilidades de aprendizaje mutuo. La inexistencia de

proyectos similares en España, hasta el impetuoso arranque de Ondarenet, en el País Vasco, no ha posibilitado compartir experiencias en un entorno que trabaja cooperativamente en otras materias comunes. Por otro lado, somos conscientes que los proyectos internacionales de depósito digital nacional que están consolidados no dedican a estas tareas los recursos necesarios para la mejora permanente de sus herramientas, mejora de la que se podrían beneficiar proyectos como Padicat.

Como se apuntaba en la definición esquemática del repositorio, el modelo que ha adoptado la BC es el sistema híbrido, en la línea de la referida tendencia generalizada en las bibliotecas nacionales, consistente en compilar masivamente los recursos digitales publicados en abierto en Internet; impulsar los acuerdos selectivos con los agentes implicados en la producción digital en Cataluña; y promover líneas de investigación específicas por medio de la integración focalizada de recursos digitales sobre determinados acontecimientos de la vida pública catalana.

Ello se traduce, actualmente y después de cuatro años de experiencia, en cerca de 450 acuerdos de colaboración con entidades de todo tipo que conforman el entramado cívico, cultural, empresarial, deportivo, etc., de la sociedad catalana: la mayoría de universidades catalanas, colegios y asociaciones profesionales, clubs y federaciones deportivas, empresas, ayuntamientos, partidos políticos, museos, centros culturales, medios de comunicación y un largo etcétera, representan los socios de Padicat.

Por otro lado, desde la puesta en funcionamiento de su portal web en el año 2006, Padicat ha apostado por el eje de trabajo que ha resultado ser el más impactante: la creación de líneas de investigación específicas por medio de la integración focalizada de recursos digitales sobre determinados acontecimientos de la vida pública catalana. Se ha procedido a elaborar recopilaciones especiales de recursos web en eventos como las elecciones autonómicas de 2006, las municipales de 2007, las elecciones generales de 2008, o las recientes elecciones al Parlamento Europeo, de 2009. Asimismo, abordando otros ámbitos de la vida pública fuera del estrictamente político, también se ha elaborado una sección monográfica basada en la música folk y tradicional catalana, en colaboración con la Escuela Superior de Música de Catalunya (ESMUC), y se prepara otra monográfico sobre museos y colecciones museísticas en Cataluña.

Estas iniciativas han tenido una muy buena aceptación por parte de los medios de comunicación, y muy especialmente de parte de los profesionales universitarios especializados en cada materia, a los que se ha conseguido «integrar» para asesorar a la BC en la identificación y selección de los recursos digitales a recopilar.

De los tres ejes mencionados anteriormente que conforman el sistema híbrido, quizá sea el relativo a la captura masiva de recursos publicados en abierto (y especialmente la captura de aquellas webs bajo dominio .CAT), el que aporta una de las pocas sombras en la realización de los objetivos iniciales del repositorio. En 2006, el proyecto firmó un acuerdo de colaboración con la Fundació puntCAT (adminis-

tradora de los dominios .cat) para acceder a los 25.000 registros bajo este dominio. Pero más allá de diversas capturas parciales de estos registros, no se ha procedido, a octubre de 2009, a un proceso de captura masiva de estas páginas web, siendo la causa principal la capacidad limitada de los recursos destinados a captura y almacenamiento de las páginas web. La mejora en este punto supone uno de los retos de futuro del repositorio, como se apuntará más adelante.

Demos ahora una mirada más focalizada a determinados aspectos de funcionamiento del sistema, para una mejor descripción del depósito Padicat.

4. FUNCIONAMIENTO DEL SISTEMA

El sistema se basa en el ciclo documental clásico de bibliotecas y servicios de información (compilación, proceso, difusión), y, aplazando un análisis somero del sistema informático que permite el ciclo para más adelante, identificamos los pilares del proceso en la captura de los recursos, la organización de los mismos, y el acceso permanente a la colección.

En lo que respecta a la captura de los recursos se llevó a cabo un ejercicio de definición del tipo de recursos digitales susceptibles de captura, así como del alcance temático del proyecto. Es evidente que la tecnología que se aplica a los sistemas de repositorio digital cambia y cambiará en el futuro de forma exponencial, por ello las variables sobre la naturaleza del recurso digital y el *software* utilizado dotan de diferente grado de complejidad a lo que conocemos como páginas web.

Sin entrar a valorar el verdadero núcleo informativo de un recurso digital (¿todo él? ¿sólo la portada? ¿algunos capítulos? ¿un vídeo incrustado en el recurso?), sí citaremos la definición usada habitualmente por los miembros del Laboratorio de Internet del CINDOC-CSIC, que servirá para definir qué entendemos genéricamente por recurso digital, o «página web»⁶: Página web, o conjunto de páginas web ligadas jerárquicamente a una página principal, identificable por una URL y que forma una unidad documental reconocible e independiente de otras, bien por su temática, bien por su autoría, bien por su representatividad institucional.

Por tanto, entendemos que una web susceptible de formar parte de la colección del repositorio deberá cumplir dos condiciones básicas: será una página web identificable por una URL, y formará una unidad documental reconocible. No importa, pues, que ese recurso digital sea en un lenguaje de programación concreto,

⁶ Interesante reflexión terminológica en: Pareja, Víctor Manuel [*et al.*]. «Desarrollo y aplicación del concepto de sede web como unidad documental de análisis en Cibermetría». En: *9as Jornadas Españolas de Documentación*, 2005. El World Wide Web Consortium (<http://www.w3.org/2003/glossary>) define en su glosario el concepto «página web» como una colección de información consistente en uno o más recursos web, pensados para ser utilizados simultáneamente, e identificados con un único URI.

o el formato del recurso sea texto, imagen, sonido, etc. De hecho, ya las primeras pruebas del repositorio ofrecieron datos cristalinos sobre el elevado porcentaje de formatos estándares en las páginas web capturadas.

Por lo que respecta a la cobertura temática del repositorio Padicat, entendemos «Patrimonio Digital» como la información electrónica publicada en Internet, en abierto o no, independientemente del formato en que se presenta esta información. Entendemos «de Cataluña» en el sentido que tradicionalmente ha tenido la bibliografía nacional de Cataluña en que se basa la política de la BC: todo aquello producido en Cataluña, o que trate sobre Cataluña.

Lo cierto es que Internet está diseñada para romper barreras políticas y hacer la información accesible universalmente. Pese a este hecho definitorio, es posible identificar partes de esa red que contengan módulos de interés de grupos concretos, a los que podemos llamar «comunidades de usuarios web», referidas a cierta temática o simplemente de interés de una comunidad concreta. A efectos prácticos, se establece la estrategia de captura en: webs bajo dominio .CAT; o webs ubicadas en servidores de Cataluña; o webs bajo dominios geográficos (.ES, .COM, .NET, .ORG, etc.) en lengua catalana; o finalmente, webs que no cumplen los requisitos anteriores, pero relacionadas temáticamente con Cataluña.

5. EL SOFTWARE DE PADICAT

La organización de los recursos web, una vez capturados, debe permitir gestionar la colección y asegurar la recuperación, toda vez que ha de preservar los contenidos digitales con las técnicas disponibles. Esta organización incluye la identificación permanente de los recursos, la aplicación de metadatos, el almacenamiento, y la preservación.

Tras la fase de análisis y testeo de programas, se determinó que el programa informático Heritrix⁷, usado por la mayoría de los proyectos como el que nos ocupa para la captura de recursos digitales, sería el utilizado por el sistema. Este programa es el encargado de recolectar las páginas web tal como las ve el usuario que navega por Internet, y almacenarlas en archivos comprimidos en formato ARC⁸. Después, NutchWax⁹ y Hadoop¹⁰ realizan un proceso de indexación de la información recolectada que permitirá, posteriormente, utilizar estos índices para localizar recursos dentro de la colección.

⁷ Heritrix (<http://crawler.archive.org/>). Un artículo fundamental, a cargo de uno de sus creadores, es Mohr, G. [*et al.*]. «An introduction to Heritrix: an open source archival quality web crawler». En: *International Web Archiving Workshop*, 2004. [Consulta: 30/09/2009] <http://www.iwaw.net/04/Mohr.pdf>

⁸ Arc File Format ([http://en.wikipedia.org/wiki/ARC_\(file_format\)](http://en.wikipedia.org/wiki/ARC_(file_format))).

⁹ NutchWax (<http://archive-access.sourceforge.net/projects/nutch/>).

¹⁰ Hadoop (<http://hadoop.apache.org/core/>).

Existen dos interfaces para realizar las consultas al conjunto de recursos capturados: Wera¹¹, que permite la búsqueda por palabras clave a través de los índices generados por NutchWax; y Wayback¹², que permite la consulta directa por URL.

El programa Web Curator Tool¹³, desarrollado por la National Library of New Zealand, se ha aprovechado como sistema de gestión documental que permite la asignación de metadatos a una parte significativa de la colección, lo que garantiza la posibilidad futura de integrar la colección en otros catálogos de la BC u otras instituciones (aunque para ello, como veremos, será necesaria una evolución considerable de las actuales prestaciones informáticas).

Todo el *software* que utiliza Padicat es de código abierto y gratuito y ha sido desarrollado por organizaciones sin ánimo de lucro asociadas al International Internet Preservation Consortium (IIPC).

Por otra parte, el sistema prevé un depósito que permita conservar todos los recursos, de manera que se tenga acceso en todo momento. Se contempla un sistema de doble copia en diferente ubicación geográfica, con una capacidad total necesaria de 20 TB en los períodos de producción y explotación del repositorio, hasta 2011.

El método de preservación merece un tratamiento en comunicaciones adicionales al presente artículo, pero somos conscientes de la problemática de las estrategias más habituales de preservación¹⁴, como la migración periódica o *refresh* de los datos (migración a nuevas versiones de los mismos programas o lenguajes, o a nuevos programas capaces de leer los anteriores), la emulación (el uso de *software*, especificaciones, etc., utilizado en el momento de la creación), la recreación (simulación por ingeniería inversa u otros métodos).

Y es que las plataformas tecnológicas utilizadas para este tipo de depósitos están muy centradas en los tres aspectos básicos de la cadena documental: la captura de recursos; la indexación de estos recursos una vez capturados; y el acceso a los recursos almacenados. Pero en menor medida en la preservación de estos recursos, territorio prácticamente virgen si lo asociamos exclusivamente a los procesos dirigidos a garantizar el acceso permanente a las páginas web capturadas y no, como sucede en la mayoría de casos, si se analiza desde una óptica más global (preservación de materiales digitalizados, preservación de ficheros ofimáticos, etc.).

En todo caso, las previsiones sobre el tipo de archivos que el repositorio debe gestionar, basadas en la actual composición del fondo de la colección, revelan que la mayor parte de los archivos corresponden a formatos estándares, que pueden simplificar la tarea preservadora al menos en las macrocifras. Así, sobre una

¹¹ Wera (<http://archive-access.sourceforge.net/projects/wera/>).

¹² Wayback (<http://www.archive.org/web/web.php>).

¹³ Web Curator Tool (<http://webcurator.sourceforge.net/>).

¹⁴ Ayre, Catherine; Muir, Adrienne. «The right to preserve: the rights issues of digital preservation». En: *D-Lib magazine*, 2004, march, v. 10, n. 3. [Consulta: 30/09/2009] <http://www.dlib.org/dlib/march04/ayre/03ayre.html>

muestra cercana a los 107 millones de ficheros, aproximadamente el 69% de ellos corresponden a formatos estándares, lo que significa el 87% del espacio ocupado actualmente: texto/html (65%), imagen jpeg o gif (4%), etc.

En cuanto al *hardware* utilizado, creemos que una somera descripción de las máquinas será suficiente para dar por vista la vertiente más física del repositorio: Padicat tiene a su disposición siete nodos HP ProLiant DL360 G4p encargados de las tareas de recolección e indexación de webs. De la búsqueda y visualización de resultados en la interfaz web se encarga un *clúster* Linux de alta disponibilidad con características de balanceo de carga de peticiones y de tolerancia a fallos en caso de desastre en los nodos que componen la plataforma. Los nodos están conectados mediante fibra a una *Storage Area Network* (SAN) y el sistema se completa con un robot donde se guardan copias de seguridad de los datos en cinta.

6. RETOS PARA EL FUTURO

El futuro de Padicat, después de una etapa que podemos considerar de nacimiento, pasa por consolidar su capacidad de crecimiento, mejorar sus procesos de trabajo y optimizar los recursos de que dispone.

En primer lugar, hace falta dimensionar la infraestructura necesaria del repositorio, adaptándola a los objetivos del sistema, o bien modificar a la baja estos objetivos. La actual estructura de *hardware* y de personal experto en el *software* utilizado no permite trabajar con la capacidad necesaria para acometer el reto de la captura global de la web catalana. El hecho de tener al CESCO como socio tecnológico sin duda deberá permitir establecer cuáles son las necesidades, y en base a éstas poder dar una respuesta tecnológica para el crecimiento exponencial que perseguimos.

En segundo lugar, es imprescindible abordar la definición de las estrategias de preservación de los ficheros que contiene el repositorio. Probablemente sea éste uno de los aspectos clave en el retorno que la BC quiere ofrecer a la sociedad. Al margen de radiografías periódicas de la web catalana, que ilustran el diagnóstico del lenguaje de programación usado en la edición digital, el sistema puede ayudar a definir cuáles son los formatos que a corto plazo sufren problemas de ilegibilidad. Una vez constatadas estas pérdidas, es posible identificar hacia qué formatos hace falta transformar los ficheros para dotarlos de una permanencia más longeva, así como los procesos que han de permitir esta transformación.

En tercer lugar, Padicat ha de seguir apostando por potenciar los ejes de trabajo que hasta la fecha han dado los mejores resultados, como son las recopilaciones especiales sobre eventos de la vida pública catalana expresados a través de Internet, y aprovechar estos centros de interés para la implicación de los colectivos expertos en el asesoramiento a la BC de los yacimientos de recursos digitales.

En cuarto lugar, el abordaje de la captura sistematizada de publicaciones en serie en Internet es un reto de futuro, que se iniciará en los próximos meses con capturas que permitirán proyectar las necesidades infraestructurales del repositorio. La

revisión del *software* existente, para posibilitar el aprovechamiento de los ficheros ya descargados en sucesivas capturas cuando éstas se repitan con mucha frecuencia, será la solución a este reto, porque será también la manera de optimizar espacio en disco, tiempo de captura, y, en definitiva, los recursos existentes.

Finalmente, pese a la estandarización de los lenguajes informáticos que se utilizan en el *software* de Padicat y el resto de proyectos similares, hace falta destacar que no es aún posible, como cabía esperar, un intercambio eficaz de registros bibliográficos, con la finalidad de poder integrar todos los depósitos existentes, o estos depósitos en otros catálogos. El uso de pasarelas y lenguajes estándares está aún en fase de implementación en el *software* del repositorio que, insistimos, es común al de la mayoría de depósitos digitales existentes en la red. De la capacidad de incidir en el desarrollo del *software* depende también la consecución de los objetivos de futuro de la BC, en su voluntad de archivar la Web catalana.

En la presentación de los objetivos iniciales de Padicat, en 2005, se planteaban los potenciales beneficios de un proyecto que se encontraba en un estadio preliminar. Cuatro años más tarde, los beneficios son plenamente vigentes desde el momento en que han llegado a ser factores críticos de éxito en la estrategia de la BC: a ojos de la comunidad bibliotecaria de Cataluña, los beneficios se centran en la integración de los documentos nacidos digitales en la bibliografía nacional, y en el contundente posicionamiento de la Biblioteca de Catalunya y sus socios de repositorio en una situación privilegiada como fuente de información de los documentos que representan, en buena medida, el futuro. Posibilidades infinitas de cooperación con las instituciones de la memoria, bibliotecas, archivos y museos de Cataluña, así como universidades y centros de investigación, e impulso y liderazgo en la confección del patrimonio digital de España. Así mismo, la ya existente relación de privilegio con el resto de bibliotecas nacionales del mundo en términos de preservación digital y depósitos nacionales. Para las instituciones, empresas, administraciones y particulares que producen páginas web en Cataluña, preservación de la propia producción y garantía de acceso, con los condicionantes que rige la ley, a los contenidos y diseños que, de otra forma, desaparecerían. Y por último, para la ciudadanía, y como se pretende en las Directrices de la Unesco, acceso abierto y permanente a los recursos que forman el Patrimonio Digital de Cataluña.

7. BIBLIOGRAFÍA

- Biblioteca de Catalunya. *Memòria del plantejament del projecte PADICAT (Patrimoni Digital de Catalunya)*. Barcelona: Biblioteca de Catalunya, 2005. <<http://www.recercat.net/handle/2072/1757>> [Consulta: 30/09/2009]
- Biblioteques digitals i dipòsits nacionals de recursos digitals*. Barcelona: Universitat de Barcelona, Facultat de Biblioteconomia i Documentació, 1999.
- Cócera, D.; Llueca, C. (2008). «Padicat: realitat i reptes de 3 anys de l'arxiu web de Catalunya». *11es Jornades Catalanes d'Informació i Documentació*. Barcelona: Col·legi Oficial de Bibliotecaris-Documentalistes de Catalunya. <http://eprints.rclis.org/archive/00013562/01/llueca_padicat_jornades_2008.pdf>. [Consulta: 30/09/2009]

- Directrices para la preservación del patrimonio digital*. Canberra: Unesco, 2003. <<http://unesdoc.unesco.org/images/0013/001300/130071s.pdf>>. [Consulta: 30/09/2009]
- Gomes, D.; Silva, M. J. (2005). «Characterizing a National Community Web». *ACM Transactions on Internet Technology*, vol 5, num 3 (Aug 2005). <<http://xldb.fc.ul.pt/daniel/gomesCharacterizing.pdf>>. [Consulta: 30/09/2009]
- Hodge, G. M. (2000). «Best practices for digital archiving: an information life cycle approach». *D-LIB Magazine*. Vol. 6, num. 1 (jan 2000). <<http://www.dlib.org/dlib/january00/01hodge.html>>. [Consulta: 30/09/2009]
- Keefer, A.; Gallart, N. (2007). *La preservació de recursos digitals: el repte per a les biblioteques del segle XXI*. Barcelona: UOC.
- Llueca, C. (2005). «Webs siempre accesibles : las bibliotecas nacionales y los depósitos digitales nacionales». *BiD: textos universitaris de biblioteconomia i documentació*, núm. 15 (des 2005). <<http://www.ub.es/bid/15lluec2.htm>> [Consulta: 30/09/2009]
- Llueca, C. (2006). «El projecte PADICAT (Patrimoni Digital de Catalunya) de la Biblioteca de Catalunya», *10es Jornades Catalanes d'Informació i Documentació*. Barcelona: Col·legi Oficial de Bibliotecaris-Documentalistes de Catalunya. <http://eprints.rclis.org/archive/00006434/01/llueca_padicat.pdf>. [Consulta: 30/09/2009]
- Llueca, C. (2006). «Archivando la Web, el proyecto Padicat (Patrimonio Digital de Cataluña)». *El profesional de la información*. Vol. 15, núm. 6, p. 473-478. <http://eprints.rclis.org/archive/00007767/01/epi_padicat.pdf> [Consulta: 30/09/2009]
- Llueca, C. (2007). «Archivando la web catalana, el proyecto PADICAT», *Clip: Boletín de la SEDIC*. Núm. 47. <http://www.sedic.es/p_boletinclip47_confirma.htm>. [Consulta: 30/09/2009]
- Torres, N; y otros (2007). «Patrimoni Digital de Catalunya, experiències del primer any», *Jornades Tècniques RedIris*. Oviedo: RedIris. <http://www.padicat.cat/docs/poster_padicat_rediris.pdf> [Consulta: 30/09/2009]

4. atala

Informazio digitalaren unibertsoa

4.^a parte

El universo de la información digital

XI

Semantikan oinarritutako bilaketak: Kyoto proiektua

Iñaki Alegria eta German Rigau

Euskal Herriko Unibertsitateko irakasleak, IXA Taldea
Profesores de la Universidad del País Vasco, Grupo IXA

Laburpena: Dokumentazioaren kudeaketa digitalean informazioa atzitzeko deskribatzaileak erabiltzeaz gain testu bera erabili ahal izatea oso interesgarria da. Deskribatzaileak beraiek ere askotan testuak dira. Hizkuntza ingeniartzatza erabiltzaile hainbat aukera zabalitzen dira datu-baseen informazioa atzitzeko garaian: atzipen eleanitza, multzokatze semantikoa, antzekotasunean oinarritutako atzipena, galdera-erantzun sistemak, informazioa inferitzea... Semantikaren inguruko aukeretan sakonduko dugu, Europako beste unibertsitate eta enprekin Kyoto proiektuan lantzen ari garen ikerketa-ildoak azalduz.

Resumen: Investigaciones basadas en la semántica: Proyecto de Kyoto. En la gestión digital de la documentación puede ser muy interesante usar el propio texto además de los descriptores. Muchos descriptores también son texto. Usando técnicas de ingeniería lingüística se abren nuevas opciones de acceso a la información de estas bases de datos: acceso multilingüe, agrupación semántica, acceso basado en similitud, sistemas de pregunta-respuesta, inferencia de información... Profundizaremos en las posibilidades basadas en semántica, exponiendo las líneas de investigación que estamos desarrollando en el proyecto europeo Kyoto.

SARRERA

Eduki digitalen hedapenarekin eskura dugun informazio-bolumena izugarri hazi da eta bilaketak zaildu egin dira. Horren aurrean bilaketen teknologia asko garatu da, informazioaren atzipen eroso eta zehatza helburu. Gaur egun, *Google* eta *Yahoo* bezalako bilatzaileen garapenaren eraginez, bilaketen teknologia heldua da informazio-behar arruntentarako, dokumentu edo paragrafo batez asetzen diren beharretarako hain zuzen ere. Beste behar batzuetarako berriz gaur egungo teknologia ez da aski, galdera zehatzak erantzuteko edo informazioaren sintesia lortzeko (dokumentu anitzen sintesia batez ere). Arazoa areago korapilatzen duen ezaugarri bat eleanitzasuna da: askotan komenigarria edo ezinbestekoa da dokumentu eleanitzen artean bilatzea edo galderak hizkuntza desberdinetan onartzea. Azalduko dugunez helburu berri hauek lortzeko testuen interpretazio semantikoa ezinbesteko bitartekoa da. Edozein kasutan ez da nahastu behar testuen prozesaketa semantikoa eta web semantikoa. Lehena, hemen zabalduko duguna, hizkuntzaren tratamendu

automatikoarekin lotuta dago, bigarrena, berriz, dokumentuen egituraketa eta etiketatzearekin. Dokumentuen azterketa semantikoetan sakontzeko asmotan Kyoto izeneko proiektu bati ekin diogu EHU eta beste unibertsitate zein enpresaren artean.

Bilaketa-teknologia eta bereziki tresna linguistikoek egiten duten ekarpena azalduko dugu hasieran artikuluko honetan. Semantikaren azterketaren inguruan ari-tuko gara gero eta, bukatzeko, Kyoto proiektuaren nondik-norakoak azalduko di-tugu.

INFORMAZIOAREN BILAKETA ETA BERE ALDAERAK

Informazioaren berreskurapena (*Information Retrieval*, IR) ohiko arlo bat izan da informatikaren garapenaren hasieratik. Konputagailuek informazio kopuru handiak biltegitratzea posible egiten dutenez, informazio hori modu zehatz, eroso eta eraginkorrean berreskuratzea beti izan da aztergai garrantzitsua. Informazioaren berreskurapenaren kontzeptua testu-masa handien biltegitratze/berreskuratzearekin lotu ohi da informatikaren munduan [2]. Datu-base dokumentalak izan dira arlo honetako aplikazio garrantzitsuenak, eta bertan lantzen dira gaian gakoa diren bi urratsak: dokumentuen indexazioa eta ondorengo bilaketa.

Internet fenomenoak bultzatu egin du arlo honen garapena, testu digitalak izu-garri ugalduta direlako. IRren ohiko aplikazioez gain (testu legalak, medikuntza-koak, hemerrotekak, dokumentazio-zentroak...) Internet/Intranet eremuko aplikazio garrantzitsuenak kokatzen dira arlo honetan: *Google* moduko bilatzaileak eta *Yahoo* moduko direktorioak.

Duela gutxi arte, tresnen abiadura motela zela-eta, hizkuntza-ingeniaritzak ez du oso paper garrantzitsua jokatu arlo honen garapenean. Dena den, tresna linguistikoak hobetu diren heinean eta dokumentu digitalen eleaniztasuna areagotzearekin batera, tresna linguistikoaren erabilpena garrantzia hartzen joan da.

Ixa taldean¹ arlo honetan lan sakona egin dugu azken urteetan, ikuspegi elea-nitz batetik, baina euskararen gaineko bilaketaren problematikari erreparatu.

Bilaketaren teknologia azken urteetako teknologia arrakastatsua izan da. Inter-neteko zabalkundearekin batera bilatzaileak eguneroko tresnak dira jende asko eta askorentzat. *Google*² eta neurri txikiagoan *Yahoo*³ eta *bing*⁴ dira bilatzaile osoen eta erabilienak. Hizkuntza nagusiak baino ez dituzte kontuan hartzen eta, ondorioz, euskara bezalako hizkuntza flexiboentzat eta merkatu txikia dutenentzat arazoak daude ohiko bilatzaileetan. Horren aurrean bilaketako gaia desitxuratu (*energia*-ri buruz diharduten euskarazko dokumentuak bilatzeko *energia** eta du gakoak era-biltzea adib.) edo hizkuntza horretarako propio egindako bilatzaile bat erabiltzea

¹ <http://ixa.si.ehu.es>

² www.google.com

³ www.yahoo.com

⁴ www.bing.com

(euskarako *elebila*⁵) [4] dira aukerak. 1. irudian bilatzaile horien adibideak ikus daitezke.

Sistema hauen teknologia aski ezaguna da duela urte batzuetatik hona. Aurkitutako dokumentuak ordenatzeko garaian, dokumentuak duen edukiaz gain hainbat faktore hartzen dira kontuan: dokumentuak nondik erreferentziatzen dituzten, zenbat aldiz eta erreferentzia egiten duen gunearen esanguratasuna (*PageRank* algoritmoa [3] izan zen Google-ren arrakastaren hasierako gakoa). Informazio hori hipertestuaren topologiaren azterketaren bidez lortzen da. Hala ere, azken urteetan bilatzaileen arteko lehiakortasuna dela-eta, bilatzaileen barne-funtzionamenduari buruzko kaleratzen den informazioa murriztagoa da. Gainera, bilatzaileek webguneetara iristeko bide nagusi bihurtu diren heinean, enpresak saiatzen dira algoritmo horiei iskin egiten lehen postutan agertzeko (horri *search spam* deitzen zaio) eta bilatzaileek publiko ez diren neurriak hartzen dituzte horren aurka.

Emaitzak ebaluatzerakoan bi neurri erabiltzen dira doitasuna (*precisión*) eta estaldura (*recall*) [2]. Lehena eskuratutako dokumentuen esanguratasuna neurtzen du, hau da, eskuratutako dokumentuen artean zenbat dira esanguratsu edo interesgarri bilaketa-beharrerako. Estaldura, berriz, sistemaren eraginkortasuna neurtzen du, hau da, zeuden dokumentu esanguratsuen artean zenbat itzuli diren. Doitasuna kalkulatzeko erraza da, dokumentalista batek eskuratutako dokumentuak aztertu eta esanguratsuak zeintzuk diren markatuta. Estaldura neurtzea oso zaila da, bildumaren dokumentu guztiak begiratu beharko liriatekeelako, eta orokorrean modu erlatiboan neurtzen da, galdera beraren aurrean sistema batek beste batek baino dokumentu esanguratsu gehiago edo gutxiago itzultzen duenetz egiaztatuz.

Bilaketa-sistemetan gertatzen diren egoera arazotsuak honako hauek dira: galdera baten aurrean dokumenturik ez aurkitzea, edo erantzun gehiegi agertzea. Lehena konpontzeko estaldura handitu behar da, semantika erabiliz adibidez (ikusi geroago). Bigarrenari aurre egiteko emaitzak ordena egokian aurkeztea da gakoa, baina galderak fintzen laguntzeko tresnak ere badaude.

Dena den informazio-beharra dokumentu edo paragrafo baten eskurapenetik hara doanean gaurko teknologia ez da aski, bilaketaren eta inguruko arloetan azken urteetan aurrerakuntza nabarmenak egon diren arren. Gaiarekin lotuta honako teknika hauek azpimarra daitezke:

- CLIR (*Cross-Lingual IR*): bilaketa eleanitzean galderak eta dokumentuak hainbat hizkuntzatan egon daitezke eta bilatzailea gai izan behar da hizkuntza desberdinetako dokumentuak erlazionatzeko [6]. Hiztegi bat nahikoa izan daiteke horretarako, baina itzulpen automatikoa gero eta gehiago erabiltzen da.
- QA (*Question Answering*): dokumentuak bilatu beharrean erantzun zehatzak lortu nahi dira sistema hauetan (nork, non, zergatik...). Helburu horrekin dokumentuen prozesaketa sakonagoa behar da, hizkuntza-teknologiak

⁵ www.elebila.com

ezinbesteko tresna izanik. Emaitzak oraindik ez dira oso ikusgarriak baina ikerketa handia egiten da arlo honetan.

- Bilaketa multimodala: digitalizazioa dela-eta, dokumentuak bilatzeaz gain, argazkiak, irratiko edo telebista/bideoko programak bila daitezke. Bilaketa-sistema hauek metadatuetan (fitxa dokumentala) oinarri daitezke edo bes-telako teknketan (inguruko testua argazkietarako, ahots-testu bihurketa, etab.); eta askotan metodo konbinatuak dira egokienak. Gai hau artikulu honen esparrutik kanpo geratzen da.
- Dokumentuen sintesia: beharra asetzeko emaitza ez da dokumentu bat, do-kumentu batzuetan adierazitakoaren sintesia baizik. Pentsa dezakegu pro-dukto batez galdetzen dugunean dokumentu (eta agian hizkuntza) anitze-tan produktu horri buruz esaten diren gauza interesgarrienak lortu nahi di-gugula.

TRESNA LINGUISTIKOAK BILAKETARAKO APLIKAZIOAK

Hizkuntza-ingeniaritza arloko tresna linguistikoek bilatzaileen doitasuna edota estaldura handitzen lagundu dezakete, batez ere hemerroteketako zein liburutegi di-gitaletako bilatzaileetan. Konplexutasunaren arabera aurkeztuko ditugu, atal bakoi-tzean oinarrizko teknologia eta aplikazioa azalduz.

Morfologian oinarritutako tresnak

Oinarrizko teknologia hitza-lema erlazioa bilatzea da. Analisia esaten zaio hi-tzetik lema edo forma kanonikoa (hiztegiko sarrera) lortzen duen eragiketari, eta sorkuntza lematik bere forma posible guztiak lortzen dituenari. Flexio handiko hizkuntzetan tratamendu morfologikoa ezinbestekoa da estaldura oso txikia ez izateko.

The screenshot shows the BILATU search engine interface. At the top, there are language options (eu | es) and navigation links (Txertatu nabigatzailearen tresna-barran, Permal, Laguntza, Hori buruz). The search bar contains the text "bilaketa aurreratu" and "Hobespenak". Below the search bar, there is a "BILATU" button and a "BILAKETA AURRERATU" button. The search results section shows "Arabic Phonology" by F. Coromilas, Inaah Ouarfaly, and Dawood, published in 1977. The page also features a Google Scholar logo and a search bar with the text "diachronic linguistics".

1. irudia

Galdera zuzentzeko proposamenak

Bilatzaileetan prozesaketa morfologikoa bi modutan egin daiteke. Lehenengoan indexatze-prozesuan hitzen orde, edo hitzekin batera, lema gorde egiten dira indizeetan. Bigarrenean hitzak gorde egiten dira, baina bilaketa egitean sorkuntza egiten da, sortutako hitz guztiak bilatuz, *edo* eragileaz konbinaturik.

Teknologia linguistikoaz egin daiteke (lema, aurrizkiak, atzizkiak, paradigmak, aldaketa fonologikoak, etab. erabiliz), edo hurbilpen sinpleago batez, azken kasu horretan *stemming* edo sasilematizazioa esaten zaio eragiketari. Bilatzaile handietan sasilematizazio sinplea erabiltzen da, baina euskararen kasuan arrisku-sua izan daiteke doitasun-galera handi sor daitekeelako.

Aipatutako *elebila* bilatzaileak morfologia integratuta du sorkuntzaren bidez [4]. Beraz, hitz bat ematen dugunean bilatzeko bere lema forma guztiak (edo garrantzitsuenak) sortzen dira estaldura egokia ziurtatzeko.

Horrez gain, morfologia eta estatistika konbina daitezke galderetan erroreak detektatzeko eta proposamen egokiak emateko, 3. irudian ikus daitezkeen moduan.

Sintaxian oinarritutako aplikazioak

Testuen analisi sintaktiko sakona konputagailuz egitea ikergaia da gaur egun. Dena den analisi partzialak lortzen dituzten programak oso hedatuta daude eta ondo funtzionatzen dute. Orokorrean IR arloan izen sintagmak bilatzen dira, informazio garrantzitsua eskaintzen dutelako: termino berriak, pertsonak, enpresak/era-kundeak, tokiak...

Oinarrizko elementu horiek eta estatistikak konbinatuz aplikazio interesgarriak sor daitezke, dokumentuak estekatuz edo multzokatuz adibidez. Gainera erantzun gehiegi itzultzen dituzten galderak fintzen lagun dezakete aplikazio horiek.

Horren adibideak dira www.clusty.com, www.quintura.com, www.kartoo.com webguneak. Horietako batzuetan emaitzak erabiltzen dira bistaratze bereziak egiteko (multzoak, erlazioak...).

Euskarazko guneen artean www.zientzia.net aipa dezakegu, lematizazioa eta multzokatzea integratzen ditu-eta.

Eleaniztasuna. CLIR

Aplikazio eleanitzak egiteko hainbat aukera daude: hiztegi elebidun digitalak, ontologia eleanitzak eta itzulpen automatikoa. Ontologia eleanitzaren adibide gisa hurrengo atalean azalduko den *WordNet* dugu. Itzulpen automatiko zehatza [6] aspaldiko amets betegabea da, baina azken urteetan aurrerapen handiak egin dira, batez IR sistemetan integra daitezkeen sistema arin eta azkarren aldetik. Doitasuna ez da oso handia, baina aplikazioaren helburua itzultzea ez denez kalitate onargarria izan daiteke IR aplikazioetarako. Antzekotasun handiko hizkuntzen artean itzultzeko edo baliabide asko duten hizkuntza nagusien artean itzultzeko, gaur programa nahiko zehatzak daude.

Euskararen kasuan, oraindik gauza asko egin behar dira baina gaztelaniatik euskara itzultzen duen lehen sistema dago eskuragarri⁶. Kalitatea oraindik ez da egokia itzulpen profesionalean erabiltzeko baina halako aplikazioetan erabil daiteke. 2. irudian adibide bat ikus daiteke.

The screenshot shows a web browser window displaying the OpenTrad website. The browser's address bar shows the URL: <http://www.opentrad.org/demo/bs/nabigatzailea.php?language=eu>. The page content is in Spanish and features several news items:

- Ratzinger: beldurraren estrategia**: Article about Pope Benedict XVI's visit to Mexico.
- Dick Cheney Irak-era iristen da ustekabe bisitan**: Article about Dick Cheney's surprise visit to Iraq.
- Recomendada berriak kontsulta egin ditzan**: A section for recommended news.

The right sidebar contains a section titled 'Orrialde hau izena aipatzea' (Cite this page) and a list of names under 'Berria da:' (News is:).

2. irudia

Opentrad itzultzailearen adibidea

CLIR egiteko hainbat diseinu egin daitezke, bilduma hizkuntza bakar batean egotea, baina galderak edozein hizkuntzatan egin ahal izatea; galderak hizkuntza bakar batean baina dokumentuak eleanitzak izatea, edo aurreko bien konbinazioa, galderak eta dokumentuak eleanitzak. Are gehiago, argazkien edo bideoaren gaineko bilaketa eleanitza egin daiteke.

Horretarako hainbat aukera daude, galderak edota dokumentuak itzultzea hitzez hitz hiztegien bitartez, edo itzultzea itzulpen automatikoaren bidez. Aukera gehiago daude, ontologia batera proiektatu daitezke galderak eta dokumentuak. Kasu horretan bilaketa semantikoaz hitz egin daiteke.

⁶ www.opentrad.com

The screenshot shows a Google Translate search interface. The search query is "rock concerts in Moscow". The translated text is "рок-концерты в Москве". The search results are categorized as "Translated results from Russian web pages" and show three results:

- English translation:** Poster **Rock Concert Moscow - Rock Video On-Line - Playbill**. Posters **rock concert rock concerts Moscow** concert by the **rock band Surrounding rock:concert poster clubs in Moscow rock concert poster punk punk rock concert metal concerts...**. www.rock-video.net/afisha.htm - 31k - [Cached](#)
- Original Russian:** Афиша **Рок Концертов Москва - Rock Video On-Line - Афиша**. Афиша **рок концертов Рок концерты москва Концерт рок группы Ближайшие рок концерты афиша клубомосквы рок афиша панк концерт панк рок концерт метал концерты ...**. www.rock-video.net/afisha.htm - 31k - [Сохранено в кэше](#)
- Tickets for the concert poster and concerts in Moscow 2009 - rock and...**. Order biletov **concerts Moscow 2009**, poster, **rock concerts** in the Olympic. www.hotconcert.ru/ - 31k - [Cached](#)
- Билеты на концерт и афиша концертов в Москве 2009 - рок и ...**. Заказ билетов на **концерты москвы 2009**, афиша, **рок концерты** в Олимпийском. www.hotconcert.ru/ - 31k - [Сохранено в кэше](#)
- Poster concerts rock bands in clubs of Moscow - the best concerts in...**. Every day in the clubs of **Moscow** goes through many of **concerts** and of course the lion's share of them are **rock concerts** of groups. Gigs in clubs in **Moscow** allow... www.hotconcert.ru/events/20/ - 26k - [Cached](#)
- Афиша концертов рок групп в клубах Москвы - лучшие концерты в ...**. Каждый день в клубах **Москвы** проходит множество **концертов** и конечно львиная доля из них это **концерты рок групп**. **Концерты в клубах в Москве** позволяют ...

3. irudia

Translated Search funtzioaren adibidea

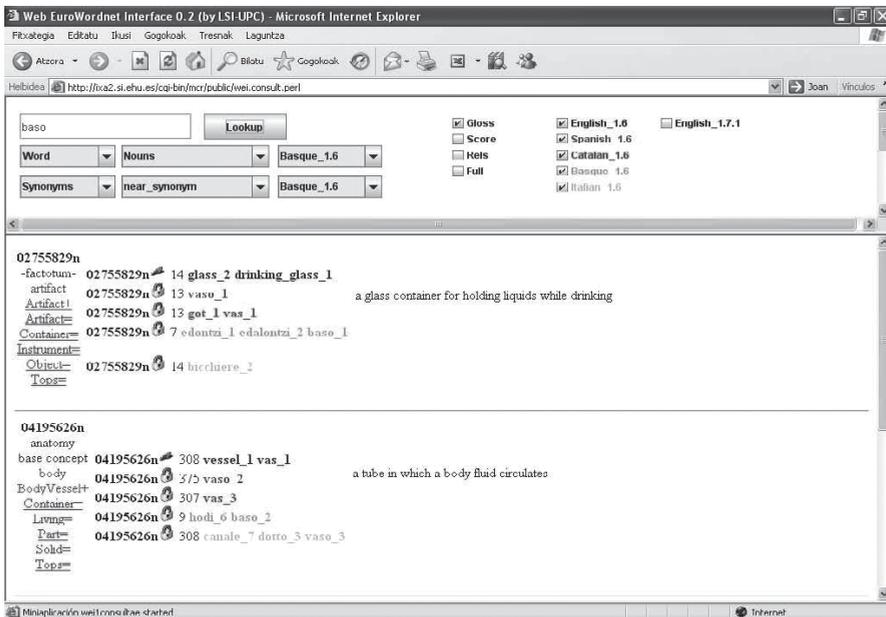
Eleaniztasunean oinarritutako aplikazioen adibide gisa 3. irudiko adibidea dugu Googleren *Translated Search* delakoa itzulpen automatikoan oinarrituta dagoena.

SEMANTIKA. BILAKETA SEMANTIKOA

Bilaketa semantikoaren bidez hitzen bidezko bilaketan dauden arazo biri erantzun nahi zaie: adiera anbiguotasuna batetik (*baso* bilatzen badugu arbolen basoa edo edateko basoari buruzko dokumentuak atzitzeko genituzke), eta sinonimia edo espezializazioa bestetik (*jatetxe* bilatzen badugu, *sagardotegi* edo *erretegi* hitzak azaltzen diren dokumentuak ez genituzke lortuko). Bilaketa kontzeptuen bitartez egingo balitz, bi arazo horiek arinduko lirateke. Problema bat dago, tratamendu semantikoa nahikoa zehatza ez bada estaldura handitu arren doitasuna galdu egiten dela.

Orokorrean kontzeptuen inbentario diren ontologiak eleaniztunak balira, orduan hizkuntzen arteko bilaketak ere hobeto egingo lirateke, kontzeptu gehienak konpartitzen baitira hizkuntzen artean. Web 2.0 delakoan erabiltzen diren folksonomiak ez bezala, analisi semantikorako taxonomia eta ontologia formalagoak erabili ohi dira, eta horien artean *WordNet* (Ingeleserako ontologia elebakarra) eredu

jarraitzen duten wordnet eleaniztunak dira erabilienak, IXA taldean beste erakunde batzuekin elkarlanean landu den *Multilingual Central Repository* (MCR) kasu [1]. MCR delakoak ingelera, gaztelera, italiara, katalanera eta euskara biltzen ditu bere baitan. Bertan hizkuntzetatik independente izan nahi duen kontzeptu inbentario bat dago, eta kontzeptu horiei buruzko informazio semantiko aberatsa jasotzen da, taxonomia egiturak barne. Azken urteetan egindako ikerketari esker, kontzeptuei buruzko hainbat informazio sartu izan da MCRen. Beheko irudian euskarazko baso-ri dagozkion bi adiera azaltzen dira, beste hizkuntzako itzulpenekin batera, eta kontzeptuaren hainbat tasun semantikorekin batera (adibidez, gizakiak egindako dela edateko basoa, edukitzeko ontzi baten propietateak dituela, instrumentu bezala erabili daitekeela, etab.). Hurrengo irudian azaltzen dira ere



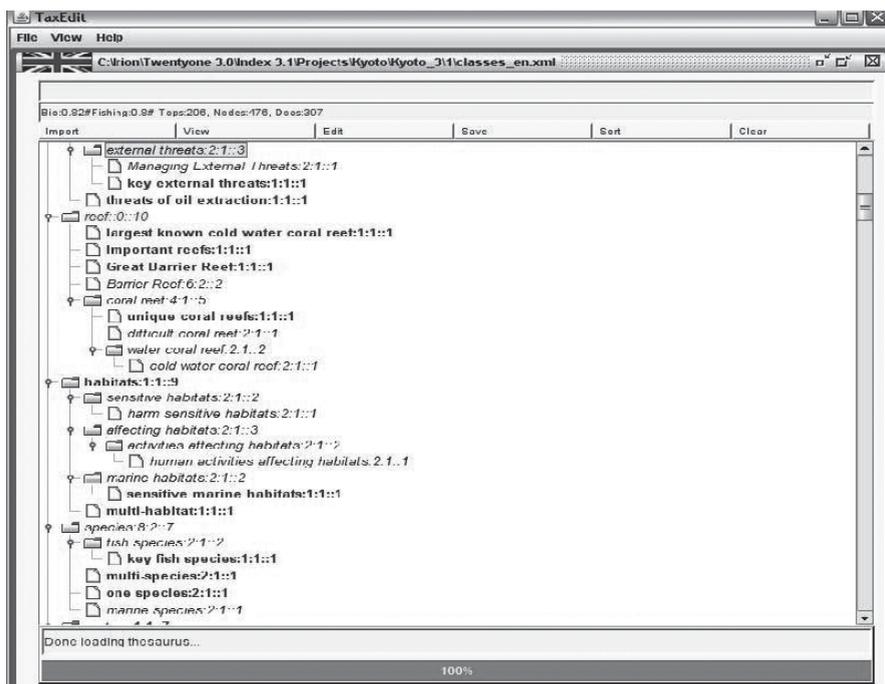
4. irudia

WordNet interfazearen adibidea (*glass-edalontzi*, *baso*)

KYOTO PROIEKTUA

Asia eta Europako hainbat erakunderen artean garatzen ari den KYOTO proiektuaren (www.kyoto-project.eu) helburua ezagutza modelatzeko eta gertaerak identifikatzeko plataforma bat garatzea da [7]. Garapena komunitatean oinarrituta dago eta hizkuntzen eta kulturen artean erabilgarri izan dadin diseinatu

da. Wiki sisteman oinarrizten den plataformaren bitartez erabiltzaileek domeinu zehaztutako terminologia adostu dezakete, terminoen erauzketa dokumentu eleantzetatik modu automatikoan lortu eta gero. Adituek termino horiek aldatu, aberastu eta erlazionatu ditzakete. Programen eta adituen lanaren ondorioz ezagutza-egitura konplexuak sortzen dira, hizkuntzarekiko neutralak direnak eta ezkutuan geratzen direnak, baina erabil daitezkeenak testu-bilduma berriak ustiatzeko eta bertatik gertaerak inferitzeko.



5. irudia

Tybot-aren emaitza

Gertaeren datu-base bat sortuko da erabileraren ondorioz eta gertaera horietan zehar bilaketak egin eta nabigatu ahal izango dute erabiltzaileek. Sistema zazpi hizkuntzatarako garatzen ari da eta ingurumena da landutako domeinua, baina beste hizkuntzatarako edo domeinutarako da hedagarria.

Hizkuntzen eta kulturen arteko elkarrengiketa lortzeko partekatutako ontologia bat erabiltzen da, zeinaren bitartez hitzak eta espresioak estekatzen diren. Ontologia errepresentazio formal eta hizkuntzatik independentea da, inferentziarako zein arrazoinamendurako erabil daitekeena. Wiki ingurune baten bitartez lankidetzatresna bat garatu da erabiltzaileek erauzitako informazioa gainbegiratu eta edita de-

zaten. Informazioaren erazketan aurretik definitutako ontologia eta berari estekatu-tako espresioak erabiltzen dira patroio ontologiko izeneko predikatuak erabiliz.

Sistema martxan jarri baino lehen adituen informazio-beharrak aztertu dira. Hauexek dira adibideetako batzuk (ingelesez):

- *Which are the most suitable areas in Europe for pro biodiversity business?*
- *What are the key biodiversity indicators in a certain area?*
- *What is the effect of hedgerows on air quality?*
- *What is the impact of dogs on wildlife?*
- *Are there huge negative effects with regard to eco-networks and alien invasive species?*

Ikus daitekeenez espresioetan termino konplexuak (*pro biodiversity business, biodiversity, eco-networks, alien invasive species*) eta erlazio kausalak (*indicators, impact, effect*) agertzen dira. Ohiko bilatzaileetan arazoak daude galdera hauetara erantzun gisa emaitza interesgarriak lortzeko. Kyoto proiektuaren helburua horixe da, gai zehatz baten inguruko galdera konplexuei erantzun egokia ematea. Horretarako hiru helburu lortzeko lanean gaude:

Osatu den testu-bildumatik terminologia eta kontzeptuak hartzen dira automatikoki. Hau egiteko *Tybot* izeneko osagaia eraiki da.

Wikyoto izeneko wiki plataforma bat eratu da adituek hizkuntza eta kulturen artean terminoen eta kontzeptuen esanahia adostu eta argi dezaten. Plataforma honen bitartez adituek terminoen hautapena egin eta sare semantikoan kokatu dezakete, geroago ezagutza hori partekatutako ontologian integra dadin.

The screenshot displays the Wikyoto Knowledge Editor interface. At the top, the browser window title is 'KYOTO Project - Wikyoto Knowledge Editor - Mozilla Firefox'. The address bar shows the URL 'http://www.group.it.cnr.it/demos/Wikyoto_Term_Editor/Term_Editor.php'. The interface is divided into several panels:

- Search for a term:** A text input field with the placeholder 'Type the term you want to search for...'.
- Browse KYOTO Term Hierarchies:** A tree view showing a hierarchy of terms starting with 'frog', including sub-terms like 'endemic frogs', 'forest frog', 'common frog', 'poison frog', 'golden poison frog', 'gopher frog', 'Dusky gopher frog', and 'bullfrog'.
- Browse Domain WordNet Hierarchies:** A tree view showing a hierarchy of terms starting with 'any of various tailless stout-bodied amphibians', including sub-terms like 'endemic frogs (3) frogs that live in...' and 'bullfrog (1) amphibian showing...'.
- Selected terms:** A section showing 'frog' with a table of WordNet pointers. The table has columns for 'LOAD', 'Lexicon', 'Sense', 'Doc. occurrences', 'SHOW OCC.', and 'Gloss'.

LOAD	Lexicon	Sense	Doc. occurrences	SHOW OCC.	Gloss
LOAD	PWN00	01829785-n	7	SHOW	[frog, toad, toad frog, anuran, batrachian, s...
LOAD	PWN00	09170045-n	7	SHOW	[frog, Gaul] (n) a person of French descent
LOAD	PWN00	03390362-n	7	SHOW	[frog] (n) a decorative top of bread or cord
- System: any of various tailless stout-bodied amphibians:** A section showing a lemma 'frog' with its internal and external references.

Lemma	Internal ref.	External ref.
ADD Lemma		
DELETE	Lemma	Sense number
DEL	frog	2

6. irudia

Wikyoto-aren itxura

Testu-erazketarako tresna eleanitz ahaltsu batek terminologia eta ontologia erabiltzen du testu bildumetatik informazio eta gertaera esanguratsuak lortzeko. Aurrekoaren ondorioz datu-bilduma partekatu bakarra sortzen da, bertan bilaketen eginbeharra modu adimentsuan burutu ahal izateko. Tresna honi *Kybot* deritzo.

Bistan da urrats horietan semantika gunea dela. Semantikan oinarritutako prozesaketa aurrera eramateko lau osagai integratzen dira:

- Hizkuntza/kulturari dagokion Wordnet-a, terminoak eta erlazioak jasotzen dituena.
- Termino horien definizioa ontologian, kontzeptu abstraktuak terminoekin lotzeko.
- Kontzeptuen definizioa, hizkuntzatik eta kulturatik neutrala dena.
- Hizkuntzarekin lotutako terminoak eta hizkuntzatik independenteak diren kontzeptuak lotzeko moduaren definizioa.

Gaur egun proiektua erdibidean dago eta oraindik ez dago emaitzarik. Edozein kasutan aipatutako elementu nagusien prototipoak eraiki ditugu eta test fasean daude.

ONDORIOAK

Artikuluaz azaldu dugunez, teknologia linguistikoak, tartean semantika, ekarpen handia egin dezake dokumentu digitalen bilaketen arloan. Bilatzaile orokor ezagunenetan oraindik gutxi erabiltzen badira ere, erabilera handitzen doa, orokorrean estaldura handitzeko oso tresna interesgarriak direlako. Tresna linguistiko hauek (sintaxia, itzulpena, semantika) ez dira maizago erabiltzen oraingoz ez direlako nahiko eraginkorrak, prozesaketa-denboraren aldetik, prozesatzeko behar den memoriaren aldetik edo doitasunaren aldetik.

Dena den gero eta gehiago erabiltzen ari dira, eta aipatutako mugak gainditzeko helburuarekin hainbat ikerketa egiten dira, tartean Kyoto proiektua, zeinaren bitartez ezagutza sakona behar duten informazio-behar konplexuei erantzun egokia eman nahi zaion.

BIBLIOGRAFIA

- [1] Agirre E., Alegria I., Rigau G., Vossen P. 2007. MCR for CLIR. *SEPLN aldizkaria, monografia TIIMM*. vol 38, 3-16. ISSN 1135-5948. <http://ixa.si.ehu.es/Ixa/Argitalpenak/Artikuluak/1174895701/publikoak/pdf>
- [2] Baeza-Yates R., Ribeiro-Neto B. 1999. *Modern Information Retrieval*. ACM Press Series/Addison Wesley, 1999
- [3] Brin S., Page L. 1998. *The anatomy of a large-scale hypertextual Web search engine Computer Networks and ISDN Systems*, Elsevier.
- [4] Leturia I., Gurrutxaga A., Areta N., Alegria I., Ezeiza A. 2007. «EusBila, a search service designed for the agglutinative nature of Basque». *SIGIR2007- iNEWS'07 workshop*. ISBN

- 978-84-690-6978-3. <http://ixa.si.ehu.es/Ixa/Argitalpenak/Artikuluak/1190627800/publikoak/pdf>
- [5] Mayor, A. 2007. *MATXIN: Erregeletan oinarritutako itzulpen automatikoko sistema baten eraikuntza estaldura handiko baliabide linguistikoak berrerabiliz*. Doktorego-tesia. Euskal Herriko Unibertsitateko; Donostiako Informatika Fakultatea. <http://ixa.si.ehu.es/Ixa/Argitalpenak/Tesiak/1196444990/publikoak/Matxin2007.pdf>
- [6] Saralegi X., Alegria I. 2007. «Similitud entre documentos multilingües de carácter técnico en un entorno Web». *SEPLN aldizkaria, 2007*. Sevilla. ISSN 1135-5948. <http://ixa.si.ehu.es/Ixa/Argitalpenak/Artikuluak/1184248312/publikoak/pdf>
- [7] Vossen P., E. Agirre, F. Bond, W. Bosma, C. Fellbaum, A. Hicks, S. Hsieh, H. Isahara, Ch. Huang, K. Kanzaki, A. Marchetti, G. Rigau, F. Ronzano, R. Segers, M. Tesconi: «KYOTO: a Wiki for Establishing Semantic Interoperability for Knowledge Sharing across Languages and Cultures», in: *Handbook of Research on Culturally-Aware Information Technology: Perspectives and Models*. IGI Global USA.

IXA taldea

Lang: eu | en

Nor gara?

- Etvera
- Aurkezpena
- Kideak

Zer egiten dugu?

- Ikerlerroak
- Argitalpenak
- Produktuak
- Proiektuak
- HAP masterra

Beste batzuk

- Estekak
- Lanpoltsa
- Karrera bukaerako proiektuak
- Pribatua

HIZKUNTZAREN AZTERKETA ETA PROZESAMENDUA MASTERRA

Kontaktua

Ixa Taldea
649 Posta kutxa
20080 Donostia
ixa@ehu.es

Euskal Herriko Unibertsitatea

IXA taldea
Hizkuntzaren prozesamendua

Bilatu nahi duzun

Hemen zaude: Ixa

PIL-PILEAN

Udako ikastaroa: Hizkuntzen kudeaketa mundu global batean **2010-Uzt-01**

Laponiako samiera hizkuntza eta euskara lankidetzan aztertzen **2010-Eka-28**

Hizkuntzaren Azterketa eta Prozesamendua. Master ofiziala **2010-2011 2010-Eka-18**

CLARIN proiektuaren bilera Euskal Herriko agenteekin **2010-Eka-14**

Berri gehiago ikusi

<http://ixa.sii.ehu.es/ixa> (1 de 2)15/09/2010 14:21:21

IXA taldea

Aurkezpena

IXA taldea Euskal Herriko Unibertsitateko ikerkuntza-taldea da, hizkuntzaren tratamendu automatikoa lan egiten duena. Momentu honetan taldeko kideak 43 gara: 32 informatikari, 8 hizkuntzalari, 2 ikerkuntzarako teknikari eta administrari bat. Gehienak Donostiako Informatika Fakultatean aritzen gara, baina beste zentro batzuetan ere bai (Bilboko Industria Ingeniaritza Teknikoko Eskola, Donostiako Irakasle Eskola...)

Argitaratu ditugun [artikulu guztiak](#) ikus ditzakezu webgune honetan, baita gure [proiektuak](#), [baliabide linguistikoak](#), [produktuak](#)...

Sortu ditugun produktu ezagunenak hauek dira: [Xuxen](#) zuzentzaile ortografikoa, [OpenTrad](#) itzultzaile automatikoa, Euskal WordNet sarea eta [ZT](#) eta [EPEC corpusak](#).

Esku artean ditugun proiektu nagusiak hauek dira: [KYOTO](#) Europako STREP proiektua; Eusko Jauriaritzaren [A motako ikertalde finkatua](#); M.E. C.eko [IMLT](#), [OpenMT-2](#) eta [KNOW2](#) proiektuak eta [TIMM](#) eta [RTH](#) ikerketa-sareak. Ikus [hemen](#) proiektu guztien lista 1988tik.

 [Gehiago irakurri](#)

Proiektuak



A motako ikertalde finkatua

TIMM OPENMT²

KNOW²



IMLT

XII

Informazio kudeaketa, sindikazioaren eta tresna folksonomikoen bidez

Gorka Julio

Elurnet enpresako ingeniari teknikoa
Ingeniero técnico de la empresa Elurnet

Laburpena: Sindikazioa aspalditik erabilgarri bazegoen ere, ez da blogen ez-tanda iritsi arte hedatu. Orain esparru hori erabat gainditu du, eta Interneteko leku gehienetan aurkitzen dugu sindikatzeo aukera. Horrek informazio iturri ugari-tan harpidetzeko aukera ematen digu, betiere informatuak egon gaitezen. Gero eta garrantzitsuagoa da informazioaren kudeaketa egokia egitea. Gure esparruko berriak ezagutu, gure enpresa, talde, elkarteari buruz esaten dena ezagutu, alertak jaso, etab. egin daiteke tresna hauen bidez. Informazioa jasotzeaz gain, informa-zio hori besteekin (talde publiko zein pribatuekin) era parte-hartzaile batean tru-katzeko aukera handiak ematen ditu. Gure edukiak gure esparrutik kanpo zabal-tzeko aukera ere ematen dute jarioek. Horien formatu estandarri esker, edukia mila lekutan jartzeko aukera ematen dute esfortzu minimo bat eginez. Gaur egun, enpresa, elkarte edo taldeek sindikazioan oinarritutako informazio kudeaketa be-rritzailea egiteko beharra dute.

Resumen: Gestión de la información a través de la sindicación y de las he-rramientas folksonómicas. La sindicación es útil desde hace tiempo pero no se ha propagado hasta el estallido de los blogs. Ese ámbito ya lo ha superado y en la actualidad en la mayoría de sitios tenemos opción de sindicarnos. Esto nos da la posibilidad de suscribirnos a una cantidad abundante de distintas fuen-tes de información para que siempre permanezcamos informados. Cada vez es más importante hacer una correcta gestión de la información. Conocer las no-ticias de nuestro entorno, nuestra empresa, grupo, conocer lo que se dice sobre la sociedad, recibir avisos... es posible a través de estas herramientas. Aparte de recibir información nos brinda la posibilidad de intercambiar información tanto con grupos públicos como privados. Los flujos también nos dan la posibi-lidad de propagar nuestros contenidos fuera de nuestros ámbitos. Haciendo un mínimo esfuerzo, gracias a estos formatos estándar tenemos la opción de poner estos contenidos en múltiples lugares. Las empresas, asociaciones y grupos de hoy en día deben basarse en la sindicación para una gestión innovadora de la información.

Web sozialaren garaian informazio uholdea oso handia izan daiteke, baina era berean hau kontrolatu eta erabili ahal izateko tresna berriak ditugu. Horien artean bi tresna garrantzitsu daude: eduki sindikazioa eta tresna folksonomikoak.

Informazioa jaso eta erabiltzeko edozein prozesutan bezala, aurrelan bat dago. Prozesu horretan beharrak detektatu, informazioa bilatu eta jaso, informazioa tratatu, informazio hori hedatu eta, berriro ere, prozesuaren analisi bat egin beharko da.

Testuingurua ere garrantzitsua da, web 2.0 edo web soziala deitu den horretan kokatzen gara. Bertan egon eta eragin ahal izateko entzun, eztabaidatu eta bukatu bertan eragiteko bideak sortu behar ditugu. Entzuteko fasean batez ere da informazio kudeaketaren inguruan guri dagokiguna, eta hor sindikazioak eta tresna folksonomikoek asko lagunduko digute.

Sindikazioa edukien eguneraketan beste pauso bat da. Informazioaren bila joan beharrean jario irakurgailuen bidez harpidetutako iturrietako informazioa eguneratzen denean bakarrik jasoko dugu. Hori posible egin ahal izateko, webguneek harpidetzeko aukera ematen duten jarioak eskaini behar dituzte; jario hauek formatu ezberdinetakoak izan daitezke, baina orokorrean bi formatu familia topatu daitezke RSS eta Atom familiak. RSS familiaren barnean mota ezberdinak badaude ere denak, bakar batean txertatzen dira. Azken finean, informazioa prozesagarria izango den formatu batzuetan ematen duten fitxategiak dira. Honi esker jario irakurgailuek hauetara harpidetu eta informazioa jaso ahal izango dute.

Jario irakurgailuak mota ezberdinetakoak izan daitezke. Web zerbitzuak edo instalatu beharreko aplikazioak izan daitezke. Web zerbitzuak erabiltzeko nabigatzaile bat beharko dugu, kontu bat sortu eta erabiltzen has gaitezke, Google Reader (<http://reader.google.com>) edota Bloglines (<http://bloglines.com>) adibidez. Besteak aldiz, gure ordenagailuan instalatu behar izango ditugu erabiltzeko. Honelakoak dira Liferea (<http://liferea.sourceforge.net>), RSSbandit (<http://rssbandit.org>) edo FeedReader (<http://feedreader.com>).

Sindikazioarekin loturik, podcast-a dago. Podcasta audio edo bideo fitxategiak lotzen dituzten jario bereziak dira. Normalean jarioan dagoen *enclosure* etiketa baten definitzen da audio edo bideo fitxategi hori zein helbidetik hartu daitekeen. Honela jario irakurgailu bereziek fitxategi hori jaitsi edo erabili ahal izango dute. Era honetako jarioetara harpidetzeko jario irakurgailu bereziak daude. Audio podcast-ei dagokionez ezagunenak Juice (<http://juicereceiver.sourceforge.net/>), iTunes (<http://www.apple.com/es/itunes/overview/>), Songbird (<http://getsongbird.com>). Bideorako aldiz Miro (<http://getmiro.com>) da interesgarriena.

Sindikazio tresnak harpidetzeko aukera ematen badute, formatu estandarrek izanik beste hainbat gauzetarako balio dute. Hainbat iturritako informazioa leku bakar batean bistaratu, jarioetako informazioa beste formatu batzuetan erakutsi, etab.

Askotan arazoa ez da informazio iturri egokiak aurkitzea baizik eta interesatzen ez zaigun informazioa iragaztea. Hori ekidin ahal izateko informazio iturriak filtratzeko aukera ematen duten beste hainbat tresna daude, horietako bat Yahoo Pipes (<http://pipes.yahoo.com>). Honi jarioetako informazioa batu, iragazi, itzuli ahal izango dugu. Oso tresna interesgarria da, baina ikasketa prozesu txiki bat beharrezkoa da.

Baina, sindikazioa erabili ahal izateko webgune edo zerbitzuak jarioak eskaini behar ditu. Interesatzen zaizkigunek ez badute sindikatzeo aukera ematen, horiek

sailkatuta eta gorde ahal izateko tresna folksonomikoak erabili daitezke. Tresna folksonomiko hauen bidez interesatzen zaigun horren URLa gorde, deskribapen bat gehitu eta deskriptore edo etiketak gaineratuko dizkiogu. Honela, etiketa sorta horrek amaieran egitura folksonomiko bat sortuko digu, etiketa laino baten bidez edota bilaketen bidez aurretik gorde dugun baliabide hori aurkitzeko aukera erraztuz.

Tresna hauek interes handiagoa dute baldin eta jende kopuru handienak erabili. Bakoitzak etiketa bidez egin dezakeen sailkapenaz gain, tresna erabiltzen duten beste erabiltzaileek egindako sailkapenak ere interesgarriak izan daitezke. Azkenean, bakoitzak egindako bildumez osatutako egitura folksonomiko bat sortzen da. Horrelako tresnak dira delicious (<http://delicious.com>) edo euskarazko Bildu (<http://bildu.net>).

Hala ere, gaur egun tresna hauek behar izateak, WWWean metadatuaren falta erakusten du. Etorkizunean webgune eta zerbitzuek informazio semantiko gehiago izango dute eta orain sindikazioa edo tresna folksonomikoekin egiten ditugun nahasketa eta aukeraketa guztiak, eskala handiagoan egin ahal izango ditugu. Mikroformatuek (<http://microformats.org>) adibidez, orain arteko html etiketak erabiliz, informazioa eduki dezaketen kode zatiei informazio semantikoa gehitzen saiatzen dira. Ekitaldiak kodetzeko hCalendar mikroformatua erabiltzen da, kontaktuak kodetzeko hCard edota kritikentzako hReview.

Baina, hauek guztiak adibide oso txikiak dira. Badira beste teknologia batzuek, RDF kasu, etorkizunean webguneetako informazioa datu base baten moduan atzitzeko aukera emango digutenak. Horretarako gure webgune eta zerbitzuak prestatzen joango beharko gara. WWWean dagoen informazioa era automatikoan atzitzeko aukera dugunean, aukera berri asko agertuko dira.

XIII

Una visión de la gestión de contenidos web y el rol del gestor de información digital

Fernando Fernández de Aránguiz

Responsable de Proyecto de la empresa ODEI
ODEI enpresako Proiektu arduraduna

Resumen: La ponencia analiza qué es un Gestor de contenidos web (WCM) y qué tipos de gestores hay, desde soluciones empresariales, de software libre o soluciones a medida. También trata la integración de estos gestores con el resto de sistemas de información, así como el rol del gestor de información digital, su perfil, conocimientos, funciones, etc.

Laburpena: Web edukien kudeaketaren ikuspegi bat eta informazio digitalaren kudeatzailearen eginkizuna. Ponentziak web edukien kudeatzaileak (WCM) zer diren eta zein motatako kudeatzaileak dauden aztertzen du, bai enpresentzat pentsatutakoak, software librekoak edota neurrira egindakoa. Kudeatzaile horiek gainontzeko informazio sistemekin nola integratzen diren ere aztertzen du, baita informazio digitalaren kudeatzailearen betekizuna, ezaugarriak, ezagutzak, funtzioak, etab. zein diren ere.

Provengo de la empresa ODEI, una empresa de Vitoria que se dedica a los servicios y a las tecnologías de la información. Desarrollamos sistemas de información de todo tipo, haciendo desde la consultoría inicial —para ver la oportunidad de llevar a cabo proyectos y cómo hacerlos—, hasta el desarrollo de estos proyectos e incluso, en algunos casos, su explotación, ya que tenemos gente experta para llevar a cabo esos sistemas. Esta sería una dimensión de nuestro trabajo; la otra dimensión sería repetir esto tantas veces como tipos de sistemas de información haya: sistemas de información generalmente documentales, temas de archivo, seguimiento de prensa, bibliotecas...; y, también, webs más convencionales, de explotación de datos, de proceso de negocio de las empresas, etc. Con esta experiencia como bagaje, voy a hablar de la gestión de contenidos web: qué es, qué modelos hay, qué herramientas se emplean y cómo se integran con otros sistemas de información; y también voy a tratar el papel del gestor de información digital: cuáles son sus funciones, qué conocimientos debe poseer, etc.

1. GESTIÓN DE CONTENIDOS WEB

Los gestores de información necesitan comunicar algo, y lo hacen a través de los contenidos web. Esos contenidos web se encajan en una página web, que está

dentro de una estructura mayor que le da sentido y que permite llegar a esa página o intentar localizarla. Esos sitios web forman parte de la World Wide Web, que sería el conjunto de todos los sitios web navegables; y esta www forma parte de Internet. A veces, Web e Internet se utilizan como sinónimos, pero tienen su diferencia. Internet sería la red que permite comunicarnos y navegar, pero esa red nos posibilita muchas otras cosas, como transferir ficheros, intercambiar correo electrónico, etc.

Pero lo más importante de todo, lo que le da sentido a todo esto, es que detrás hay un usuario, hay un consumidor de esa información. Vemos cómo usuarios y gestores de información —aunque tengan unas capas tecnológicas intermedias— están muy unidos: si hay un gestor de información es porque en el otro lado hay un usuario que va a demandar esa información.

1.1. Qué es la gestión de contenidos web

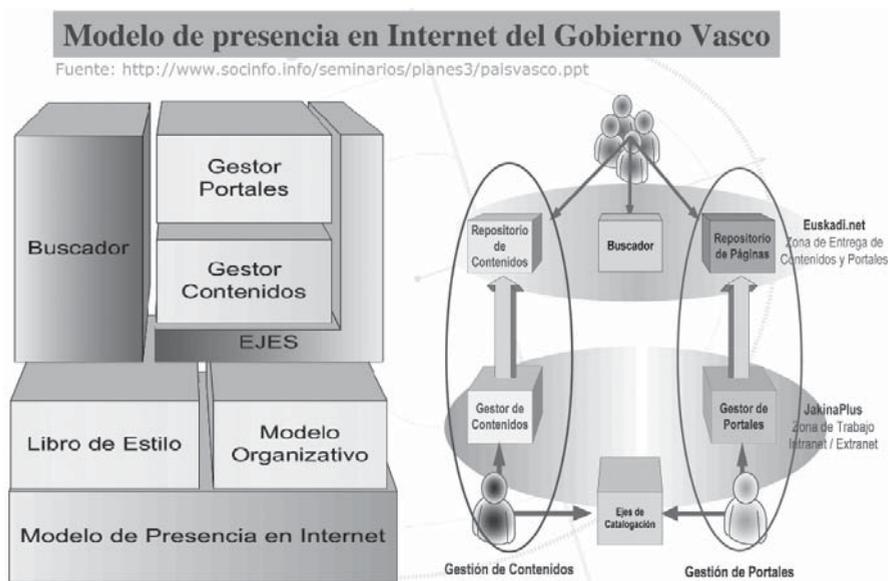
Primero vamos a mencionar cómo son esos contenidos web, cómo se encajan, para explicar a continuación qué son los gestores de información. Respecto a la gestión de contenidos, veremos un modelo muy completo, que Jokin Olaizola conoce bien, por haberlo creado él mismo. A continuación hablaremos de herramientas y de integración con otros sistemas de información.

En cuanto a la necesidad de la gestión de contenidos web, cada organización, cada empresa, cada institución tiene una actividad principal o un negocio que llevar adelante. Todo se modela a partir de procesos, en función de la información y sus flujos, muchas veces canalizados a través de los documentos. Hay una necesidad clara de difundir la información, en primer lugar al entorno de la organización; en un segundo nivel a clientes y proveedores; y en un tercer nivel a clientes potenciales, público objetivo o público en general. Respecto a los procesos, tenemos también la necesidad de gestionarlos a nivel interno y externo. Para todo ello, uno de los mecanismos que existe, y que ahora está en boga, es el de los sitios web. Por todo ello, la gestión de contenidos web no es solamente hacer páginas web, es hacer páginas web con un objetivo: difundir una información o llevar adelante una serie de procesos.

1.2. Modelo de gestión

Como modelo completo de gestión de contenidos web, podemos consultar el modelo de presencia en Internet del Gobierno Vasco.

En este modelo hay unas herramientas principales, que serán básicamente un gestor de portales y un gestor de contenidos, que harán su labor en primera línea. Estas herramientas estarán coordinadas con unos ejes de catalogación que permitirán documentar esos contenidos, y que posibilitarán hacer unas búsquedas más fáciles, apoyadas por un buscador, que es también una simple herramienta,



que puede utilizar distintas técnicas (por indexación de palabras o aprovechando previas catalogaciones). Todo ello con un libro de estilo, que nos permitirá unificar la forma de presentación de las cosas, y un modelo organizativo que nos dirá quién hace qué. En algunas organizaciones una sola persona hará todo el trabajo, mientras que en organizaciones grandes y con muchos recursos las distintas tareas estarán especializadas, como en el modelo del Gobierno Vasco en Internet.

Visto desde la parte interna, habrá una gestión de contenidos, gestores que con la herramienta de gestor de contenidos vayan generando contenidos, y habrá gestores de portales, que con una visión más amplia organizarán esos portales. Todo ello para que la ciudadanía pueda acceder a esa información, en el caso del Gobierno Vasco a través de Euskadi.net, accediendo directamente a las páginas o buscando la información.

1.3. Herramientas

La herramienta principal para la gestión de contenidos son los sistemas de gestión de contenidos web, *Web Content Management Systems* (WCMS). Las herramientas han ido evolucionando, desde las muy sencillas que había antes, muy ligadas a la informática, a las herramientas que hay ahora, que permiten hacer cosas muy buenas a usuarios sin experiencia y difundir la información de una manera

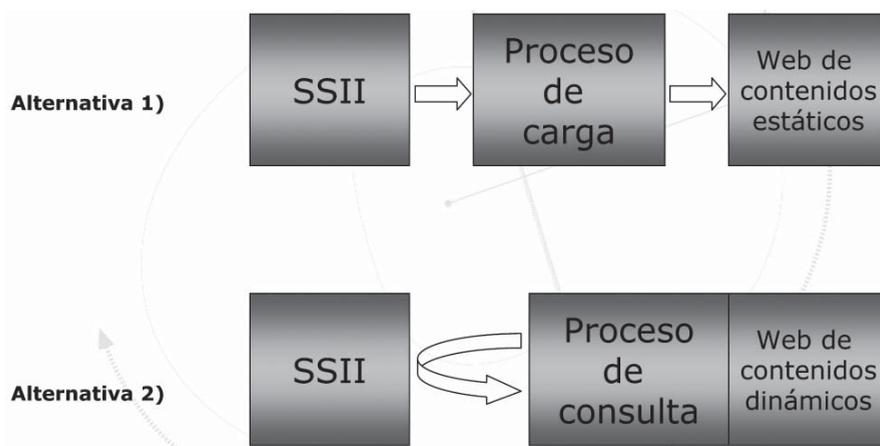
muy sencilla. Tienen una edición muy fácil y no hace falta conocer HTML ni otro tipo de lenguajes. Permiten hacer una gestión distribuida de la edición de los contenidos. Ellos mismos van evolucionando y nos van ofreciendo nuevas herramientas y, además, se van encargando de cumplir ciertas normas que son muy importantes, como las referidas a la accesibilidad.

En el mercado encontramos diferentes alternativas. Hay soluciones empresariales (como Documentum, Interwoven...) que son grandes herramientas sobre las que hay que construir el sistema de gestión de contenidos específico que queremos. Éstas no son cosas que podamos ver. Las que sí se pueden ver, y a las que todos tenemos acceso, son soluciones de software libre, herramientas como Drupal, Jomla, Plone, etc. que permiten a un usuario no informático llevar adelante una web con unas características casi profesionales. Eso sí, es necesario ajustarse a lo que el programa da, y no se puede cambiar; pero lo que da ya es bastante. Hay un sitio web llamado OpenSourceCMS <<http://opensourcecms.com/>> donde están puestas todas estas herramientas de software libre de gestión de webs, para que el usuario las pruebe directamente o bien las descargue en su PC o en su servidor. Es un sitio muy interesante para quien quiera explorar el mundo del software libre de los gestores de contenido. Tiene la virtud de haber conseguido agrupar iniciativas, y quien tiene algo nuevo lo publica aquí, porque es su mejor ventana, su mejor escaparate.

También existen soluciones a medida (como J2EE, ASP.net...), porque no a todo el mundo le resulta válido el producto ya cerrado como es un *open source*, y algunos quieren integrarlo con otros sistemas de información de la organización.

1.4. Integración con otros sistemas de información

Una empresa o una gran organización tendrá seguramente muchos sistemas de información y de gestión funcionando, como por ejemplo la planificación de sus recursos, sus procesos, las relaciones con los clientes... Si es una institución, una administración pública, tendrá ya su propia gestión de expedientes, tendrá incluso herramientas horizontales de uso como son los sistemas de gestión documental que darán soporte a otros sistemas de información. En definitiva, no siempre los contenidos web se generan de cero, hay veces que esa información ya está en los sistemas de gestión, y lo que queremos publicar hacia fuera es sólo una parte. Para estos casos hay básicamente dos alternativas. La primera sería elaborar unos procesos de carga que nos permitan alimentar la web —generalmente una web de contenidos estáticos donde lo que se generan son páginas HTML— a base de una información que en su momento, en los sistemas de información, era dinámica. La segunda alternativa consiste en desarrollarle a la web unos procesos de consulta que vayan contra nuestros sistemas de información, y de ahí recuperar los datos que necesitamos mostrar, y que generalmente será un subconjunto de los que tengamos.



Todo esto en cuanto a la introducción a la gestión de contenidos web desde la perspectiva de una empresa como la nuestra, con los proyectos y necesidades con los que nos encontramos.

2. PAPEL DEL GESTOR DE INFORMACIÓN DIGITAL

2.1. Funciones del gestor de información digital

Es difícil categorizar el papel del gestor de información digital, ya que es una disciplina que no lleva tanto tiempo como otras (donde se puede decir lo que es un administrativo, lo que es un ingeniero de caminos, etc.). La gestión de contenidos web está a caballo de muchas disciplinas y no hay una que la centralice. De hecho, si vamos a buscar o queremos poner ofertas de empleo de esta área en los medios de comunicación, tenemos dificultades para ubicarlas, y a veces están en las categorías de Comunicación, Medios y Periodismo, a veces en Diseño y Creatividad, otras puramente en Informática... No hay un lugar concreto en la clasificación, pero lo que sí hay es una demanda real.

En los sitios pequeños vamos a encontrar al hombre-orquesta que tendrá que hacer de todo, porque no se puede tener mucha gente, ni gente especializada en las distintas cosas. Mientras que en una gran organización, donde las actividades están segmentadas y hay especialización, el gestor de información se dedicará a hacer una o dos cosas concretas.

¿Qué funciones vemos que se están demandando cuando nos piden personas que realicen esta actividad?

Una labor importante es la selección de contenidos. No todos los contenidos de la web son originales, ni tampoco todos las webs muestran contenidos propios,

:: Ofertas de empleo: Requisitos

❖ **Gestor de contenidos web**

- ✓ Se requiere Gestor de contenidos para varios portales web.
- ✓ Sus labores serán:
 - **Actualización de webs y blogs** con contenidos originales
 - **Gestión de las notas de prensa** recibidas de cara a su posible publicación
- ✓ Su objetivo principal será mantener las páginas que se le asignen actualizadas.
- ✓ El perfil ideal requiere una persona que tenga un nivel cultural medio-alto. La capacidad de trabajo en equipo es fundamental.
- ✓ Se valorará positivamente el conocimiento de Inglés.
- ✓ Es necesaria la experiencia previa en gestión de contenidos y/o redacción.

❖ **Redactor@ jefe**

- ✓ Asumirá las funciones de coordinación así como la **edición de noticias** de carácter económico para importante diario digital
- ✓ Requisitos:
 - Experiencia en **medios de internet**
 - Experiencia en dirección de equipos
 - Se valorará conocimiento del idioma inglés

❖ **Gestor de contenidos web**

- ✓ ... bajo la supervisión de los responsables de comunicación y gerencia, se encargará de **elaborar, volcar y actualizar** los contenidos del nuevo sitio web Agenda Viva Digital, a través de un programa de gestión de contenidos. Deberá buscar, manejar, clasificar y seleccionar documentación, directorios y múltiples...

❖ **Gestor de contenidos web**

- ✓ ...bajo la supervisión del coordinador y de la dirección de la Fundación Félix Rodríguez de la Fuente, será el encargado de agilizar la construcción y desarrollo de la plataforma Web del Proyecto de Convergencia del Mundo Rural y la Biodiversidad. Deberá **manejar, clasificar y seleccionar documentación y múltiples enlaces, así como gestionar foros, Blogs,...**

:: Funciones del gestor de información digital

- ❖ **Selección de contenidos. Dossiers.**
- ❖ **Redacción de contenidos**
- ❖ **Actualización de contenidos**
- ❖ **Organización de contenidos**
- ❖ **Administración del portal**
- ❖ **Diseño gráfico**
- ❖ **Catalogación**
- ❖ **Digitalización. OCR.**
- ❖ **Traducción**



son frecuentes las recopilaciones y la selección. Hay mucha información y, a veces, el valor añadido no está en generar cosas nuevas sino en seleccionar lo que interesa a ciertas personas. Por ejemplo, en webs que se dedican a hacer dossiers de prensa, seguimiento de prensa, una de las labores sería la selección de contenidos.

Por supuesto, también estarían las labores de redacción de contenidos y la actualización de esos contenidos; pueden existir diferentes figuras, una que redacta y otra que actualiza, esta última tiene una labor más de edición, más mecánica. Además estaría la organización de contenidos en un nivel superior, puesto que es necesaria una visión más global de lo que va a ser la web y de lo que se quiere comunicar. Asimismo está la administración del portal en sus aspectos más técnicos pero, incluso, en cuanto a organizar quién va a hacer qué cosas, cómo distribuir los contenidos, etc.

Por supuesto está todo el mundo del diseño gráfico, desde el diseño gráfico sencillo (que se ocupa de cómo poner bonito lo que ya existe) hasta la parte creativa, para concebir una web atractiva. En este sentido, no es lo mismo hacer una web de la portada de un periódico, que tiene que ser atractiva pero también muy funcional, que hacer una web de un museo o de otro tipo donde la parte creativa sea más importante.

La catalogación es un elemento importante a la hora de hacer contenidos web; no son contenidos propiamente dichos, pero son los elementos que luego nos van a permitir buscar la información de una manera sencilla. Luego habrá labores más técnicas o más auxiliares, como sería la digitalización, el hacer el OCR —es decir, extraer las palabras de esos textos, que se digitalizan para que se puedan buscar

y no sean meras imágenes digitalizadas—, la traducción, etc. son funciones muy deslavazadas, muy sueltas, porque todavía el mundo de la gestión de contenidos web no está muy regulado.

2.2. Conocimientos del gestor de información digital

Por lo que se refiere a los conocimientos del gestor de información digital, debemos tener presentes una serie de factores importantes: el usuario, Internet, la web, los sitios web, las páginas web y los contenidos web.

Respecto al usuario, el gestor de información digital debe tener en cuenta la accesibilidad y la usabilidad, es decir, el usuario tiene que poder acceder a los contenidos, no sólo físicamente sino también intelectualmente. Si alguien está haciendo una web para difundir información del sida, por ejemplo, y la está haciendo para difundirla a toda la población, el texto tiene que ser accesible a todo el público, tiene que estar escrito en términos sencillos y comprensibles. Hay sitios web muy interesantes donde consultar temas de accesibilidad y usabilidad.

En relación a Internet, es preciso tener claros una serie de conceptos, como por ejemplo saber diferenciar los distintos servicios y no confundirlos.

Acerca de la World Wide Web, hay que comprender lo que es el hipertexto, conocer el consorcio que regula todo este ámbito y que permite la aparición de estándares para que al final todos nos comuniquemos, ya que si nos comunicamos es porque hay unos estándares por debajo.

Respecto a los sitios web, hay que entender que hay distintos tipos de sitios: webs corporativas —que tratan de explicar cuál es el negocio de una organización y pretenden desde informar hasta hacer transacciones—; pasando por los blogs —que tienen un gran calado social, como refleja, por ejemplo, la utilización en publicidad de eslóganes como «la nueva generación blog»—; o los wikis, que también son importantes, con la representación abanderada de la *Wikipedia* como resultado de la colaboración y la interacción. La *Wikipedia* muestra que se pueden generar contenidos de calidad con el esfuerzo de todos, compitiendo con otras organizaciones más tradicionales, como puede ser la *Enciclopedia Británica*, que tienen mayor rigor pero a las que les falta la inmediatez que da el wiki. Por ejemplo, cuando apareció el Sudoku, a los pocos días la *Wikipedia* ya decía lo que era, ponía las reglas, las instrucciones de cómo resolverlos, ejemplos... Habría que ver si la *Enciclopedia Británica* lo ha incluido y, en su caso, cuánto ha tardado en hacerlo. También hay otros tipos de webs como Twitter, etc. aunque casi nos estamos desviando hacia la mensajería instantánea más que a los contenidos web.

Sobre los sitios web también hay que entender temas como el diseño, la arquitectura de información, la creación, el mantenimiento, los distintos niveles o los distintos públicos, si va dirigido a Internet, Intranet o Extranet. Por ejemplo, la Intranet de la empresa ODEI aprovecha la tecnología web para difundir todos los contenidos interesantes de la empresa y agrupar todos los procesos importantes

(desde canales de comunicación a poder rellenar los partes de actividad, controlar los procesos de calidad...).

Sobre las páginas web, lo básico será conocer su edición, mantenimiento, hojas de estilo, lenguajes de marcado... Afortunadamente, hoy día esto no es tan necesario, aunque a veces puede ser útil tener un conocimiento básico de HTML para refinar algún contenido.

2.3. Tipos de contenidos web

Respecto a los tipos de contenidos, tendremos texto, imagen, vídeo, audio, animaciones flash, ficheros en general (presentaciones, ficheros comprimidos...), otro tipo de contenidos como mapas, etc. También está la catalogación, que no es propiamente un contenido web pero está muy relacionada.

2.3.1. Formatos

En cuanto a tipos de contenidos web y sus formatos, es importante la utilización de estándares porque nos van a dar dos grandes beneficios. Uno, que todos nos puedan entender con sus navegadores o con sus dispositivos de lectura. Y otro, que perdure en el tiempo, porque si estamos utilizando estándares abiertos siempre se podrán desarrollar aplicaciones que sigan leyendo esos ficheros. Sin embargo, si utilizamos formatos que pertenecen a empresas muy concretas, si esas empresas desaparecen y nadie mantiene el lector, la información se irá perdiendo.

TIPOS DE CONTENIDOS	FORMATOS PRINCIPALES
Texto	HTML, CSS. Estándares W3C
Imagen	jpg, gif, png, svg
Vídeo	asf, avi, flv, m4v, mov, mp4, mpg, wmv
Audio	aac, mp3, ogg, wma, rm
Flash	swf
Ficheros	Ofimáticos PDF Comprimidos
Catalogación	Etiquetas (indexación social), descriptores, web semántica
Otros: Mapas	

2.3.2. Herramientas de edición

Las herramientas de edición para manejar y editar el texto de los webs van desde un simple bloc de notas hasta editores más avanzados como podría ser Adobe Dreamweaver, un editor profesional de páginas web.

TIPOS DE CONTENIDOS	HERRAMIENTAS DE EDICIÓN PRINCIPALES
Texto	Bloc de notas, Notepad+..., NVU, Aptana, Bluefish, Adobe Dreamweaver
Imagen	Paint ,..., Paint.NET, GIMP, Adobe Photoshop
Vídeo	Camtasia , MS Movie Maker ,..., Adobe Premiere, Pinnacle Studio
Audio	Grabador de sonidos de Windows , Audaci
Flash	SWiSH Max, Adobe Flash
Ficheros	Ofimáticos (Open Office, MS Office) PDF (Adobe Acrobat, PDF Creator,...) Comprimidos (WinZIP, WinRar, 7zip)
Catalogación	Editor de texto (etiquetas <meta>), WCMS

Por lo que se refiere a la imagen, la podemos gestionar desde el Paint del sistema operativo y hacer cosas sencillas, hasta utilizar programas más potentes como el GIMP, que sería el equivalente al PhotoShop en software libre, o el propio Adobe PhotoShop.

En cuanto a vídeo, hay programas que nos permiten capturar vídeo de nuestra pantalla, como sería Camtasia. Se trata de un programa muy cómodo, sobre todo para hacer tutoriales y presentaciones, ya que nos permite grabar todo lo que estamos haciendo en la pantalla, y según vamos navegando y hablando vamos generando un vídeo. Para editar vídeos de una manera sencilla tenemos el Windows MovieMaker así como programas más tradicionales, más clásicos, como el Adobe Premier y otros más profesionales.

En el caso del audio, pasaremos desde la grabadora de sonidos del sistema operativo (un poco limitada) a programas más elaborados, como Audacity, un programa de software libre muy sencillo de manejar y bastante útil, que nos permite editar la información y convertir unos formatos a otros, porque los formatos, el peso y el volumen de los ficheros son temas importantes.

Respecto a Flash, hay unas herramientas ya muy clásicas, muy estándares, para elaborar las animaciones Flash, como el Adobe Flash. En cuanto a ficheros, lo normal es que si son ficheros ofimáticos se manejen con el Open Office software libre, o con su equivalente comercial que es el Office de Microsoft. Para elabo-

rar PDFs tenemos el Adobe Acrobat o distintas herramientas de software libre que también tratan el tema del PDF; lo mismo ocurre con los ficheros comprimidos. En el tema de catalogación, si estamos editando nosotros directamente los webs, utilizaremos un editor de texto y pondremos etiquetas *meta* a nivel interno, o bien directamente el gestor de contenidos web nos permitirá poner *tags*, poner clasificaciones, etc.

2.3.3. Albergue de contenidos web

¿Dónde albergar estos contenidos que vamos generando? Para quien disponga de un gestor de contenidos web en su organización, todo esto le resultará transparente. Para un usuario autónomo, que empieza a generar sus propios contenidos web, hay que señalar que para los webs de texto no hay un sitio específico, pero los proveedores de acceso a Internet directamente dan posibilidades de albergue.

TIPOS DE CONTENIDOS	ALBERGUES DE CONTENIDOS WEB PRINCIPALES
Texto	Proveedor de acceso a Internet (ISP)
Imagen	Flickr, Picasa
Vídeo	YouTube, Google video, Blip.tv
Audio	Internet Archive, Castpost, Podomatic, Poderato
Flash	No se suele usar un albergue específico. El mismo que para el texto.
Ficheros	RapidShare, GigaSize, DepositFiles Presentaciones: slideshare
Catalogación	Donde esté el contenido indexado

Las imágenes las podemos guardar en Flickr y en Picasa. A veces nos encontramos con problemas de espacio en estos albergues de los proveedores de acceso a Internet pero, afortunadamente, podemos dejar los contenidos multimedia en otro sitio y enlazarlos desde nuestro web. Para albergar vídeo tenemos YouTube y Google Vídeo. En audio tenemos otros sitios específicos (Internet Archive, Castpost...); no así para albergar Flash, que se alberga directamente donde lo tengamos. Para los ficheros sí hay sitios específicos, porque el tamaño importa (y si en un momento dado queremos colgar un libro digitalizado que ocupa 80 megas no lo va a soportar cualquier sitio) y tendremos sitios muy útiles que permiten dejar ficheros muy grandes y compartirlos. Por último, la catalogación se albergará donde esté indexada la información.

2.3.4. Repositorios y buscadores

¿Dónde encontrar repositorios, dónde publicar o dónde buscar información de interés respecto a los contenidos web?

TIPOS DE CONTENIDOS	REPOSITORIOS Y BUSCADORES PRINCIPALES
Texto	Google, ..., Hemerotecas digitales , OAI , E-LIS , Google Scholar, ...
Imagen	De pago: iStockphoto , Dreamstime , Fotolia , JupiterImages , PhotosToGO Gratuitos: YotoPhoto (Buscador), La galería de imágenes de Office Online, El archivo fotográfico digitalizado de la revista LIFE, El archivo fotográfico de la NASA, Fotografías de Flickr , Stock.xchng (SXC)
Vídeo	YouTube , Google video , Blip.tv
Audio	PodCastellano
Flash	Búsqueda avanzada de Google, Infografías de la Revista Consumer
Ficheros	Presentaciones: slideshare Búsqueda avanzada de Google
Catalogación	del.icio.us (enlaces favoritos), Flickr (fotos), Technorati (blogs)

En cuanto al texto, el clásico es Google pero también hay sitios especializados, por ejemplo, las hemerotecas digitales, que en el caso de *La Vanguardia* ofrece información desde 1881. Tenemos repositorios de información donde, utilizando estándares de acceso abierto, se pueden dejar textos científicos, textos divulgativos... como por ejemplo, *E-LIS* un repositorio especializado en temas de Documentación.

La imagen tiene gran importancia en la web, antes se ponía una foto sin más pero ahora es un tema que se cuida mucho (para ver su relevancia se puede consultar *El arte de presentar* <<http://elartedepresentar.com/>>). ¿Dónde conseguir fotos? Hay sitios de pago y sitios gratuitos, cuando se quiere algo serio, algo profesional, no resulta caro pagar por una foto comparado con la opción de tener que hacerla personalmente o contratarla. Hay sitios como *iStockphoto* donde directamente se puede lanzar una búsqueda de fotos y suministran un conjunto de imágenes del tema de distintos tamaños, precios... Otros sitios interesantes pueden ser la revista *Life*, que pone sus archivos en la red, la NASA, Flickr donde los usuarios ponen fotografías en formato abierto, etc.

Para conseguir vídeos, está YouTube entre otros. Para Flash no hay un repositorio concreto, la propia revista *Consumer* tiene bastantes infografías sencillas e interesantes que nos dan ideas de lo que se puede hacer con Flash, y cuando se pincha sobre ellas se va viendo la animación.

Por lo que se refiere a los ficheros, en *Slideshare* podemos subir presentaciones o materiales parecidos a presentaciones, ficheros en Open Office o PDF, y compartirlos. Es un sitio muy bueno, donde se pueden encontrar muchísimas presentaciones de calidad (incluso plantean concursos). A las presentaciones se les puede añadir un audio, y convertirlas en *slidecast*. Se puede contactar con mucha gente a través de este tipo de herramientas 2.0, porque no es simplemente dejar los ficheros; dejas el contenido, lo catalogas, vas creando grupos de afinidad, porque cuando se ve mi presentación al lado se ofrecen otras parecidas. Hay una interacción interesante y todo es gratuito.

En cuanto a catalogación, no hay un sitio específico de catalogación, pero apoyado en ella podemos acceder a distintos sitios, a enlaces, a favoritos que tengamos, etc.

3. ODEI Y LA GESTIÓN DE CONTENIDOS WEB

En ODEI nos dedicamos a desarrollar webs tanto estáticas como dinámicas. Como ejemplo de dinámica tenemos la web del *Eustat*, donde el contenido web es más un contenido de datos que un contenido elaborado, y donde se puede acceder a la información de forma dinámica, generando gráficos o mapas a la demanda. Desarrollamos sistemas de información para Internet, Extranet e Intranet —bien sean a medida, WMCS u Opensource—, diseñamos webs, nos ocupamos de la integración web con otros sistemas de información, desarrollamos sistemas de gestión de negocio, servicios de elaboración de contenidos web, servicios de digitalización y OCR, cedemos personas que gestionan contenidos web, incluso todo un centro de documentación... Para terminar, quiero decir que en ODEI estamos abiertos a la colaboración con la Universidad del País Vasco.

XIV

El acceso al conocimiento global: la contribución de Google

Luis Collado

Responsable de Google Book Search, Scholar y News Archive Search para España y Portugal
Espainia eta Portugaleko Google Book Search, Scholar eta News Archive Search-eko arduraduna

Resumen: Internet se ha convertido en muy pocos años en una potentísima e imprescindible herramienta para acceder a niveles de información y contenidos impensables hasta ahora por otros medios. Hemos pasado de la escasez de información al exceso de información y fuentes. Analizaremos cómo Google está incorporando información y contenidos que están universalizando el acceso al Conocimiento y cómo está ayudando a encontrar, descubrir, analizar y categorizar los contenidos accesibles en la Red.

Laburpena: Ezagutza globalerako sarbidea: Googleren ekarpena. Oso urte gutxian Internet tresna ezinbestekoa eta sendoa bihurtu da, orain arte beste bitarteko batzuen bidez lortezinak ziren informazio eta edukietara heltzeko. Lehen, apenas geneukan informaziorik, eta orain, berriz, informazio eta informazio iturri gehiegi ditugu. Ikusiko dugu Google nola gehitzen ari den informazioa eta edukiak, eta nola, hori dela eta, ezagutza ia denon eskura dagoen; halaber, aztertuko dugu Google nola ari den laguntzen sarean dauden edukiak bilatzen, aurkitzen, analizatzen eta kategorizatzen.

0. INTRODUCCIÓN

En esta comunicación vamos a hablar de Internet, en general, y de cómo se realiza el acceso al conocimiento, a los contenidos, a través de Internet y en concreto a través de Google. Intentaremos fijar alguna idea muy básica de lo que son los contenidos digitales, la información y los contenidos en la red. Y terminaremos hablando de las tendencias de futuro.

Internet es algo que está en nuestras vidas; las personas con menos de 20-22 años lo tienen asumido pero, quienes hace tiempo que dejamos atrás esa edad, casi no nos hemos dado cuenta de la gran transformación que ha supuesto en nuestras vidas la aparición de Internet. Hace pocos años la conexión a la Red era una quimera, un trabajo ímprobo; ahora, con la banda ancha, con un cable estamos conectados. Internet no supone un problema de conexión para cualquiera de nosotros; en todo caso lo suponen las redes wifi, porque todavía no están muy desarrolladas, pero no el acceso universal a Internet.

Otro de los grandes ejes que ha desarrollado Internet, como medio de comunicación y de información, es la democratización de las herramientas de pro-

ducción. Hasta hace muy poco tiempo, quien tenía algo que decir al resto de la sociedad, al resto del mundo, tenía que utilizar medios muy tradicionales y muy cerrados a la persona de la calle. Ahora, a través de las nuevas tecnologías, cualquiera de nosotros puede generar información y difundirla, darla a conocer de la manera más sencilla. Con lo cual no sólo se ha extendido la generación de contenidos, sino también su difusión —con los problemas que ello implica y que posteriormente abordaremos—.

Además, la tecnología, es decir, las máquinas, las memorias que se necesitan para almacenar toda la información que está fluyendo por Internet, también se han universalizado, fundamentalmente porque los costes de esos aparatos se han reducido de manera muy drástica. Podemos poner el ejemplo de la Biblioteca del Congreso de los EEUU, llena de libros en papel que ocupan un espacio. Allí en concreto hay 20 millones de libros, que están ocupando miles y miles de metros cuadrados de estanterías. Si todos estos libros estuvieran digitalizados, el espacio para guardar las memorias, los discos donde estarían esos contenidos, no ocuparía más de cuatro o cinco metros cuadrados. Es fácil imaginar lo que supone de concentración de información y también de ahorro de costes, porque el coste de las memorias es mínimo (el mayor coste vendría de emplear horas/persona para el proceso de digitalización). Es fácil comprender el ahorro que supondría la digitalización de esos contenidos y la facilidad para hacer uso de ellos.

1. ACCESO A LOS CONTENIDOS A TRAVÉS DE GOOGLE

Internet es el medio más simple y más «tonto» que hay. En este momento lo están utilizando 1.400 millones de personas en todo el mundo: hay investigadores de muy alto nivel que hacen una utilización muy específica de Internet; hay muchos profesionales utilizando Internet para hacer muchas cosas; y está el ama de casa de mi pueblo que utiliza Internet para consultar la crítica de la película que va a ver esa noche. Internet vale y tiene que valer para todos esos tipos de uso. Por lo tanto, una de las cosas que deberíamos pensar cuando hablamos de Internet es que, por supuesto, puede ser una herramienta de acceso al conocimiento, a todo tipo de contenidos, pero también puede ser una herramienta de lo más sencilla, de lo más simple que puede haber, porque hay mucha gente utilizando Internet para ese tipo de usos.

Internet es algo que usamos todos los días, a todas horas y para todo. Si, además, hablamos de los más jóvenes, ya no sólo es para buscar información, es también para hablar con los amigos, para quedar el fin de semana, para compartir información y fotos con la cuadrilla a través de Facebook, etc. Incluso para hacer cosas que hasta hace muy poco tiempo no se podían hacer más que en otros medios que existían, por ejemplo, ver televisión. Para gente por debajo de los 20 años, el consumo de la televisión tradicional empieza a ser mínimo; cuando quieren ver algo, se conectan directamente a Internet, se bajan el último capítulo de la serie que les gusta, se van a alguno de los portales que permiten ver el programa que les interesa, en el momento en que les interesa y sin publicidad, no como en una cadena de televisión tradicional.

En Internet los usuarios hacen de todo: se comunican, localizan lugares (sitios, salas de cine...), utilizan Internet como un mercado comprando y vendiendo cualquier tipo de cosas; porque también es un mercado, desde el punto de vista de las empresas y de los consumidores; pero, fundamentalmente, lo que hacemos todos es buscar contenidos, buscar información en Internet; y los contenidos en Internet están accesibles a estos 1.400 millones de usuarios. Esta gran cantidad de usuarios se corresponde con una gran diversidad de tipologías, y eso es algo que cuando empezamos a enfrentarnos a Internet tenemos que tener en cuenta, que hay una gran diversidad de usuarios, igual que hay una gran diversidad de contenidos.

En Google siempre utilizamos el mismo ejemplo para hablar de la gran cantidad de contenidos: Internet es un gran paisaje lleno de mucha información, de muchos detalles, de muchas píldoras de información y, a partir de ahí, cada uno de nosotros puede estar dirigiendo la atención a un sitio u otro de ese paisaje; hay muchos detalles, incluso hay mucha información que no se ve —es lo que llamamos la web invisible— y hay que empezar a bucear para encontrar lo que hay debajo.

Pero, Internet no es nada si yo no hago nada; yo, usuario, me estoy enfrentando a Internet, pero Internet no me da más que este paisaje. ¿Quién tiene la decisión, el poder, el control para manejar este volumen de información y obtener aquello que más le interesa? Pues cada uno de nosotros. Cuando digo que Internet es el medio más simple, es que si yo usuario me siento delante de un ordenador con acceso a Internet, lo que encuentro es eso. Y le pido «bueno, Internet, dame algo», pero Internet no responde. Es decir, soy yo, con lo que yo hago, con la decisión que como usuario tomo en cada momento, quien va a extraer las mayores ventajas de lo que hay en la red. Soy yo, usuario, frente a Internet; con mis manos empiezo a buscar y descubrir aquellos contenidos que a mí me interesan. Nadie me lo va a dar, porque hay tanta información en Internet que es absolutamente imposible que alguien pueda gestionar, organizar, orientar toda la información e, incluso, recomendar lo que es bueno y menos bueno. Soy yo, usuario, el que empiezo a buscar y poco a poco a encontrar aquellos detalles que a mí me puedan interesar. Si yo no hago este proceso interactivo con Internet, no encontraré o no conseguiré pasar de ese gran paisaje, lleno de detalles, al punto de información, al contenido específico que a mí, usuario, me interesa.

Porque, para bien y para mal, la información en Internet está absolutamente fragmentada. Ya son incontables los cientos, los miles de millones de páginas web que están editadas en todo el mundo. En Google somos incapaces ni siquiera de indexar todo lo que hay. Nos conformamos con estar indexando entre el 60 y el 70 por ciento de la información que hay en Internet. Hay muchas cosas a las que es imposible acceder. Si hablamos de blogs, que también es un elemento de información donde se están utilizando continuamente herramientas, aportando información a Internet, estamos hablando de que existen más de 150 millones de blogs en todo el mundo. Si hablamos de revistas y de repositorios digitales de información, estamos por encima de los 200.000. Cuando hablamos de revistas o de repositorios hablamos de información de cierto nivel, por lo

tanto, hay un volumen de información tan enorme y tan fragmentado que es absolutamente imposible gestionar.

¿Cuál es la idea que tenemos en Google ante esta situación? Lo que pretendemos hacer de alguna manera es organizar toda esta información, y además, hacerlo de la manera más sencilla para que sea accesible y útil para cualquier persona. Realmente esto es una quimera, es en lo que trabajamos pero es absolutamente imposible de lograr. Cuando decimos «hacerla accesible y útil» estamos pensando, sobre todo, en esa gran masa de 1.400 millones de usuarios que están continuamente buscando, haciendo cosas en Internet. Por lo tanto, quizá a un cinco por ciento de esos usuarios que son los investigadores de alto nivel o a un 15 por ciento, que son los profesionales más destacados, las herramientas de Google le resultarán simples, demasiado sencillas, demasiado básicas. Porque el objetivo que tenemos en Google cuando hablamos de organización se dirige hacia los 1.400 millones de usuarios, y no sólo a la elite de los profesionales o investigadores que puedan utilizar Internet como una fuente de documentación.

Además, en Google estamos trabajando en otra vía. Entendemos que hay mucha información muchos contenidos en Internet pero, realmente, de todos los contenidos que nuestra sociedad ha generado a lo largo de la historia, seguramente en Internet no hay más que el 10% o el 15%. Porque Internet no lleva más de 10-15 años funcionando de una manera generalizada y, por lo tanto, hay muchos contenidos generados antes de la aparición de Internet que no están accesibles a través de la red. En Google tenemos también el objetivo de incorporar toda esa información progresivamente a este nuevo medio que es Internet.

Tenemos que dejar una cosa clara: Google no es más que un medio, somos un buscador, independientemente de que también hagamos otras cosas. Un buscador significa que hay muchos contenidos en Internet, que hay muchos usuarios deseando utilizar esos contenidos y necesitan una puerta de entrada que les ayude a encontrar el detalle específico en ese paisaje al que aludíamos antes. En consecuencia, y aunque algunas veces se nos tacha de ello, Google no tiene entre sus objetivos generar contenidos propios, no somos autores, no somos creadores ni editores de contenidos. No transformamos ni decimos lo que es bueno o menos bueno; únicamente, lo que intentamos es organizar para que los usuarios lleguen a los contenidos que les interesan de la manera más sencilla y rápida; pero no tenemos la autoridad para decir lo que es bueno o menos bueno en Internet. Si hay algún contenido que sale en los primeros resultados de Google y no interesa al usuario que está buscando, pues hay que entenderlo, pero eso no quiere decir que Google no esté haciendo su trabajo de búsqueda y encuentro de información, porque en ningún caso Google está editando la información que hay en Internet.

En Google tenemos un reto: estamos hablando de muchos usuarios y mucho contenido y, sin embargo, la página de entrada al buscador es la página más sencilla, la más «tonta», la más minimalista que hay en Internet. Sólo es el logotipo y una cajita de búsqueda, nada más, porque es lo que los 1.400 millones de usuarios necesitan. No necesitan un gran portal de entrada con mucha información

más o menos organizada, muchos menús etc. Con esta cajita de búsqueda tan simple cualquiera puede llegar a encontrar la información que necesita. ¿Cómo? interactuando con el ordenador. Es algo tan simple —y ésta es una de las razones del éxito de Google— como que yo tecleo una palabra, un par de palabras, y a través de Google empiezo a encontrar resultados, contenidos, que pueden corresponderse a esas palabras que yo estoy buscando. Por supuesto, esto vale para los 1.400 millones de usuarios, y también vale para el investigador del CERN de Europa, para un proyecto del más alto nivel. Nuestra filosofía es que tiene que valer para los 1.400 millones de usuarios, o dicho en palabras de mi madre: «las cosas hay que hacerlas para los tontos y de esa manera nos enteramos los tontos y también se enteran los listos». Esa es la filosofía que tenemos. Para el usuario que tiene más conocimiento y más experiencia, Google también facilita formas que optimizan las búsquedas, por ejemplo, algo tan básico como utilizar las comillas en las búsquedas, o emplear la caja de búsqueda como calculadora, diccionario o convertidor de moneda. Es algo tan simple o intuitivo como eso, preguntar en la caja de búsqueda lo que nos interesa. Cuando empecemos a incluir palabras en la cajita de búsqueda, debemos pensar de la manera más sencilla, más intuitiva. De esa manera encontraremos los mejores resultados, de la manera más rápida.

Porque lo difícil es lo que hay detrás: silos y silos de información de todo tipo. Pero eso ya es un problema de Google, el gestionar, manejar toda esta información para ofrecer a cada usuario los mejores resultados para su búsqueda. Ese es el problema de los ingenieros, de los informáticos, de los técnicos que están trabajando detrás. El usuario que utiliza esta caja tan simple tiene que interactuar con ella de la manera más sencilla y más intuitiva. Por lo tanto, en Internet, la facilidad, la gran ventaja de utilización, es para el usuario; y el problema es para los técnicos que están detrás manejando la situación.

2. EVOLUCIÓN EN LOS CONTENIDOS

Hablando de contenidos específicamente, ha habido una evolución en Internet y en lo que se puede hallar en Google a lo largo de estos diez años y medio de antigüedad que tiene la compañía. Hasta hace muy poco tiempo, Google sólo podía ofrecer resultados que estaban en páginas web, es decir, que otras personas habían creado, introducido en un formato compatible con Internet, en un lenguaje compatible en la web. Pero eso ya es ayer. Ahora a través de Google se puede estar leyendo el periódico de hoy mismo o se puede estar leyendo un periódico editado hace 50 años, se puede estar leyendo libros, se puede estar viajando por el mundo a través de Google Earth sin moverse de delante de la pantalla del ordenador y ver unos detalles en algunos monumentos, piedras, edificios, ciudades, paisajes, etc. que si estuviéramos allí seríamos incapaces de ver con nuestros ojos. También se puede viajar por el firmamento y se puede viajar por debajo de mares y océanos. En definitiva, se pueden hacer cosas que eran impensables hace muy poco tiempo. Por supuesto, se puede buscar cualquier tipo de material audiovisual: series de te-

levisión, ver aquello que todo el mundo comenta en el bar de la facultad y que aparecía en un programa de televisión... si lo buscamos en YouTube seguramente lo encontraremos. Todo eso ya se puede hacer en Internet, y seguramente ya todos lo estaréis haciendo.

¿Lo que se va a poder hacer mañana en Internet? Ni nosotros mismos sabemos qué se va a poder hacer, ni qué va a ser Internet dentro de unos cuantos años. No tenemos la suficiente experiencia ni la tecnología desarrollada para saber lo que será Internet dentro de 15 ó 20 años.

Cuando hablamos de buscar contenidos a través de Google, de buscar conocimientos, tenemos algunas fuentes como el buscador Google Scholar, bastante utilizado en el ámbito universitario. Ahí hay todo tipo de información científica, técnica, académica, buscable y encontrable: acceso a revistas científicas y técnicas, *journals*, cualquier documento de investigación de cualquier equipo de investigación de cualquier universidad del mundo, apuntes, etc. Es un gran repositorio y, como tal, un gran cajón de sastre con mucha información científico-técnica y académica.

Una herramienta que está todavía en pruebas es Knol, se trata de un repositorio que Google ha puesto a disposición de cualquier persona, para que incorpore «píldoras de información» que puedan ser de interés para el resto de la comunidad de usuarios. Si yo estoy buscando, por ejemplo, información sobre mecánica y busco en español, puedo tener ciertas píldoras de información que contienen ese tema que estoy buscando. Es casi como la Wikipedia, pero con una gran diferencia. Todos generalmente utilizamos Wikipedia para encontrar una información, una explicación muy rápida. Uno de los problemas que tiene Wikipedia es que no identifica la fuente que ha generado la información. Esta es la gran diferencia con la herramienta de Knol que Google ha desarrollado, y que da información como Wikipedia —todavía de menor nivel porque está en pruebas—, pero identificando claramente al autor de la información a la que se está accediendo, lo que puede determinar un mayor o un menor prestigio de la información. Además, a diferencia de Wikipedia, tenemos al lado las estadísticas de cada una de las píldoras de información: las vistas, si hay comentario de otros usuarios, cuándo se editó la información, un *rating* hecho por el resto de los usuarios que han accedido a esa información, etc. Al fin y al cabo, es dimensión 2.0 alrededor de píldoras de información, de píldoras de conocimiento, que la Wikipedia tampoco tiene.

Otro acceso a información —a contenido, a conocimiento— a través de Google es YouTube. Es muy conocida, pero seguramente pocos habrán buscado en YouTube clases de psicología en la universidad. YouTube va más allá de ser un mero canal para ver vídeos de entretenimiento que sacan en la televisión. Si yo busco «clases de psicología en la universidad», voy a tener algún vídeo grabado y subido por alguien para ver una clase de psicología universitaria. YouTube también es una herramienta de acceso a conocimiento, lo que ocurre es que habitualmente no se utiliza para acceder a clases magistrales de un profesor, a prácticas, a laboratorios virtuales, etc. pero todo ese contenido también está en YouTube.

Otra herramienta de acceso a contenidos es Google News, se trata de un agregador de noticias, de titulares. Ofrece diferentes posibilidades, por ejemplo, si yo busco «Clark Gable army» — cuando el actor se tuvo que alistar en el Ejército norteamericano en la II Guerra Mundial — puedo buscar en todos los datos, puedo ir a 1943 y puedo estar accediendo a un periódico, a una noticia, tal y como se publicó hace más de 50 años. Hasta ahora, si yo quería ver este periódico, el *St. Petersburg Times*, hubiese tenido que viajar a una biblioteca americana o a la sede del periódico, pedir el microfilm, meterlo en un aparato de visionado y empezar a buscar para encontrar ese ejemplar antiguo del periódico con todo lo que ello supone. Son grandes las ventajas que ahora nos ofrece Internet para estudiar, para investigar, para documentarse. A través de lo que Google está ofreciendo, podemos acceder a prensa histórica, haciendo quizá una mera búsqueda podemos encontrar y ojear el periódico tal cual fue editado, sencillamente porque Google ha trabajado con el periódico o con una biblioteca para digitalizar e incorporar a Internet esos contenidos que hasta ahora no estaban dentro de Internet. Y, a partir de ahí, me puedo mover por el periódico como si estuviera utilizando el lector de microfilm.

Tenemos también el buscador de libros, y si busco «mecánica de fluidos» puedo encontrar toda una serie de libros que me hablan de mecánica de fluidos. Tenemos ahora mismo más de siete millones de libros digitalizados; y yo, sin moverme de delante de la pantalla del ordenador, puedo hacer con un libro lo mismo que haría en la biblioteca o en la librería. En primer lugar, he descubierto el libro sin preguntarle al bibliotecario o al librero, he cogido el libro de la estantería, lo he traído a mi pantalla del ordenador y puedo ojearlo como si estuviera en una biblioteca o en una librería.

Todo eso es lo que se puede hacer con Google. Todo eso es lo que Google está haciendo para incorporar conocimiento generado por nuestra sociedad a lo largo de la historia a este nuevo medio que es Internet. Todo este conocimiento está ya accesible a través de Internet, y está accesible para 1.400 millones de personas que están utilizando Internet. ¿Dónde está el reto para que esos 1.400 millones de personas accedan al tipo de conocimiento que les interesa? Es un reto tecnológico, la clave está en una tecnología que identifique palabras. Nos manejamos en Internet a través de las palabras, y la tecnología tiene que facilitar a esos 1.400 millones de personas que, utilizando el idioma que utilizan habitualmente —y volvemos al carácter intuitivo de Google— puedan acceder a los conocimientos que hay en Internet. ¿De qué manera? ¿Qué facilita la tecnología? La tecnología lo que tiene que hacer es digitalizar los documentos, pero no sólo digitalizarlos, reconocer palabras, es decir, trabajar en OCRs que reconozcan palabras, indexar esas palabras y guardarlas en los servidores junto con los documentos originales. De tal manera que un usuario en cualquier idioma empieza a teclear palabras (sea en Scholar, sea buscar periódicos, sea buscar libros, etc.) y encuentre resultados, contenidos, conocimientos, en las palabras y en el idioma que ha buscado. El idioma no es un problema en Internet; es un problema para los técnicos que están detrás, pero no para los 1.400 millones de usuarios que utilizan Internet.

Es tan sencillo como empezar a usar en Internet el idioma que yo utilizo habitualmente o que necesito para mi labor de investigación.

3. IDEAS BÁSICAS SOBRE LOS CONTENIDOS EN INTERNET.

Lo primero y fundamental es que hay que estar presentes: si queremos crear algo y difundirlo, hay que hacerlo a través de Internet. Con la gran ventaja de que el formato con el cual yo he creado y producido esos contenidos ya no es una barrera, puedo generar contenidos en papel y habrá tecnología que los pase a entorno Internet. Tampoco necesito etiquetar cualquier documento, porque el OCR me identifica no sólo las etiquetas que alguien ha decidido, sino todas y cada una de las palabras integradas en cada una de las páginas de un documento; por lo tanto ya no hace falta etiquetar.

La segunda clave de los contenidos en Internet es que no vale con estar; hay tanto en Internet que lo que hay que hacer es hacerse ver, y esa es una cosa muy importante que no todo el mundo tiene en cuenta cuando vuelca o va a volcar contenidos en Internet. No sólo hay que estar, es que hay que hacerse ver, o de lo contrario el contenido que hemos volcado se pierde en el maremágnum de cosas que hay en la red. Esto es fundamental.

La tercera clave es la usabilidad. No se trata de ser tan minimalistas como la caja de búsqueda de Google, pero las cosas hay que hacerlas de la manera que sean más útiles y más atractivas para los usuarios a los cuales nos dirigimos. Uno de los grandes errores, cuando alguien genera contenidos hacia Internet y los vuelca, es que da la sensación que hay que crear unas páginas web llenas de muchos contenidos, muchos menús, muchas imágenes, flash por un lado, *banner* por otro, etc. La realidad es que eso hace que los usuarios salgan de esas páginas web. Los usuarios buscan simplicidad y utilidad; el exceso de información es un enorme problema para hacer llegar mis contenidos a mis potenciales usuarios en Internet.

Y una última idea muy básica: Internet es el medio más infiel que existe. Lo estamos viendo con la prensa. Cuando comprábamos la prensa en papel, comprábamos el periódico de siempre; con Internet eso se está perdiendo, y lo que importa ya es acceder a la última información, al último titular, a quien lo haya hecho más rápido y más útil. Por lo tanto, ya no hay tanta lealtad a una marca como, por ejemplo, hasta hace muy poco tiempo había en prensa. ¿Cómo lo hacemos en Internet? Entrando en Favoritos, seguramente tenemos la edición digital del periódico que siempre hemos comprado. Pero en favoritos tenemos 15-20 fuentes de información, y cuando queremos leer prensa en Internet directamente estamos leyendo titulares en 15 o 20 ediciones digitales de prensa; porque los usuarios de Internet no se casan con nadie. Ya no vale de nada conseguir que vengan a nuestros contenidos, si no conseguimos hacer que vuelvan un tiempo después a buscar más información, actualizaciones, etc. Eso es algo que muchas veces cuando desarrollamos contenidos hacia Internet olvidamos; y es que nos parece que si nos han encontrado una vez (por tener una página web o

algo en Internet, un repositorio o cualquier cosa), ya estamos absolutamente visibles y la gente va a venir a nuestro sitio sin mayor problema.

4. CÓMO HACER VISIBLES LOS CONTENIDOS

Veamos a continuación algunas preguntas que hay que hacerse siempre que queramos hacer visibles contenidos en Internet.

Hay que tener mucho cuidado en el diseño, en la estructura de los contenidos, intentando en la medida de lo posible eliminar todo tipo de barreras, cortafuegos, registros, *passwords*, etc. En Internet, cualquier tipo de barrera hace que la gran mayoría de los usuarios directamente escapen. Una simple petición de registro, aunque sea de acceso gratuito a un contenido, hará que un gran porcentaje de usuarios directamente se vaya y no vuelva más a nuestro sitio.

También es importante tener muchos vínculos, y que los contenidos que nosotros volquemos a Internet no sólo sean nuestros, sino que sean también una puerta de acceso a muchos más sitios de contenidos. Tener muchos vínculos a otros sitios web favorecerá que nuestros contenidos sean más visibles y más utilizados.

Conviene fomentar mucho la interactividad entre lo volcado a Internet y los usuarios que lo van a utilizar. También hay que tener en cuenta la velocidad de carga de la página. Es necesario cuidar la relevancia de los contenidos, ver qué puede estar interesando a mis usuarios, qué necesidades pueden tener, para así volcar la información que realmente les interesa.

Todo lo que acabamos de mencionar es básico, sin embargo, muchos desarrolladores de contenidos por Internet olvidan gran parte de estos temas cuando desarrollan contenidos hacia el nuevo medio.

5. LOS E-BOOKS

Y, hablando de futuro, debemos tratar un tema muy actual, que está muy en boga, y que son los *e-books*, otra forma de acceso al conocimiento, a la información. ¿De qué hablamos cuando decimos *e-book*? ¿Estamos hablando de un aparato de lectura —éste que se empieza a ver en muchos sitios—? Aparatos donde yo puedo llevar un volumen de información, de documentación, impensable hasta ahora; aparatos absolutamente portátiles en cuya pantalla se puede leer muy bien, con una batería de larga duración, etc. Este tipo de dispositivos nos facilitan el acceso a contenidos digitalizados sin necesidad de ordenador, ni siquiera portátil.

¿Un *e-book* es un archivo PDF? Cuando se habla de un *e-book* —es decir, de contenidos digitalizados y accesibles sea en formato libro o cualquier otro formato— muchos usuarios piensan que con generar PDF de un documento que estaba en papel y con ponerlo ya es suficiente; pero a lo mejor hay muchas más cosas que se pueden hacer. ¿Un *e-book* es también un programa, es un fichero, es un XML, un HTML en cualquiera de sus formatos que pueden dar acceso a la infor-

mación? Hay mucha gente que desarrolla contenidos de forma electrónica de esta manera, no PDF-s directamente, sino que crea un programa que tiene los contenidos, y además tiene muchas más cosas que se pueden hacer con esos contenidos.

Lo que está cada vez más claro es que cuando hablamos de contenidos electrónicos, ya sean libros, revistas, etc. tenemos que separar algo que hasta ahora iba unido; cuando hablábamos de libros o revistas, hablábamos siempre del papel. A partir de ahora, hay que separar lo que es contenido-texto de los formatos en los que está, sean aparatos, sea papel, sean ordenadores, etc.

6. TENDENCIAS DE FUTURO

El futuro vendrá por aquí, ya no va a ser imprescindible el ordenador para acceder a la información digital, ya que hay dispositivos de lectura —y habrá cada vez más— con grandísima capacidad, legibilidad de pantalla muy buena, color y diversidad de contenidos, todo ello en un aparatito como éstos que ya estamos viendo. Pero además, a través de los móviles de última generación (como un iPhone o cualquier modelo basado en el sistema Android) es posible acceder a libros, a repositorios digitales de información, llevarlos en el móvil y consultarlos a través del móvil sin mayor problema. Por lo tanto, ya no necesitaremos de un ordenador de sobremesa o portátil para acceder a los contenidos.

La idea en la que está trabajando mucha gente —por ejemplo Google— es que el problema sea técnico, pero la ventaja sea para los usuarios. La fórmula en la que estamos trabajando, y a la que llamamos «acceso *online*», consiste en que cualquier píldora de información, cualquier contenido digitalizado, cualquier conocimiento que ya está en Internet, sea accesible a través de cualquier dispositivo con conexión a Internet (por lo tanto, no necesito memoria, capacidad de memoria); desde cualquier lugar, en cualquier momento y para cualquier tipo de contenido, independientemente del aparato o del dispositivo que en cada momento esté utilizando. Por ejemplo, en la oficina trabajaré con el portátil con acceso a Internet, en el autobús o en el tren consultaré contenidos a través del móvil y, de vuelta a casa, podré seguir consultando esos contenidos a través de la PlayStation, con conexión wifi.

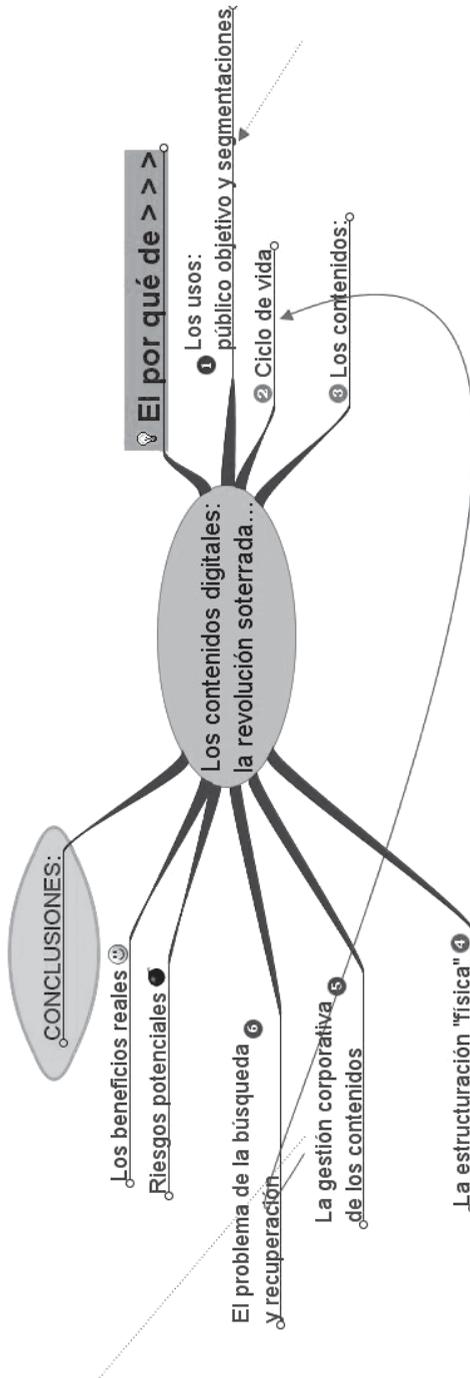
En Google creemos que el futuro del acceso a los contenidos digitalizados va por ahí: la no dependencia de un dispositivo sino el acceso multidispositivo, en cualquier momento y desde cualquier lugar. Porque entendemos que la clave va a ser ésta: cualquiera de nosotros, cuando quiera y donde quiera acceder a información, lo va a poder hacer, desde cualquier parte y desde cualquier tipo de dispositivo.

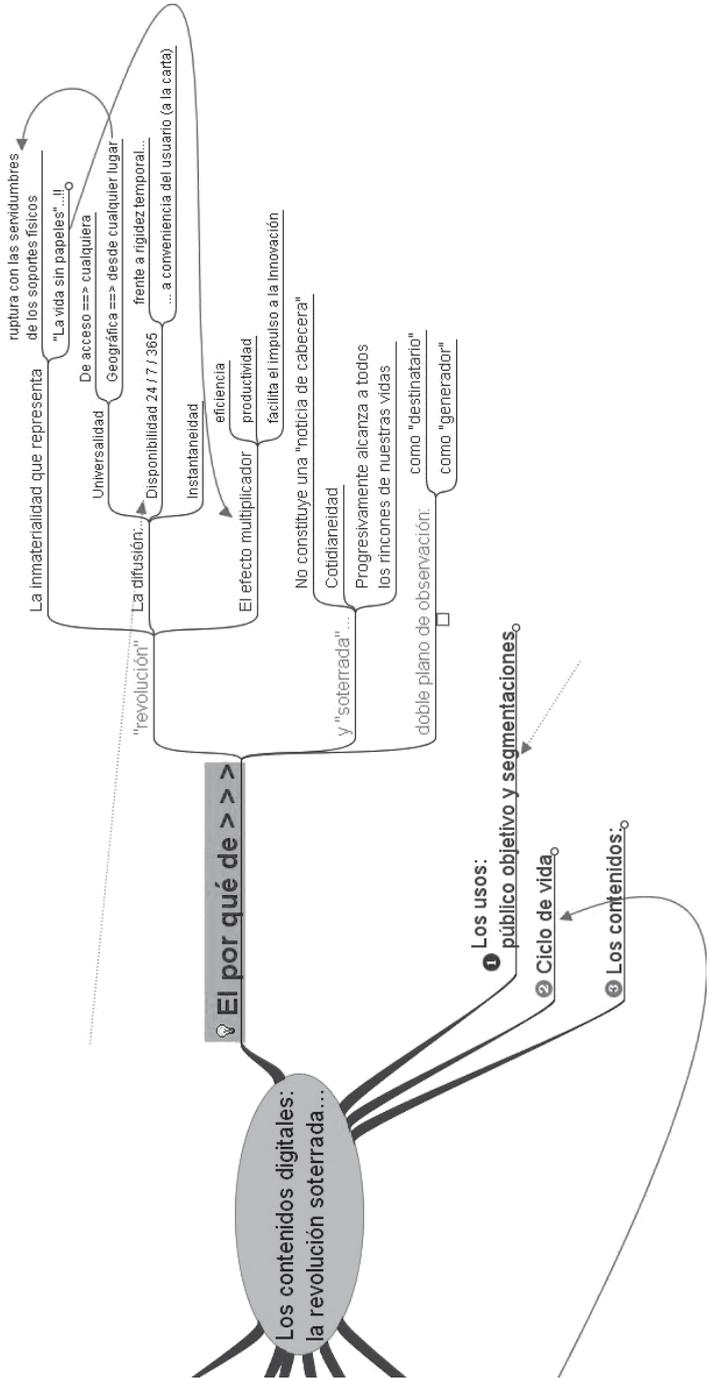
Apéndice

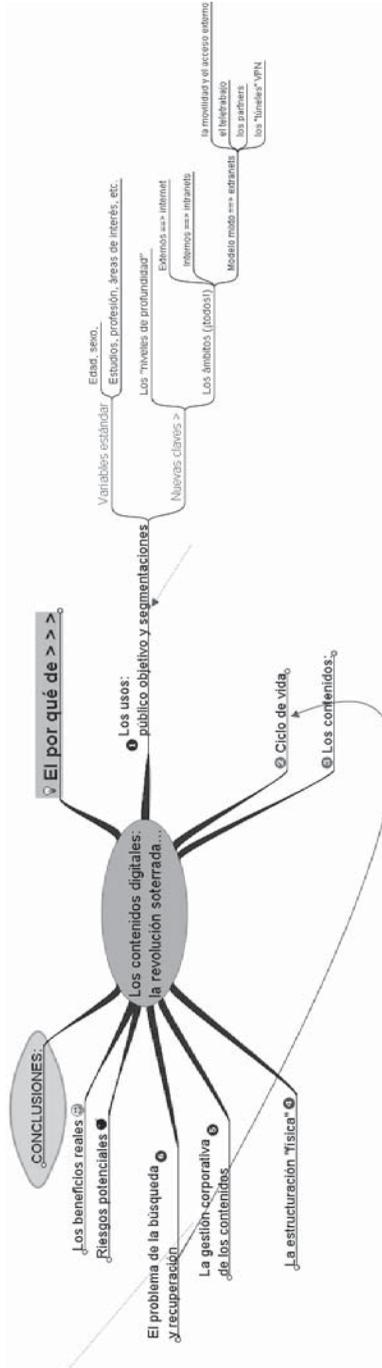
Los contenidos digitales: la revolución soterrada... (esquemas)

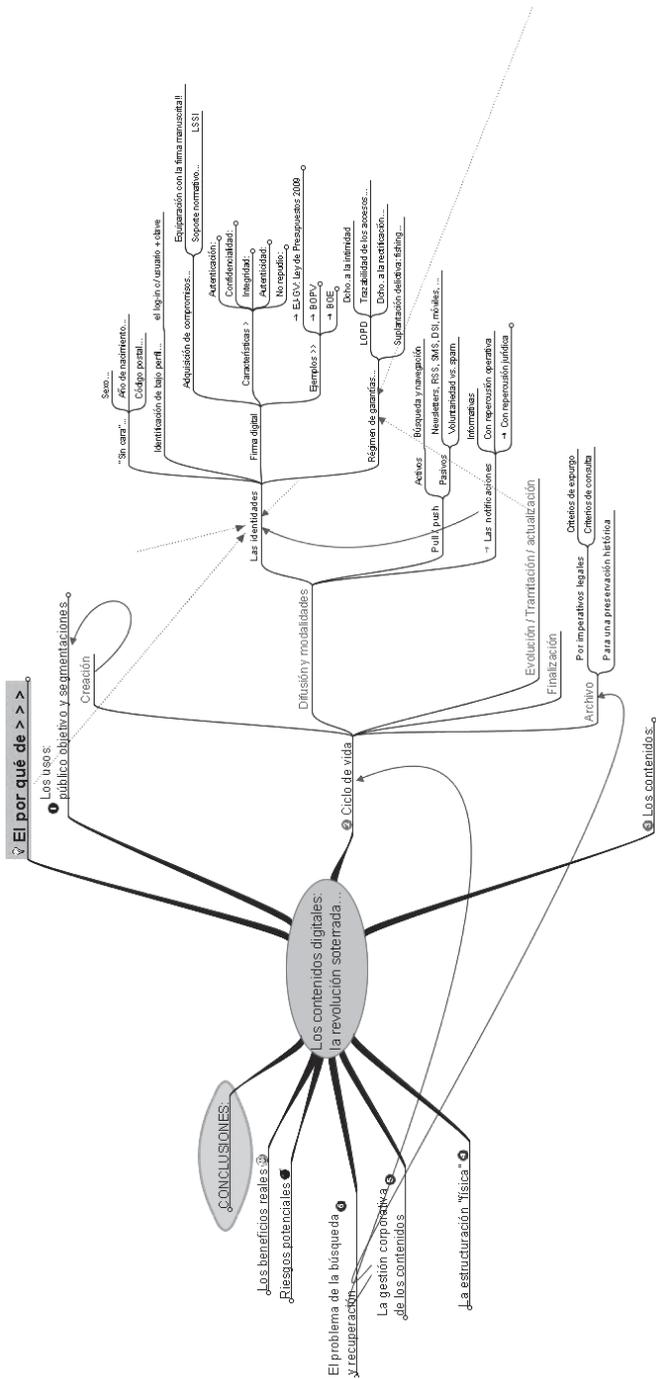
Jokin Olaizola Azkarate

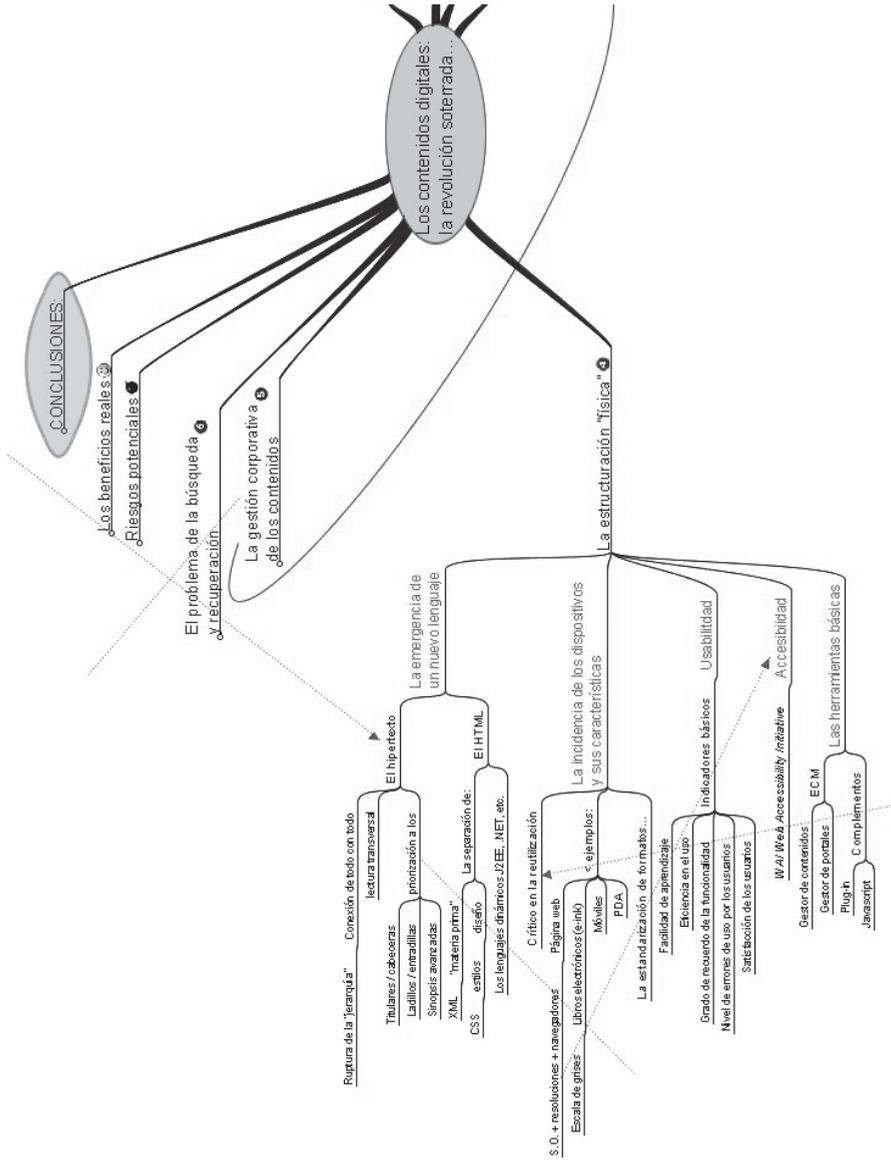
Gestor Web en el Gobierno Vasco
Web kudeatzailea Eusko Jaurlaritzan

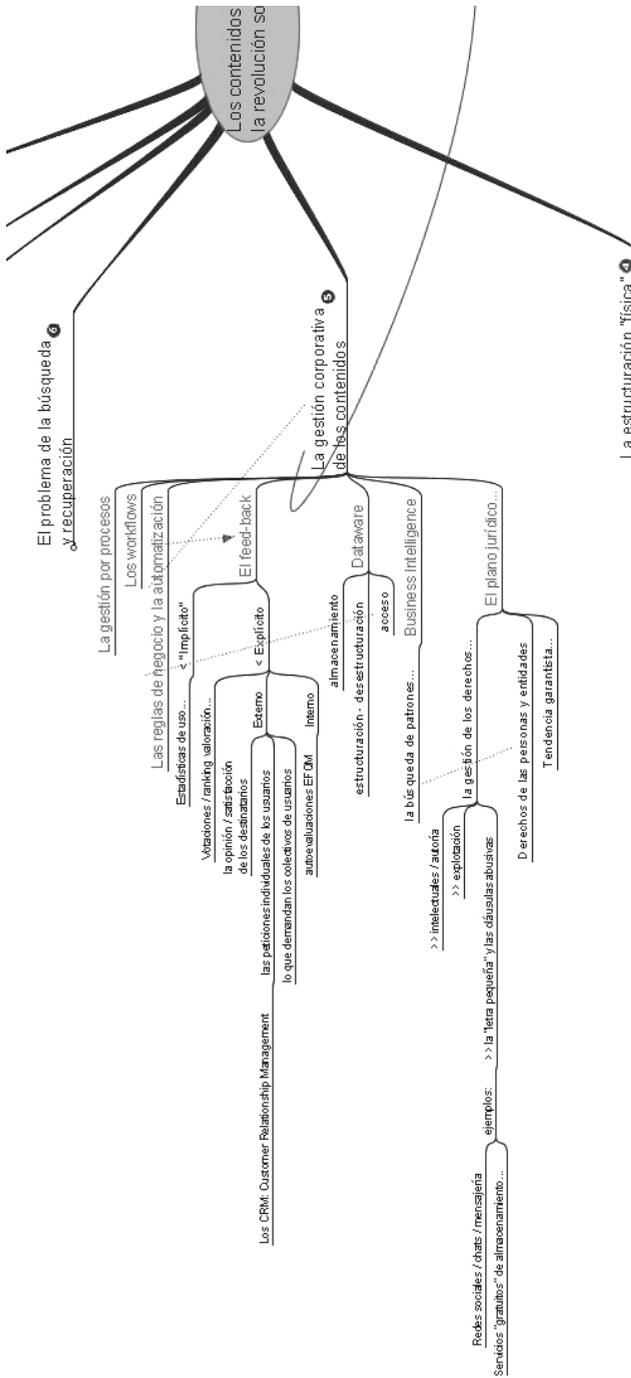


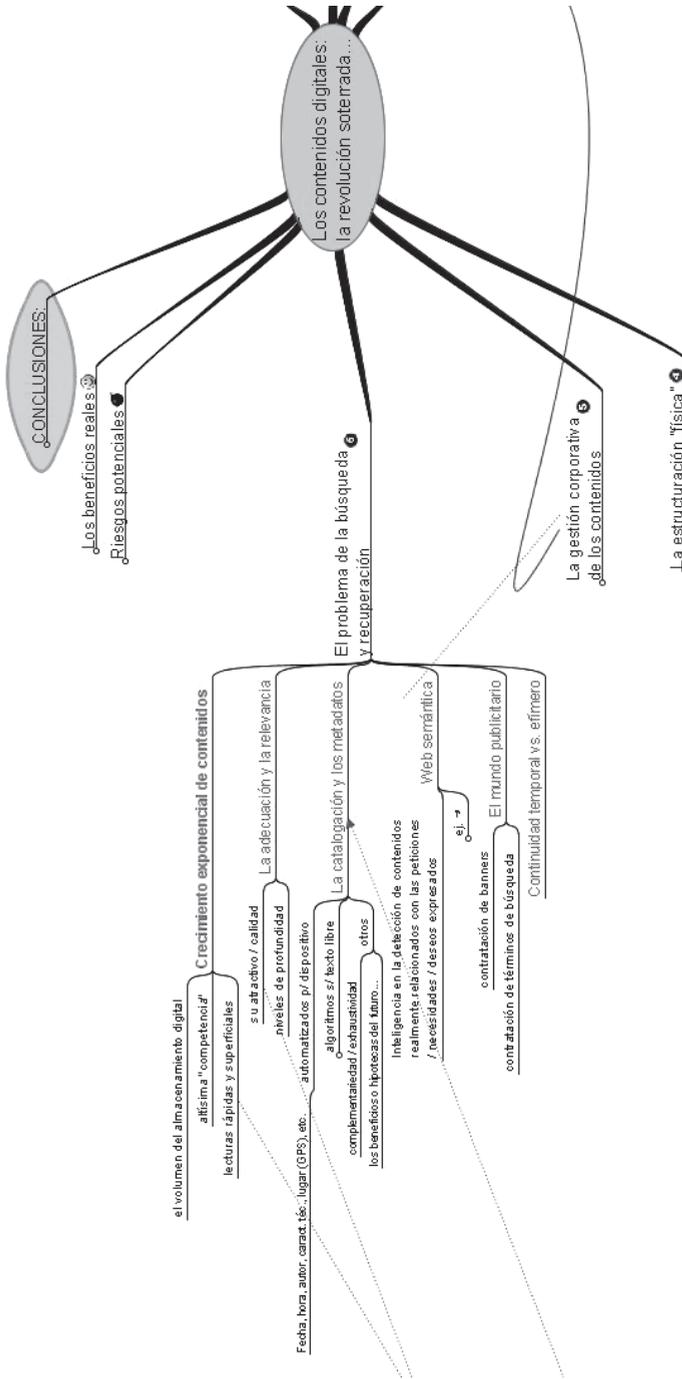


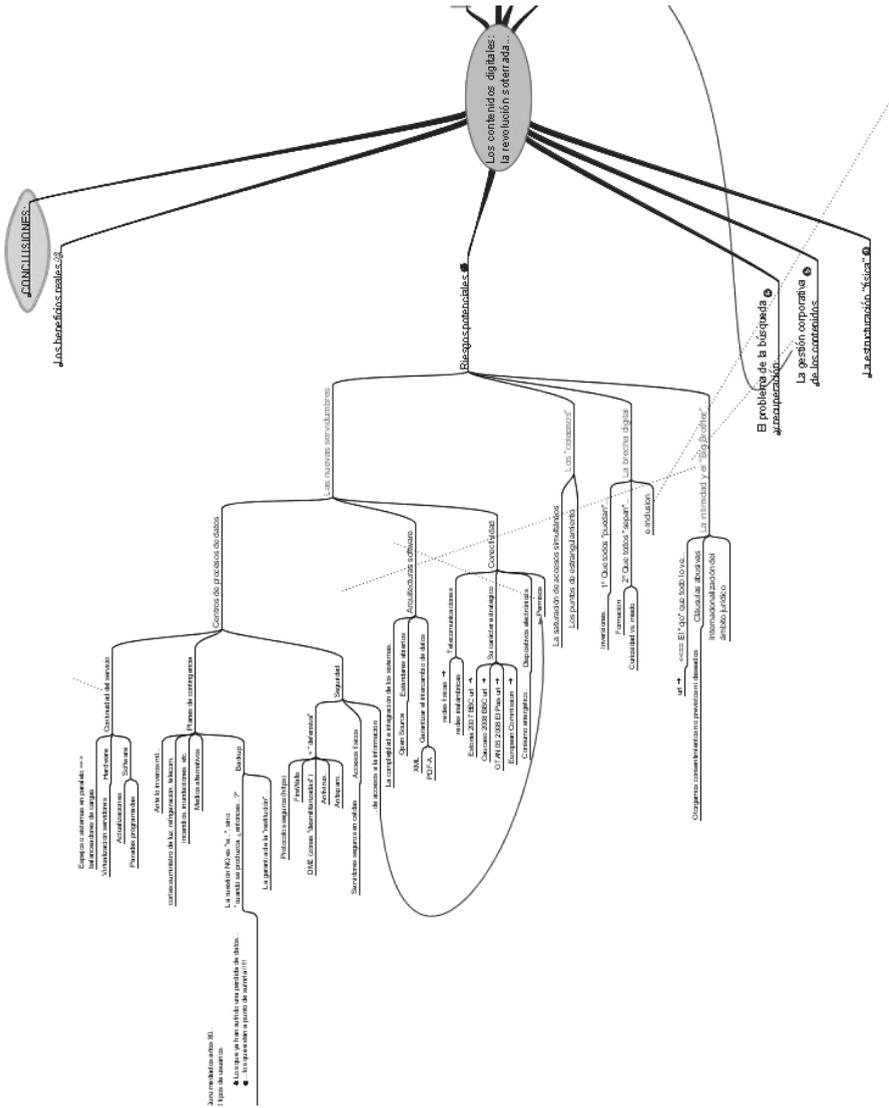


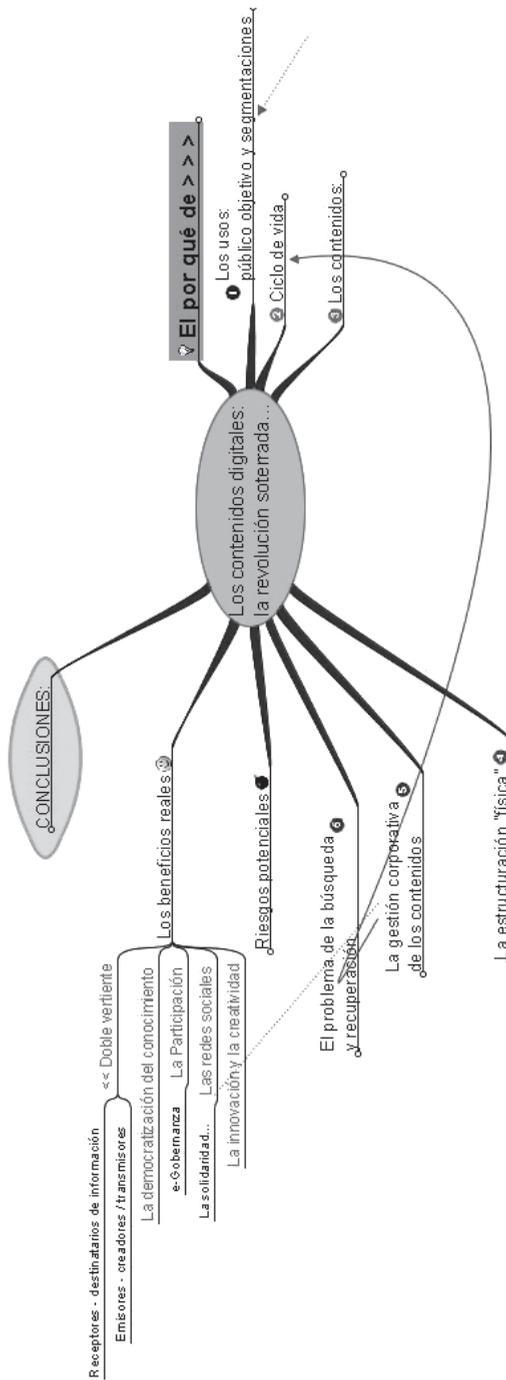


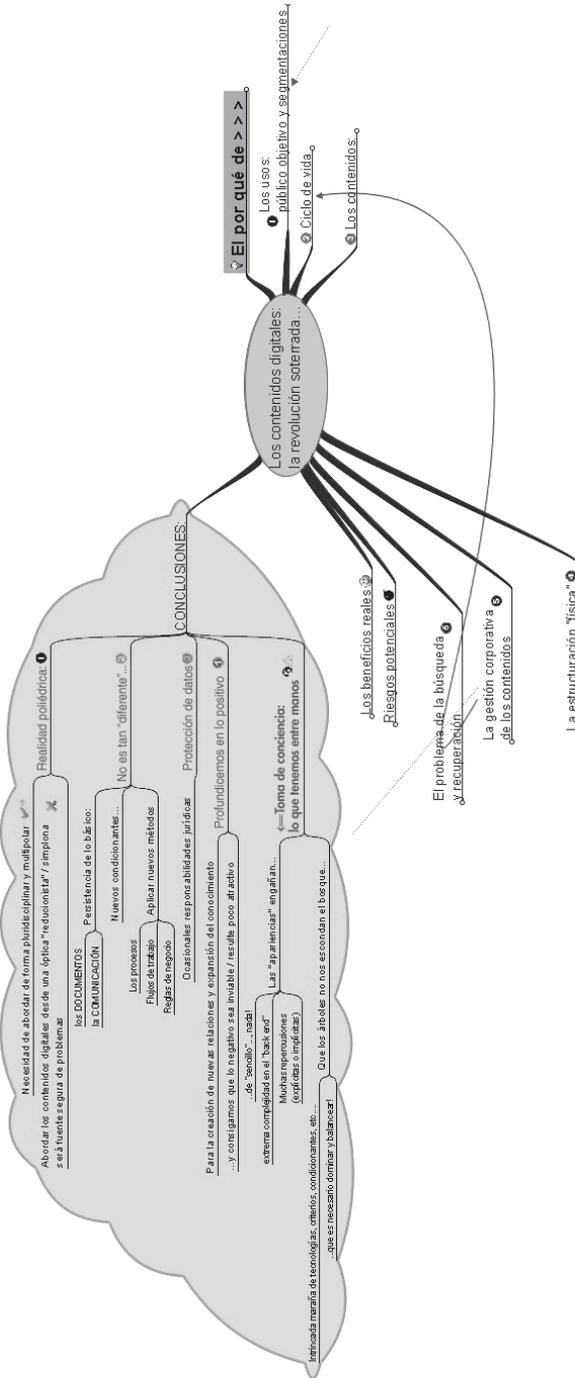














EUSKO JAURLARITZA
GOBIERNO VASCO

HEZKUNTZA, UNIBERTSITATE
ETA IKERKETA SAILA
DEPARTAMENTO DE EDUCACIÓN,
UNIVERSIDADES E INVESTIGACIÓN

berria

Diario
Noticias de Alava
Eres lo que lees

Odei
Servicios de Tecnologías
de la Información

Q
2 0 0 9

