
Introdución ás Bibliotecas Dixitais

Nieves R. Brisaboa / Carme Fernández Pérez-Sanjulián
brisaboa@udc.es / carme@udc.es
Universidade da Coruña

1. Introdución

Se revisarmos o panorama das tecnoloxías da información e das telecomunicacións durante os últimos anos, resulta evidente que a velocidade do desenvolvemento destas disciplinas ten medrado exponencialmente ao tempo que os prezos do material asociado a tales tecnoloxías (*chips*) se rebaixaban na mesma medida. Este feito, a sorprendente suba de prestacións e calidade ligada ao espectacular descenso de prezos, non ten precedentes na historia tecnolóxica da humanidade e, aínda que para todo o mundo se ten feito evidente a entrada masiva dos computadores na vida ordinaria, probabelmente só os expertos se decatan de até que punto este proceso é abraiante. A modo de exemplo, pódese lembrar que o computador que levou a primeira persoa á lúa tiña 8 kilobytes de memoria (non interesa detérmolos aquí a precisar que ou canto é un kilobyte de memoria, abonda dicir que o computador con que se está a escribir este traballo ten 256.000 kilobytes, 32.000 veces máis), mais, ao mesmo tempo, eses 8 kilobytes tiveron un custo de millóns de dólares, mentres que, hoxe, con 40.000 pesetas, pódense comprar os 256.000 kilobytes deste computador. Se a tecnoloxía do automóbel tivese experimentado esa mesma progresión, agora conduciríamos vehículos capaces de desenvolveren unha velocidade de 1.600.000 Km/hora e o prezo do coche sería inferior ao dun refresco.

Semella pouco ortodoxo comezar un artigo científico cuns exemplos deste tipo, mais é difícil explicar doutro xeito a persoas alleas ao mundo da informática até que punto o desenvolvemento destas tecnoloxías é un feito verdadeiramente sorprendente e inusitado.

O abaratamento dos prezos e o incremento en potencia dos computadores na década dos 80 fixo deles unha ferramenta que, pola súa versatilidade, se converteu en imprescindible para todo tipo de empresas e profesionais, mais a aparición e xeneralización da Internet na década dos 90 ofreceu todo un mundo de novas posibilidades ao pór a traballar conxuntamente as tecnoloxías informáticas coas das telecomunicacións. Deste xeito, comezouse a utilizar o termo TIC (Tecnoloxías da

Información e das Comunicacións) para denominar a conxunción destas dúas tecnoloxías.

Internet proporcionou unhas posibilidades de publicación de contidos de todo tipo (desde documentos a música, imaxes ou vídeos) a un público ilimitado por un prezo ridículo. Aínda hoxe as posibilidades da Internet están a ser apenas albiscadas, mais, con certeza, Internet vai producir cambios radicais no noso modo de acceder a todo tipo de informacións e, xa que logo, no xeito de nos relacionar, traballar, estudar, divertir, consumir, etc. A acumulación de tantas mudanzas vai traducir nunha modificación dos hábitos que rexen tanto a nosa vida como os da sociedade en que vivimos, con influencia só comparábel á que tiveron no seu día a introdución da radio, primeiro, e da televisión, despois, ou a popularización do automóbel.

É esa capacidade que proporciona a Internet de publicar contidos a que abriu a porta ao xurdimento e proliferación das Bibliotecas Dixitais. Mercé á Internet é posíbel que obras valiosas e antigas se fagan accesíbeis para o público en xeral, permitindo dar ao rico patrimonio documental unha difusión até agora impensábel. Do mesmo modo, Internet pode facer posíbel que no último recanto da nosa xeografía, simplemente con ter luz e teléfono, se teña acceso a todos os catálogos bibliográficos do mundo que estean dixitalizados e conectados á Rede e, mesmo, ao contido das obras das Bibliotecas Dixitais. Tamén, grazas á tecnoloxía das Bibliotecas Dixitais, é hoxe posíbel ler a última edición de calquera xornal, de *A Nosa Terra* por exemplo, desde calquera lugar do planeta.

Resulta evidente que a hipótese de publicar obras de literatura, obras científicas, xornais, edicións facsimilares de documentos antigos (fotografados dixitalmente páxina a páxina), etc. abre un mundo de posibilidades para a divulgación de tales documentos e informacións; porén, a pesar dos esforzos que se teñen feito por parte de institucións públicas e privadas, moitas destas accións pasan desapercibidas por non contaren co asesoramento tecnolóxico axeitado. Do mesmo xeito, iniciativas ben fundamentadas tecnoloxicamente tampouco resultan satisfactorias por seren levadas adiante con pouco coñecemento das necesidades, intereses e competencia dos investigadores ou do público a que ían dirixidas. Estamos perante un campo en que cómpre que as equipas de traballo sexan multidisciplinares, contando con especialistas procedentes tanto das áreas tecnolóxicas como das Humanidades.

Con este traballo téntase, de entrada, expor os conceptos chave sobre Bibliotecas Dixitais e, despois, describir algunhas das ferramentas informáticas que subxacen nesta tecnoloxía. Esbózanse, así mesmo, algúns dos retos de investigación en que se están a realizar esforzos.

2. Bibliotecas Dixitais

2.1. Algúns conceptos previos

É importante diferenciar os conceptos de Bibliotecas automatizadas ou informatizadas, Bibliotecas Dixitais e Bibliotecas Virtuais.

O primeiro concepto, o de Bibliotecas automatizadas ou informatizadas, refírese a bibliotecas tradicionais onde as tarefas de catalogación, elaboración de fichas, control de empréstitos etc. se realizan co apoio dun programa informático. Nestes casos a informática é unicamente unha ferramenta que facilita as tarefas habituais do persoal da biblioteca. Alén diso, se existir un catálogo informático, este ha facilitar a procura de obras ou documentos e se, ademais, ese catálogo fose accesíbel vía Internet, unha persoa podería desde a súa casa ver se unha obra en concreto está na biblioteca, se está emprestada ou dispoñíbel, etc. A Biblioteca da Universidade da Coruña, como a de moitas outras universidades actuais, ten as facilidades antes citadas, mais isto non é suficiente para afirmar que se trata dunha Biblioteca Dixital .

Unha Biblioteca Dixital é aquela en que os seus fondos están dixitalizados (no apartado 4.1.1 aclárase este termo). Os seus fondos están, por tanto, almacenados en soportes informáticos.

Por suposto, é perfectamente posíbel pensar en bibliotecas bimodais, isto é, bibliotecas en que os fondos físicos (en formatos tradicionais: libros, xornais, revistas...) vaian sendo paulatinamente dixitalizados para garantir, dunha banda, un máis doado acceso e, doutra, a súa conservación.

En todo caso, unha Biblioteca Dixital debe ofrecer aos seus usuarios algún meio informático de acceso aos seus fondos, ben sexa para a súa lectura en pantalla ou ben para obter unha copia dos mesmos baixándoos vía Internet. Obsérvase que o concepto de Biblioteca Dixital non implica de ningún modo que o acceso aos seus fondos sexa vía Internet: de feito, é perfectamente pensábel que só se poida acceder aos mesmos desde computadores situados na propia biblioteca e que a base de datos que conteña os fondos dixitalizados non sexa accesíbel desde ningún outro lado. Agora ben, calquera Biblioteca Dixital pode ter unha interface de acceso vía web.

Unha Biblioteca Virtual é aquela Biblioteca Dixital que ofrece, como meio de acceso aos seus fondos dixitais, unha interface de usuario sofisticada que simula unha biblioteca real no seu aspecto e funcionamento. Así, ao acceder a unha Biblioteca Virtual vese representada unha biblioteca cos seus caixóns de arquivo, que se poden abrir para facer procuras, os seus andeis cos libros, que se poden coller, ou taboleiros onde é posíbel ler e escribir anuncios e noticias

(<http://rosalia.dc.fi.udc.es/cicyt/>). Ademais, tales bibliotecas ofertan servizos que, en xeral, teñen a ver coa explotación do ciberespazo e das posibilidades que fornece Internet (*chats, e-mail*, servizos de noticias, etc.).

Actualmente a rede física que soporta a Internet ten moitos tramos que, por estaren soportados no vello cableado da telefonía, non permiten grandes velocidades de transmisión de dados (cando menos, desde os domicilios particulares até o provedor da Internet). Como consecuencia da baixa velocidade da Rede, interfaces tan sofisticadas como as requeridas en contornos de realidade virtual son pouco axeitadas para a Internet debido a que tales interfaces requiren grande cantidade de información gráfica, e mesmo sonora, que tarda en viaxar pola Rede. Polo tanto, á parte do custo de programar as complexas interfaces de realidade virtual necesarias para unha Biblioteca Virtual, as limitacións da Internet desaconsellan tamén ese tipo de interfaces, de aí que a maior parte das bibliotecas accesíbeis vía Internet sexan Bibliotecas Dixitais e non Bibliotecas Virtuais propiamente ditas.

En todo caso, a diferenza entre unha Biblioteca Dixital e unha Biblioteca Virtual radica en que a súa interface de acceso ao público utilice máis ou menos medios propios de realidade virtual (isto é, que tente, en maior ou menor medida, recrear un contorno virtual de biblioteca). Como é evidente que este parámetro é moi subxectivo, na literatura técnica véñense usando con frecuencia os termos de Biblioteca Dixital e Biblioteca Virtual como sinónimos. Con estas definicións a *Biblioteca Virtual Miguel de Cervantes* ([http:// www.cervantesvirtual.com](http://www.cervantesvirtual.com)) tería sido, nas súas orixes, máis unha Biblioteca Dixital con acceso vía Internet que unha Biblioteca Virtual; porén, os últimos servizos implementados para usuarios (taboleiros de anuncios, foros de debate, etc), case a aproximarían á definición de Biblioteca Virtual.

Ao longo dese traballo referirémonos ás Bibliotecas Dixitais como termo xenérico, pois o feito de posuíren unha interface de usuario vía web que reconstrúa, con máis ou menos suceso, o aspecto e funcionamento dunha biblioteca real a modo da realidade virtual non cría ningunha diferenza nos aspectos tratados neste traballo. Doutra banda non hai un acordo explícito na literatura técnica sobre as definicións aquí propostas sendo, de feito, moi frecuente que os termos de Biblioteca Dixital e Biblioteca Virtual sexan usados de xeito indistinto.

Entre as Bibliotecas Dixitais operativas entre nós poderíase sinalar a existente na web do Centro Superior Bibliográfico de Galicia (CSBG) que permite consultar e ver en liña tanto a colección de prensa antiga como a cartografía de Galiza (<http://biblioteca.cesga.es>). No que atinxe á literatura galega ou en galego, cómpre citar os esforzos desenvolvidos, desde o grupo galego.21.org polo Proxecto Tomiño (<http://www.galego.21.org/tominho/>) que tenta trasladar, de xeito gratuíto, os textos da literatura galega en formato dixital á Rede e tamén os traballos realizados pola Asociación Galega de Tradutores desde a súa Biblioteca Virtual (<http://www.BiVir.com>).

2.2. Consideracións sociolóxicas

As bibliotecas teñen sido tradicionalmente gardadoras, depositarias e conservadoras dos documentos dunha determinada comunidade. No cumprimento desta función, as bibliotecas serviron aos intereses económicos, ideolóxicos e políticos das comunidades en que naceron, que as financiaron e que, por tanto, tamén as controlaron. Só unha visión inxenua do papel das bibliotecas podería situalas fóra das forzas e intereses socio-económicos das sociedades que as promoveron.

No caso de bibliotecas privadas, por exemplo bibliotecas de empresas ou de corporacións, é a política das propias sociedades a que decide que documentos preservar e a quen darlle acceso aos mesmos. No caso de bibliotecas públicas, as políticas de investimentos (que fondos mercar e cantos exemplares de cada obra), de empréstitos (a quen emprestar e por canto tempo), de horarios de apertura, etc. teñen condicionado o impacto dos diferentes documentos gardados nas bibliotecas, mediatizando ou desenvolvendo, segundo o caso, o seu papel como depositarias, promotoras e divulgadoras de cultura. En resumo, tanto as bibliotecas públicas como as nadas en ámbitos completamente privados e fechados, estiveron sempre suxeitas a intereses, motivacións, crenzas e prexuízos, elementos todos eles que teñen dirixido os seus investimentos e políticas de acceso.

Evidentemente, as Bibliotecas Dixitais están xa e estarán no futuro baixo o dominio das entidades que as promovan, do mesmo xeito que sempre o estiveron as bibliotecas tradicionais. Con todo, nas Bibliotecas Dixitais os mecanismos de control terán de ser outros, ou o que é o mesmo, as Bibliotecas Dixitais abren novos retos e posibilidades de influencia neste marco socio-político que cómpre non esquecer.

En primeiro lugar, serán os poderes públicos, no caso de Bibliotecas Dixitais públicas, os que determinen quen poderá acceder aos fondos, que coleccións se deberán dixitalizar e de onde van saír os recursos que financien tal dixitalización. As decisións sobre que dixitalizar poden ser tan transcendentales para un futuro non moi distante, como o foron no seu momento as opcións tomadas a respecto de que obras mercar e cales non, ou que documentos catalogar e facer accesíbeis ao público e cales abandonar á súa sorte nos armacéns.

Por outro lado, as bibliotecas sempre serviron a un ou máis grupos sociais aglutinados (necesariamente) arredor do emprazamento físico da biblioteca, con referentes culturais e intereses relativamente comúns. A limitación imposta pola distancia xeográfica resulta evidente se observarmos a localización das Bibliotecas Nacionais, sempre situadas na capital ou en lugares hipoteticamente centrais, de cara a favorecer, na teoría, o acceso ás persoas usuarias. Para alén disto, a cantidade e calidade dos fondos dunha biblioteca estaban tamén limitados polo tamaño, importancia, riqueza e intereses da comunidade á cal servían. Estas limitacións vense radicalmente modificadas polas Bibliotecas Dixitais.

As Bibliotecas Dixitais, ao posibilitaren o acceso aos seus fondos vía Internet, poden servir a comunidades moito máis amplas. Ábrese a posibilidade de que comunidades con intereses especiais, mais con membros distantes xeograficamente, poidan ter Bibliotecas Dixitais que sirvan aos seus intereses (por exemplo, unha ampla Biblioteca Dixital de textos lidos para invidentes será factíbel grazas a que todas as persoas cegas da mesma comunidade lingüística poden ser usuarios potenciais desa Biblioteca Dixital e facer, por tanto, economicamente interesante a súa creación). Doutra banda, bibliotecas con fondos xerais para unha ampla comunidade de lectores (pensemos nunha biblioteca pública dunha grande cidade) poden repensar a súa política de investimentos, mercando menos exemplares das obras e utilizando os fondos así resgatados en dixitalización, co fin de serviren a comunidades máis amplas de potenciais usuarios web.

Con todo, os condicionantes sociopolíticos do desenvolvemento das Bibliotecas Dixitais non só se van manifestar nas políticas que implementen as persoas encargadas de tales bibliotecas ou os poderes públicos en xeral, senón que tamén van ser exercidas a través do propio desenvolvemento tecnolóxico. Actualmente están a xurdir diferentes solucións tecnolóxicas para a creación, alimentación, xestión e mantemento de bases de datos documentais e interfaces de acceso ás mesmas. É probábel que, co tempo, certas solucións gañen forza ou mesmo se estandaricen, mentres que outras vaian ficando esquecidas.

De todos xeitos, ao mesmo tempo que existe unha tendencia centrípeta nos procesos de desenvolvemento tecnolóxico de cara a lograr a estandarización, tamén é certo que nestes mesmos procesos se producen tendencias centrífugas que abren portas a variadas solucións e manifestacións adaptadas a diferentes culturas e necesidades. Aínda é demasiado cedo neste campo tecnolóxico para saber se, a longo prazo, as Bibliotecas Dixitais han permitir que diferentes culturas (e entidades) achen solucións adaptadas ás súas propias necesidades ou se, pola contra, estes novos espazos dixitais servirán para promover unha uniformización cultural que utilice como excusa necesidades de estandarización tecnolóxica. Neste sentido, é moi interesante o informe realizado no taller sobre Aspectos Sociais das Bibliotecas Dixitais accesíbel en (<http://www.gslis.ucla.edu/DL/>).

En calquera caso, é de esperar que as Bibliotecas Dixitais non só vaian estar influenciadas no seu desenvolvemento polos factores socio-políticos e económicos antes citados, senón que tamén cheguen elas propias a determinaren, en certa medida, os devanditos factores, cumprindo un papel de transmisoras de cultura e servindo á difusión dun amplo abano de ideoloxías, opinións, linguas, etc. Dito doutro xeito, é posíbel que as Bibliotecas Dixitais axuden a conformar a sociedade futura e esperemos, ademais, que esa sociedade estea caracterizada por unha maior equidade no acceso ao coñecemento e á información. Tal e como a imprenta fixo no seu día, esta tecnoloxía abre a posibilidade de lograr, cun prezo mínimo, un

nivel de difusión de todo tipo de culturas e informacións (textuais e multimedia) nunca acadado na historia da humanidade.

De calquera maneira, a simple aparición da tecnoloxía dixital para o almacenamento e divulgación de documentos fai xurdir algunhas interesantes cuestións: como van influenciar as Bibliotecas Dixitais a industria da información?, cal vai ser o negocio da prensa ou como vai funcionar o negocio editorial?, que mudanzas se van producir nas políticas de dereitos de autor para permitir a divulgación dixital de textos?...

3. Precedentes de Bibliotecas Dixitais

3.1. Evolución histórica das bibliotecas

A partir do momento en que na historia da Humanidade se desenvolve a escrita e xorden os primeiros documentos, aparece, con eles, o interese por conservalos e, xa que logo, as primeiras grandes bibliotecas.

Sucesivos avances tecnolóxicos foron modificando a forma de producir e editar textos e, así, tal e como se pode apreciar no gráfico, aumentou extraordinariamente o número de documentos dispoñíbeis ao igual que as técnicas para a súa catalogación e arquivo nas bibliotecas.

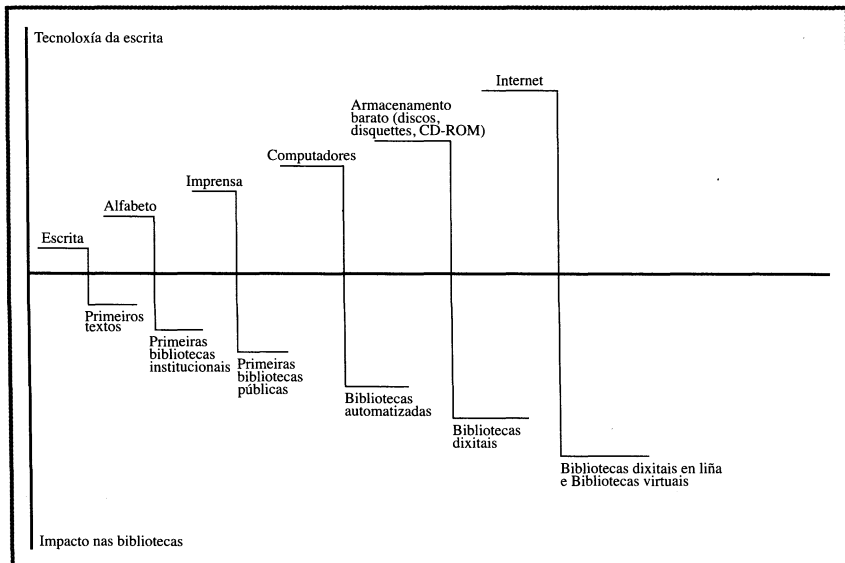


Fig. 1: Evolución das Bibliotecas

A aparición da Informática propiciou a utilización de computadores nos procesos de catalogación, consulta, control de empréstimos, fondos, etc. dando lugar ás bibliotecas automatizadas, xa definidas. Posteriormente, e ligado ao rapidísimo proceso de abaratamento dos dispositivos de almacenamento (discos, disquetes, CD-ROM...) que fixo asequíbel a dixitalización dos propios documentos, xorden as Bibliotecas Dixitais, que, coa aparición da Internet, se fixeron accesíbeis a un público máis amplo.

É de esperar que a progresiva mellora da Internet e o incremento do ancho de banda (velocidade da Rede) faga posíbel que as Bibliotecas Dixitais pasen a ser Bibliotecas Virtuais propiamente ditas. Así mesmo, a maior velocidade da Rede tamén fará posíbel a visualización no navegador web (*Netscape* ou *Internet Explorer*, por exemplo) de vídeo ou a audición de son, o que hoxe é tecnoloxicamente factíbel, mais pouco eficiente por mor das limitacións do soporte físico da Rede (escaso ancho de banda) e, tamén, pola reducida velocidade dos *modems* caseiros que, en xeral, non superan os 56.000 *bytes* por segundo.

3.2. Estado actual da tecnoloxía de Bibliotecas dixitais

Desde a xeneralización do uso da Internet fíxose evidente o importante papel que as Bibliotecas Dixitais poderían xogar na difusión de todo tipo de contidos ao facerse posíbel o acceso aos mesmos para a comunidade internacional de usuarios. A partir de entón a investigación para o desenvolvemento de tecnoloxías de soporte das Bibliotecas Dixitais ten crecido de xeito exponencial.

As Bibliotecas Dixitais presentan problemas non resoltos aínda e que urxe enfrontar de cara a unha plena utilización das posibilidades teóricas das mesmas. Nos últimos tempos, numerosos proxectos de investigación centráronse en diferentes aspectos sen acadaren solucións definitivas. Entre estes aspectos cómpre sinalar: o deseño e desenvolvemento de interfaces de usuario que faciliten ao público a expresión das características dos documentos que están a procurar, o desenvolvemento de técnicas de localización de documentos segundo a súa temática (*Text Retrieval*), a Federación de Bibliotecas Dixitais dispersas xeograficamente baixo unha única interface de usuario nun único enderezo web, técnicas de clasificación e indexación automática, técnicas de procesamento de documentos non textuais (como mapas e contidos multimedia), servizos ao público e un longo etcétera.

A Universidade Urbana-Champaign (Illinois) está a coordinar seis importantes proxectos de Bibliotecas Dixitais (<http://dli.grainger.uiuc.edu/national.htm>), cada un deles desenvolvido por unha universidade estado-unidense diferente (como se pode ver, das máis prestixiosas a nivel internacional no campo da tecnoloxía). Estes

proxectos, de seu, resultan moi ilustrativos e permiten exemplificar os retos prioritarios na investigación no campo das Bibliotecas Dixitais. A seguir presentamos unha breve descrición de cada un deles, pois serven perfectamente para expor as liñas de traballo máis urxentes nesta área tecnolóxica :

- Universidade de Berkeley (California): *Sistemas de Planificación Medioambiental e Información xeográfica*. Este proxecto pretende desenvolver servizos de introdución de documentos e dados, indexación automática dos mesmos e recuperación e procura intelixente de documentos segundo a súa temática. Con todo, o interese deste proxecto radica en que considera que a Biblioteca Dixital está formada por un conxunto de bases de datos documentais dispersas xeograficamente (isto é, trátase dunha Biblioteca Dixital Distribuída). <http://elib.cs.berkeley.edu>
- Universidade de Santa Bárbara (California): *Proxecto Alexandria*. Este proxecto desenvolverá servizos para Bibliotecas Dixitais Distribuídas onde os documentos representan mapas e outros tipos de información xeorreferenciada. <http://alexandria.sdc.ucsb.edu>
- Universidade Carnegie Mellon: *Biblioteca Dixital de Vídeo e Multimedia*. Este proxecto céntrase no desenvolvemento de tecnoloxía para realizar procuras e recuperacións de informacións non textuais senón multimedia (especificamente vídeos). <http://www.informedia.cs.cmu.edu>
- Universidade Urbana-Champaign (Illinois): *Federación de Bases de Datos Documentais de Literatura Científica*. O obxectivo deste proxecto é a creación dunha Biblioteca Dixital de inmensas proporcións, capaz de atender a millares de usuarios simultáneos. <http://dli.grainger.uiuc.edu/default.htm>
- Universidade de Michigan: *Axentes Intelixentes para a Procura de Información*. Este proxecto, baseado en tecnoloxía de axentes (pequenos programas autónomos e intelixentes que interactúan entre si), pretende achar solucións aos problemas de realización de procuras de documentos sobre bases de datos documentais distribuídas, ao tempo que se facilita o traballo dos usuarios no que se refire á expresión das súas consultas. <http://www.si.umich.edu/UMDL>
- Universidade de Stanford: *Mecanismos de Interoperación para Servizos Heteroxéneos*. Este proxecto traballa no tema de Federación de Bibliotecas Dixitais (aspecto este que, pola súa especial relevancia, se desenvolve na última parte deste artigo). <http://walrus.stanford.edu/diglib>

Unha visión de conxunto interesante sobre este tema é a proporcionada por Lesk (1997), así como os artigos de alta calidade que se poden achar nos libros de actas das Conferencias sobre Bibliotecas Dixitais da ACM. Pódese consultar información on-line sobre estes temas en *D-Lib Magazine* (Friedlander 1998) e nas publicacións da *National Science Foundation* (NSF) (<http://www.nsf.gov>) e da *National Aeronautics and Space Administration* (NASA) (<http://technology.nasa.gov>). Por outra banda, resulta inescusábel citar o excelente traballo que, ao longo dos últimos anos, vén realizando Edward A. Fox e a súa equipa (Fox 1999).

4. Contidos e Servizos

As Bibliotecas Dixitais son consideradas por moitos a lóxica evolución das bibliotecas tradicionais para se adaptaren á sociedade da información. Desde esta óptica, a peza chave das Bibliotecas Dixitais serían os contidos (textuais e/ou multimedia), xa que estes foron e son o elemento crucial das bibliotecas tradicionais. Esta focaxe esquece que a tecnoloxía dixital ofrece a posibilidade de mudar radicalmente cando, como, para que e quen pode usar as bibliotecas.

Nos primeiros tempos todos os esforzos de creación de Bibliotecas Dixitais estiveron orientados á selección, manexo e explotación de contidos, aínda que, conforme as experiencias de Bibliotecas Dixitais foron madurando, foise concedendo máis e máis atención aos servizos ofertados aos usuarios. Estes servizos están a ser máis demandados que os contidos en si.

Cómpre lembrar que o aspecto máis popular da Internet é a súa capacidade para conectar persoas, ben vía correo electrónico ou ben vía servizos de novas e intercambio de opinións, asíncronos, como os *news*, ou síncronos, como os *chats*. Parece evidente que o suceso da Internet vén da man de que a xente quere comunicarse, colaborar ou intercambiar ideas e opinións sobre temas diversos.

As Bibliotecas Dixitais deben tirar partido das capacidades propias da Internet, así como do xa aprendido a través da observación do comportamento do público na Rede; deben, ademais, ofertar servizos para facilitar a comunicación e o intercambio de opinións entre usuarios no rico contorno, desde o punto de vista informativo, que é unha Biblioteca Dixital. Así, servizos de procura de información para clarexar calquera punto a debate nun *chat*, ou servizos de lectura colaborativa serán máis determinantes do suceso dunha Biblioteca Dixital do que os seus contidos.

Nas próximas seccións detállanse algúns aspectos relevantes sobre os contidos das Bibliotecas Dixitais e posteriormente descríbense os servizos xerais máis salientábeis.

4.1. Contidos

Aínda que o termo Biblioteca Dixital nos leva automaticamente a pensar en texto dixitalizado, e a maior parte das Bibliotecas Dixitais hoxe en día aínda contan cos textos como contidos básicos, é preciso termos en conta que moito do interese das Bibliotecas Dixitais radica na súa capacidade de almacenaren e proporcionaren de xeito directo outro tipo de contidos. Entre eses contidos de natureza non textual están, por suposto, os multimedia (imaxes, audio e vídeo), mais, probabelmente, non van ser estes contidos os máis demandados nos próximos anos. É moi posíbel que os contidos estrela das Bibliotecas Dixitais sexan aqueles que proporcionen interactividade.

As Bibliotecas Dixitais ofrecen un contorno en que obxectos dinámicos e eventos dixitais atopan o seu lugar natural, non parecendo haber límite para os froitos da creatividade humana neste espazo animado de expresión dixital. Así, novos contidos que non serían posíbeis en bibliotecas convencionais comezan a ser producidos masivamente: textos animados, corpus hipertexto e hipermedia, vídeo baixo demanda, simuladores e mundos virtuais máis ou menos sofisticados...

Os contidos multimedia interactivos van cada día máis lonxe, combinando medios diferentes para proporcionaren aos usuarios mecanismos de control e poderen así elixir entre múltiples opcións, quer dicir, para proporcionar cada vez unha máis variada interactividade.

Outra das fontes de interactividade para un usuario a que recorren as Bibliotecas Dixitais son os demais usuarios. Xorden, deste xeito, todo tipo de escenarios colaborativos para o traballo ou o lecer, lugares de encontro para o debate (*Chats*) ou de lectura colaborativa, así como servizos demandados ás Bibliotecas Dixitais; mais disto falaremos na sección consagrada aos servizos das Bibliotecas Dixitais.

Todo o que contén unha Biblioteca Dixital son “obxectos dixitais”. En principio todo pode dixitalizarse, desde un texto a un vídeo pasando por música ou imaxes fotográficas, mais cómpre explicar que se entende por dixitalización para comprender como poden ser almacenados contidos tan diferentes todos eles na mesma Biblioteca Dixital.

4.1.1. Conceitos básicos sobre dixitalización

A dixitalización significa que o texto, son, imaxe ou vídeo de que se tratar vai ser representado integramente utilizando tan só dous signos, o *0* e o *1*, que son os denominados *bits*. Para a dixitalización de texto utilízase un formato de codificación coñecido como código ASCII. Con este código cada carácter (letra maiúscula, minúscula, número do 0 ao 9, signo de puntuación, letra acentuada, etc.) é representada por unha serie de 8 *bits* consecutivos. Cada sucesión de 8 *bits* recibe o nome de *byte*. Con 8 *bits* tan só se poden codificar 256 caracteres diferentes ($2^8 = 256$), o que resulta insuficiente para representar, mesmo, todos os caracteres existentes nos idiomas europeos (onde existen gran variedade de caracteres acentuados con distintos tiles e varios alfabetos). Unha primeira solución a esta limitación é que o código ASCII teña variantes axeitadas aos diferentes idiomas, aínda que unha solución máis global e recente é a proporcionada por UNICODE, onde o número de *bits* usados para codificar cada carácter é de 16. Dado que con 16 bits poden facerse 32.000 combinacións diferentes ($2^{16} = 32.000$), UNICODE proporciona un sistema de codificación único para todos os alfabetos existentes e aínda garda combinacións (códigos de 16 bits) non usadas para que se poida estender no futuro o conxunto de caracteres representados.

Do mesmo xeito, á hora de dixitalizar unha imaxe, esta divídese nunha sucesión de puntos de moi pequeno tamaño (*pixels*) e a cor de cada un deses puntos codifícase cunha sucesión dun número determinado de *bits* que varía segundo o formato de dixitalización da imaxe. Os formatos máis habituais de dixitalización de imaxes son GIF, TIFF e JPEG.

O son dixital utiliza a mesma estratexia: a onda espectral do son descomponse no conxunto de sinais que corresponde aos seus compoñentes acústicos (descomposta de Fourier) e cada un divídese en mostras de curtísima duración para, despois, codificar mediante series de *bits* o ton, timbre e intensidade de cada unha destas. Formatos de son serían AU, WAV AIFF.

O vídeo segue un proceso semellante: cada fotograma codifícase como unha fotografía estática e o son é codificado dentro do mesmo formato. Formatos comúns de vídeo dixitalizado son Quick Time, AVI, MPEG, Shockwave, etc.

Centrándonos nos textos, obviamente aqueles producidos nun computador son xa textos dixitais, pero se do que se trata é de dixitalizar un texto impreso o proceso a seguir obriga a obter unha imaxe dixital da páxina en cuestión, ou o que é o mesmo, pasar esa páxina por un aparello de escáner. Con este proceso obtense unha representación dixital da páxina como se de unha fotografía se tratar, isto é, non se

teñen letras senón *pixels*. Para converter a imaxe dixital da páxina nun texto dixitalizado é preciso utilizar despois un OCR (*Optical Character Recognition*). O OCR é un *software* que é capaz de recoñecer na imaxe dixital da páxina as manchas en forma de caracteres que forman as letras do texto. Como resultado da utilización do OCR obtense o texto dixitalizado.

Como é obvio, os OCR só poden recoñecer formas de caracteres predeterminados, de tal xeito que non existen hoxe OCR capaces, xa non de interpretaren letra manuscrita, senón nin sequera caracteres de imprenta do século pasado.

Normalmente os OCR recoñecen ben unha serie de caracteres, mais, en xeral, é necesaria a supervisión humana para garantir que o texto obtido sexa correcto. Loxicamente, canto máis antigos ou estraños foren os caracteres do texto orixinal ou máis deteriorado estea o documento orixinal pasado polo escáner (con manchas ou liñas que entorpezan o labor do OCR), máis traballo humano será necesario para corrixir o texto dixitalizado obtido automaticamente polo OCR.

4.1.2. Linguaxes de Marcado

Os textos dixitalizados poden ser simplemente cadeas de caracteres (un carácter tras outro, do primeiro ao derradeiro, entendendo que os signos de puntuación, os espazos en branco que separan palabras e as marcas de punto e á parte son tamén caracteres), mais resulta evidente que se os textos están almacenados, por así dicilo, como chorros de caracteres, vai ser difícil a súa representación axeitada na pantalla do ordenador. Non haberá xeito de distinguir, por exemplo, que un conxunto continuado de caracteres dentro do documento (palabras seguidas) é un título de capítulo, que outro é unha nota a rodapé ou que outro diferente é un poema que está insertado no meio do texto en prosa. Fica claro que a representación dos textos como simples cadeas de caracteres faría imposible a súa representación (edición) axeitada. Así, calquera procesador de textos garda información sobre a fonte e tamaño dos caracteres, o aspecto dos mesmos (negra, cursiva, etc), sobre o sangrado, as marxes, etc.

Cando xurdiu a Internet fíxose necesario adoptar un xeito de indicar as características que permiten representar os textos adecuadamente. Era preciso que todos os navegadores da Internet (para PC, Mac, LINUX, UNIX, etc) fosen capaces de entenderen e representaren as páxinas web. Xorde así HTML, que é un conxunto de etiquetas que permiten especificar nun texto as fontes, aspecto e tamaño dos caracteres, formato de marxes, táboas, etc. Todas as etiquetas sitúanse entre os signos <> (por exemplo, <TABLE> indica comezo dunha táboa,

indica comezo de texto en negra, etc); calquera característica do texto debe ir situada entre súas etiquetas emparelladas, a primeira indicando o inicio da característica e a segunda indicando o fin. As etiquetas de finalización comezan sempre polo signo /. Así, </TABLE> indica o fin dunha táboa e indica o fin da letra negra.

Alén disto, HTML (*HyperText Markup Language*) permite sinalar a presenza de enlaces entre partes do texto, pois xurdiu xa orientado ao hipertexto propio das páxinas web. Na figura apréséntase un texto marcado en HTML e o aspecto que tería ese mesmo texto ao ser lido cun navegador web como *Netscape* ou *Internet Explorer*.

Obsérvase a correspondencia entre as marcas na figura 2 co aspecto do texto da figura 3.

```
<HTML>
<HEAD>
<META HTTP-EQUIV="Content-Type" CONTENT="text/html; charset=windows-1252">
<META NAME="Generator" CONTENT="Microsoft Word 97">
<TITLE>ESTOU ESCRIBINDO O ENCABEZAMENTO</TITLE>
</HEAD>
<BODY>

<FONT SIZE=5><P ALIGN="CENTER">ESTOU ESCRIBINDO O ENCABEZAMENTO</P>
<P ALIGN="CENTER"></P>
</FONT><FONT SIZE=2><P ALIGN="JUSTIFY">Agora estou escribindo un texto normal
pero intercalo algunhas palabras en <B>negra </B>e en<B> </B><I>cursiva</I><B>. </P>
</B><P ALIGN="JUSTIFY"></P>
<P ALIGN="JUSTIFY">Ademais introduzo algunha palabra de
</FONT><FONT SIZE=5>maior</FONT><FONT SIZE=2> tama&ntilde;o.</P>
<P ALIGN="JUSTIFY"></P>
<P ALIGN="JUSTIFY">&nbsp;</P>
<P ALIGN="JUSTIFY">Agora vou representar un cadro:</P>
<P ALIGN="JUSTIFY"></P></FONT>
<P ALIGN="CENTER"><CENTER><TABLE BORDER CELLSPACING=1 CELLPADDING=4 WIDTH=378>
<TR><TD WIDTH="47%" VALIGN="TOP">
<FONT SIZE=2><P ALIGN="CENTER"><B>Columna 1</B></FONT></TD>
<TD WIDTH="53%" VALIGN="TOP">
<B><FONT SIZE=2><P ALIGN="CENTER">Columna 2</B></FONT></TD>
</TR>
<TR><TD WIDTH="47%" VALIGN="TOP">
<FONT SIZE=2><P ALIGN="JUSTIFY">Contido11111111111111</FONT></TD>
<TD WIDTH="53%" VALIGN="TOP">
<FONT SIZE=2><P ALIGN="JUSTIFY">Contido222222222222222222</FONT></TD>
</TR>
</TABLE>
</CENTER></P>

<FONT SIZE=2><P ALIGN="JUSTIFY"></P>
<P>&nbsp;</P></FONT></BODY>
</HTML>
```

Fig. 2: Texto en HTML



Fig. 3: Imaxe do texto lido usando Netscape

Así pois, o contido dunha páxina web é un texto con marcas HTML que indican o formato que o texto debe ter. Porén, o que os usuarios ven, ao se conectaren mediante o seu navegador á páxina web en cuestión, é o texto xa axeitadamente formatado, sen visualizaren as marcas de formato. Isto quer dicir que todo navegador da Internet le, entende e interpreta as marcas HTML e formata os textos segundo o indicado nas mesmas.

HTML permite incluír imaxes, animación, enlaces con outros textos ou con outras partes do texto, etc. Actualmente hai no mercado moitas ferramentas que permiten a creación de páxinas web ou, dito doutro xeito, que facilitan a colocación das marcas nun texto. O propio e coñecido *Microsoft Word* dá a opción de gardar os documentos con formato HTML. Un documento escrito en *Word* e gardado como HTML será visto como texto con marcas HTML se é aberto co Block de Notas de *Windows* (que só interpreta caracteres ASCII e que, polo tanto, non interpreta as marcas HTML como tales marcas indicadoras dun certo formato senón que as representa como unha sucesión de caracteres). Porén, se ese mesmo ficheiro for lido con *Internet Explorer* aparecerá en pantalla xa formatado e non se verán as marcas HTML.

Se os documentos que se queren gardar na Biblioteca Dixital son simples cadeas de caracteres, non se poderán representar axeitadamente para que os usuarios os vexan. Neste caso, HTML podería ser unha alternativa, pois o documento seleccionado polo usuario conectado á biblioteca, ao serlle enviado en HTML, sería automaticamente formatado polo seu navegador web e a súa representación visual sería perfecta.

Porén, HTML ten serias limitacións para o seu uso como formato de almacenado dos documentos nunha Biblioteca Dixital. As limitacións veñen de que HTML está demasiado orientado ao formato final de presentación. Así, dentro dun documento HTML tampouco é posíbel distinguir o título dun capítulo de outro texto que teña o mesmo tipo e tamaño de letra. Isto débese a que as marcas de HTML indican tan só o aspecto con que os caracteres entre pares de marcas deben ser representados.

Para responder a este problema xorde XML (*eXtensibel Markup Language*). XML non é un conxunto de marcas, como HTML, que poden ser inseridas no meio do texto; pola contra, XML é unha metalinguaxe, xa que pode ser usada para criar novas linguaxes de marcado adaptadas a unhas necesidades concretas. Con XML é posíbel definir o conxunto de marcas que se van usar.

Imos imaxinar, por exemplo, que se propón unha Biblioteca Dixital en que se pretende gardar poemas ademais doutros tipos de documentos. Supoñamos que temos interese non só en gardar dados sobre o autor do poema e a data de creación, senón tamén en diferenciarmos os versos e as estrofes, alén de indicarmos cada vez que apareza un nome propio dentro dos poemas.

Para podermos sinalar nos poemas todos eses elementos definimos (fig. 4) un DTD (*Document Type Definition*) onde especificamos as marcas que imos usar cos poemas e, de paso, a estrutura que estas marcas deben ter (o seu aninamento, é dicir, que pares de marcas deben ir dentro doutros pares de marcas).

<!ELEMENT poema	(título, autor, corpo, data) >
<!ELEMENT título	(#PCDATA)>
<!ELEMENT autor	(#PCDATA)>
<!ELEMENT corpo	(cuadratermo)+>
<!ELEMENT cuadra	(verso, verso, verso, verso)>
<!ELEMENT termo	(verso, verso, verso)>
<!ELEMENT verso	(#PCDATA nome)*>
<!ELEMENT nome	(#PCDATA)>
<!ELEMENT data	(#PCDATA)>

Fig. 4: Exemplo de DTD

Un DTD especifica as marcas dispoñíbeis para anotar cada un dos elementos relevantes para nós dun poema, especificando ademais para cada elemento os subelementos que pode ou debe ter. Así, neste DTD estamos a dicir que todo poema debe ter un *título*, seguido dun *autor*, seguido do *corpo* do poema e que, ao final, deberá ir a *data*. Dicimos, tamén, que o *título* é un texto libre (#PCDATA) e que o *corpo* deberá estar composto por un número indeterminado de *cuadras* e *ternos*, aínda que, polo menos, ha ter un deles, isto é, o *corpo* ten de ter un ou máis *cuadras* ou *ternos* (sinálao o + ao final da declaración). Especificase, tamén, que unha *cuadra* debe ter exactamente 4 versos e que unha *terna* non pode ter máis que 3. O *verso* é un texto libre (#PCDATA) que, ademais, pode ter 0 ou máis *nomes* (especificao o *). A respecto do seu aninamento, estase a indicar, por exemplo, que unha marca </termo> non pode aparecer se fica aberta algunha marca de verso, isto é, non sería posíbel o seguinte ordenamento de marcas:

...<verso> ...</termo></verso>

Apreséntase a seguir (fig. 5) un poema anotado segundo as marcas definidas no DTD da figura 4.

```
<?xml versión="1.0"?>
<poema>
  <título>"Soneto Já Antigo"</título>
  <autor>(Álvaro de Campos)</autor>
  <corpo>
    <cuadra>
      <verso>Olha, <nome>Daisy</nome>: quando eu morrer tu hás-de</verso>
      <verso>dizer aos meus amigos aí de <nome>Londres</nome>,</verso>
      <verso>embora não o sintas, que tu escondes</verso>
      <verso>a grande dor da minha morte. Irás de</verso>
    </cuadra>
    <cuadra>
      <verso><nome>Londres</nome> p'ra <nome>Iorque</nome>, onde
                                     nasceste (dizes</verso>
      <verso>que eu nada que tu digas acredito),</verso>
      <verso>contar áquele pobre rapazito</verso>
      <verso>que me deu horas tão felizes,</verso>
    </cuadra>
    <terno>
      <verso>em bora não o saibas, que morri...</verso>
      <verso>Mesmo ele, a quem eu tanto julguei amar,</verso>
      <verso>nada se importará... Depois vai dar </verso>
    </terno>
    <terno>
      <verso>a notícia a essa estranha <nome>Cecily</nome></verso>
      <verso>que acreditava que eu seria grande... </verso>
      <verso>Raios partam a vida e quem lá ande!</verso>
    </terno>
  </corpo>
  <data>(1922) </data>
</poema>
```

Fig. 5: Exemplo de texto en HML

Todo documento XML ten tres partes diferentes (Bradley, 1998):

- **Declaración:** Todo documento XML ten de comezar, obrigadamente, por unha declaración onde se informe da versión de XML que se está a usar, da táboa de caracteres (variante de ASCII como *iso-latin1* o *iso-greek-3*, por exemplo). Así a primeira liña dun documento XML debe comezar por algo como:

<?xml version="1.0" encoding="....." ?>.

- **DTD (Document Type Declaration):** Como xa se explicou, o DTD é o conxunto de declaracións que especifican como debe ser un tipo de documento. Esta parte é opcional no documento XML. Ás veces o que hai é unha referencia ao ficheiro no que está o DTD.
- **Contido:** É o texto marcado propiamente dito. Contén, por tanto, o texto e as marcas de inicio e cerramento de cada parte significativa do documento segundo o estipulado no DTD.

Se temos os textos da biblioteca marcados en XML segundo unha (ou varias) DTD definidas *ad hoc*, poderemos facer automaticamente análises máis finas dos documentos, procedemento que non sería posíbel co texto sen marcar nin con marcas HTML. Por exemplo, poderanse procurar poemas de máis de 7 versos, ou que tiveren dous cuartetos, ou aqueles en que aparecer algún nome propio, etc.

Evidentemente, un documento XML non aporta ningunha información a respecto de como formatalo para a súa presentación, mais isto é outra das súas vantaxes. Para dar formato de presentación a un documento XML cómpre usar unha *folla de estilo*.

Nas *follas de estilo* especificase para cada parte do texto (entendendo por parte do texto a que fica entre a apertura e o cerramento dunha etiqueta das definidas no DTD) como se quere representar. Deste xeito podemos ter moitas “follas de estilo” e utilizar en cada momento aquela que o usuario xulgar máis axeitada. Así, cunha folla de estilo pódese pedir que o nome do autor apareza baixo o poema en letra *courier* 12 e aliñado á esquerda, que os nomes propios aparezan en cursiva e negra ou que entre unha estrofa e a seguinte haxa unha liña en branco. Con outra folla de estilo pódese indicar que os poemas aparezan con letra pequena e en dúas columnas, e así sucesivamente.

Outra das funcionalidades de XML é que pode ser usado para representar, de xeito integrado e uniforme, nun único documento XML acorde a un determinado DTD, o texto transcrito e os datos bibliográficos asociados ao mesmo. Ese texto e os seus datos bibliográficos, de estaren gardados nunha base de datos, estarían de xeito,

disgregado. Así, en ocasións tíranse datos e textos de bases de datos e represéntanse de forma integrada nun único documento XML para os transmitir de xeito conxunto. Tamén á inversa, documentos XML construídos conforme a un DTD son susceptíbeis de seren lidos por un programa que vaia metendo cada elemento do documento no seu sitio correspondente na base de datos.

4.2. Servizos

Aínda que as Bibliotecas Dixitais nacen centradas nos documentos dixitais que teñen que conter é cada vez máis claro o xiro que se está a producir no sentido de considerar as persoas que usan a biblioteca como entidades centrais das Bibliotecas Dixitais.

Serán as políticas que se implementen nas Bibliotecas Dixitais as que determinen quen pode acceder, a que contidos e a que servizos, así como a cantidade, calidade e clase dos servizos implementados. Neste apartado vanse presentar algúns dos servizos máis comúns nas Bibliotecas Dixitais actuais.

4.2.1. Servizos de Procura

Calquera Biblioteca Dixital debe ofertar na súa interface de usuario (sexa esta vía web ou non) a posibilidade de acceso aos fondos dixitais que contén, aínda que talvez aconteza que o acceso estea graduado segundo o nivel de permisos dos usuarios, de xeito que non todos os fondos sexan accesíbeis a todo o mundo.

A procura dos documentos pode estar implementada ao modo das bibliotecas tradicionais, onde, en xeral, só se ofrece busca por autor/a (o primeiro ou todos), título e tema. Porén, as Bibliotecas Dixitais poden apresentar posibilidades máis interesantes, entre as que destacan as seguintes:

- **Procura por todos os campos:** Trátase de que todos os campos sobre os que hai información poidan ser usados para a realización de buscas. Deste xeito sería posíbel realizar procuras como esta: “obras de autores/-as nados/-as en certo país que tivesen publicado un romance de máis de 250 páxinas nunha editorial X entre 1980 e 1985”.
- **Procura por contido:** A procura por contido é, sen dúbida, un dos campos de investigación máis activo desde a aparición da Internet. Na terminoloxía técnica internacional chámase *Text Retrieval* (recuperación de textos) e abrangue un amplo abano de métodos e técnicas orientados a permitir que se recuperen da base de datos (isto é, que se achen como resultado de procuras) textos que teñan no seu contido certos temas. Unha Biblioteca Dixital que tivese implementadas técnicas eficientes de *Text Retrieval*

podería permitir que os usuarios realizasen procuras como: “Novelas do século XX en que se describa algún instrumento musical” ou “Textos en que se describa algunha receita de bruxaría” ou “Novelas do XIX en que se describa unha batalla”.

O interese de ter a posibilidade de realizarmos este tipo de procuras é evidente: sen ir máis lonxe, pódense imaxinar as enormes posibilidades deste servizo en Bibliotecas Dixitais con fondos antigos ao serviren de base para traballos de investigación de antropólogos, historiadores, etc.

Moitas Bibliotecas Dixitais actuais teñen implementados servizos básicos de procura por contido que, en xeral, non van máis lonxe da busca de documentos que conteñan certa(s) palabra(s) ou cadea(s) de caracteres (raíz dunha palabra). As técnicas de *Text Retrieval* están en diferente nivel de desenvolvemento segundo a lingua de que se tratar. O inglés é, sen dúbida, a lingua que conta con máis ferramentas básicas (dicionarios, *Thesaurus* de sinónimos e antónimos, redes de asociación semántica de palabras, lematizadores, etc.) entre aquelas necesarias para implementar técnicas de *Text Retrieval*. Os idiomas portugués e español teñen aínda un longo camiño por andar, embora xa conten con dicionarios electrónicos e ferramentas de lematización. No que atinxe ao galego está todo por facer.

4.2.2. Servizos de intercomunicación de usuarios

Estes son, sen ningún tipo de dúbida, despois dos servizos de procura por múltiples campos, os máis implementados nas Bibliotecas Dixitais actuais. Vendo o suceso da Internet á hora de atraer usuarios grazas ás posibilidades de intercomunicación que brindaba, as Bibliotecas Dixitais decidiron ofrecer este tipo de servizos decontado. Entre estes servizos destacan os xa moi coñecidos de *mail*, *chat* e grupos de noticias (os taboleiros de anuncios non se inclúen aquí, pois son máis ben un servizo de espallamento de información).

A calidade destes servizos e a importancia que as diferentes Bibliotecas Dixitais lles dan varía amplamente, así como o aspecto das interfaces de usuario que se proporcionan para acceder aos mesmos.

4.2.3. Servizos de espallamento de información

Entre estes pódense citar, avisos aos usuarios de novos fondos, creación de servizos ou a aparición de grupos de interese xurdidos no seo da Biblioteca Dixital. Poden, tamén, existir servizos de espallamento de información proporcionada por outros usuarios, polos administradores da biblioteca ou por empresas asociadas (publicidade). O espaxamento da información pode ser xeral (a todos os usuarios que se conecten) ou específica (a grupos de usuarios segundo criterios de agrupamento elixidos por eles

proprios, por exemplo, grupos de interese xurdidos libremente na biblioteca) ou segundo perfís de usuario predefinidos como se detalla no apartado seguinte.

Un interesante servizo deste tipo podería ser ofrecer a valoración que sobre certa obra dan os usuarios que a foron lendo, ou estatísticas sobre os libros a que máis xente ten accedido ou que foran mellor valorados (pola totalidade dos usuarios/-as ou por aqueles ou aquelas cun certo perfil).

Evidentemente nesta categoría de servizos cabería o feito de proporcionar enlaces a páxinas web doutros subministradores de información ou Bibliotecas Dixitais (enderezos de prensa, etc.)

4.2.4. Servizos asociados a perfís de usuario

A idea de base é que diferentes usuarios teñen diferentes intereses, gostos, etc. e que estas preferencias son, até certo punto, constantes. Trátase, polo tanto, de que a biblioteca sexa capaz de identificar a persoa que se está a conectar e actúe en consecuencia ofrecéndolle:

- **a interface de usuario** co aspecto axeitado (a que a persoa usuaria deseñou ou elixiu na última visita á biblioteca).
- **as noticias** sobre novos fondos do interese dese usuario/-a específico/-a (pénsese, por exemplo, en usuarios infantís a quen a biblioteca presenta interfaces de funcionamento máis sinxelo, cunha presentación máis atractiva e colorida e que, ademais, avisa de novos fondos axeitados á súa idade).
- **Os enlaces a páxinas web, servizos de noticias e chats** polos que o/a usuario/-a amosou interese en pasadas visitas á biblioteca.
- **A lista de obras xa buscadas** en pasadas visitas para que quen entrar de novo na biblioteca poida proseguir a súa lectura sen ter que voltar a realizar a procura.

Parece claro que para que un servizo deste tipo sexa posíbel cómpre que os usuarios se identifiquen ao accederen á biblioteca ou, cando menos, que opten entre ser usuarios anónimos, con interface de usuario xenérica, ou usuarios rexistrados na biblioteca, gozando, xa que logo, dos beneficios de contaren cun perfil específico. Evidentemente estas consideracións son tamén relevantes para os servizos de espallamento selectiva de información.

O perfil de usuario pode establecerse a partir de datos subministrados pola/o propia/o utente, aínda que tamén (ou ademais) a través do comportamento desta persoa a última vez que visitou a biblioteca (que consultas realizou, en que foros de debate ou noticias participou, que enlaces a prensa realizou, etc.).

4.2.5. Servizos de valor acrescentado

Baixo este nome inclúense unha serie de servizos que teñen pouco en común entre eles, excepto axudar á/ao usuaria/o a utilizar os fondos da biblioteca. Englóbanse aquí servizos de tradución automática, de preparación de bibliografías, etc.

Algunhas bibliotecas inclúen entre estes servizos procura de opinións ou críticas a certas obras enviadas por outros usuarios. Por exemplo, para unha persoa que ten dúbidas sobre cal vai ser o seguinte romance que quere ler, pódelle interesar saber cales foron ben avaliados por usuarias/os que cualificaron anteriores lecturas comúns de maneira semellante á súa, ou cal foi a avaliación feita por personalidades coñecidas ou por usuarios/as concretos.

5. Federación de Bibliotecas Dixitais

Nos últimos anos as Bibliotecas Dixitais proliferaron por todo o mundo ao tempo que se tornaba evidente a necesidade de aglutinar esforzos para, dunha banda, non repetir traballo xa feito (de catalogación e dixitalización das mesmas obras nas diferentes bibliotecas) e, doutra, facilitarlle ás/aos posíbeis usuarias/os a obtención dos documentos que estivesen a procurar.

A Federación de Bases de Datos é un motivo clásico na investigación en Informática e os seus resultados pódense aplicar case directamente ás Bibliotecas Dixitais. A arquitectura básica dunha Federación de Bases de Datos é a que aparece na figura 6.

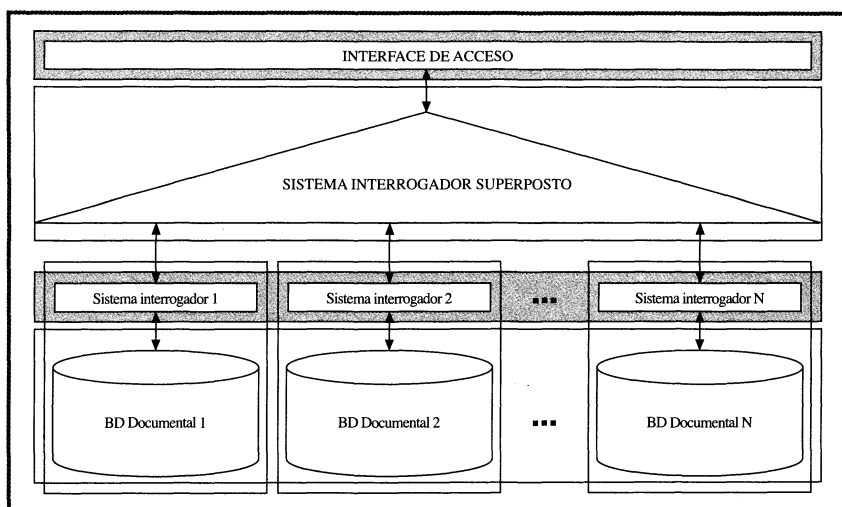


Fig. 6: Arquitectura básica dunha Federación de Bases de Datos

Obsérvase que o sistema federado facilita ao público un único enderezo web de conexión. Neste enderezo ofrécese unha interface de usuario para acceder ás Bases de Datos Documentais que soportan as bibliotecas dixitais. Esta interface de acceso permite que, de xeito máis ou menos sinxelo, unha persoa poida expresar as súas restricións de consulta, isto é, as características dos documentos que está a procurar. Só os documentos que satisfíxeren as restricións expresadas polo usuario serán recuperados polo sistema.

As restricións de consulta expostas polo usuario na interface web de consulta son recibidas polo Sistema Integrador Superposto e redirixidas a cada unha das Bases de datos das Bibliotecas Dixitais federadas. Como tales Bases de Datos poden estar soportadas por sistemas de xestión de bases de datos diferentes (e que falen distintas linguaxes) e o conxunto de datos que cada unha pode conter pode ser tamén moi variado, o Sistema Integrador Superposto deberá adaptar a consulta expresada polo usuario a cada unha das Bases de Datos. Ademais cada Base de Datos ten que ter un sistema interrogador adaptado ás características da mesma. Polo tanto, un sistema de federación de bibliotecas dixitais ten que contar coas 4 capas superpostas que se amosan na figura 6.

5.1. Precedentes de Federacións de Bibliotecas Dixitais

Unha das primeiras Federacións de Bibliotecas Dixitais foi a NCSTRL (*Networked Computer Science Technical Report Library*) (cf. Davis 1995 e Davis/Lagoze 1996), que é un sistema de busca de informes técnicos realizados en universidades (norte-)americanas. O sistema proporciona unha única interface de usuario, pero a procura realízase nas bases de datos de informes técnicos de todas as universidades. Un sistema similar é *Harvest* (Bowman 1994), que segue a mesma filosofía de procura distribuída. Pódese citar tamén o sistema AQUARELLE (<http://aqua.inria.fr/Aquarelle>) que proporciona o acceso federado ás bases de datos catalográficas dun conxunto de museos europeos, e a Biblioteca Dixital ERCIM (<http://www.iei.pi.cnr.it/DELOS/>) que federa Bibliotecas Dixitais europeas sobre Informática e Matemáticas.

Máis complexo é o proxecto *Gutenberg* (<http://www.promo.net>) que asocia moitas Bases de Datos bibliográficas. Este é un proxecto que trata de promover a creación de Bases de Datos documentais de acceso gratuito e facilitar a elaboración e o acceso ás mesmas. Os usuarios poden utilizar datos bibliográficos (autor, título, etc.) como campos de procura, mais tamén poden buscar documentos sobre un tema determinado.

O proxecto *PhiloBiblon* (<http://sunsite.Berkeley.EDU/PhiloBiblon/phhm.html>) está concibido para a busca en tres bases de datos diferentes: manuscritos antigos galegos-portugueses, manuscritos antigos cataláns e manuscritos antigos españois.

Estas tres bases de dados non conteñen textos transcritos senón unicamente un xogo de palabras-chave que describen cada documento. A procura pódese realizar sobre todas as bases de dados ou só sobre unha delas.

A Federación de Bibliotecas Dixitais é unha área de investigación moi activa, pois as Bibliotecas Dixitais presentan problemas novos que a Federación de Bases de Dados non levantaban. Por un lado, é necesario que se poidan explotar as capacidades de recuperación de textos das Bibliotecas Dixitais federadas (Brisaboa 2001a, 2001b) e, por outro, cómpre desenvolver tecnoloxía que permita federar Bibliotecas Dixitais cuxos fondos estean en linguas diferentes (Brisaboa 2001c).

6. Conclusións

Aínda que no momento presente as posibilidades da Internet están a ser só albiscadas, xa está a producir mudanzas no noso xeito de acceder a todo tipo de información. Cabe supor que este proceso vai continuar de xeito imparábel e que vai provocar radicais mudanzas nos hábitos sociais.

Un exemplo concreto de influencia da Internet na sociedade son as Bibliotecas Dixitais. De termos en conta que esta tecnoloxía, como xa se dixo, permite obter, cun prezo mínimo, un nivel de difusión de todo tipo de culturas e informacións (textuais e multimedia) nunca acadado na historia de humanidade, cómpre desenvolver todas aquelas liñas de traballo que posibiliten a introdución de todo tipo de materiais nas Bibliotecas Dixitais.

Os condicionantes sociopolíticos do desenvolvemento das Bibliotecas Dixitais non só se van manifestar nas políticas que implementen os encargados de tales bibliotecas ou os poderes públicos en xeral senón que tamén van ser exercidas polo propio desenvolvemento tecnolóxico. Neste campo cómpre que as equipas de traballo, para seren efectivas, sexan multidisciplinares, integradas por especialistas procedentes do campo da Tecnoloxía e mais das Humanidades.

É posíbel que as Bibliotecas Dixitais axuden a conformar a sociedade futura e, se así for, moi probablemente esa sociedade debería estar caracterizada por unha maior equidade no acceso ao coñecemento e á información, mais isto só vai ser unha realidade se a sociedade, todos nós, utentes e científicos, orientamos o desenvolvemento destas Tecnoloxías nun sentido aberto e non excluínte de ningún colectivo cultural, lingüístico ou social.

Referencias bibliográficas

- Bowman, C. / Danzing, P. / Hardy, D. / Manber, U. / Schwartz, M. (1994): "Harvest: A Scalable, customizable discovery and access system",
<http://www.cs.colorado.edu/harvest/>.
- Bradley, N. (1998): *XML Companion* (Reading, MA, USA: Adison Wesley).
- Brisaboa, N. R. / Durán, M. J. / Penabad, M. R. / Places, A. S. (2000): "A Collaborative Framework for a Digital Library" en *Sixth International Workshop on Groupware CRIWG'2000*: 104-111 (Los Alamitos, CA, USA: IEEE Computer Society Press).
- Brisaboa, N. R. / Places, A. S. / Rodríguez, F. J. (2001a): "Arquitectura para Federación de Bases de Datos Documentales basada en Ontoloxías" en Torres Rojas, F. / Araya Monge, J. E. / Sandoval Sánchez, Y. (ed.): *4º Workshop Iberoamericano de Engenharia de Requisitos e Ambientes de Software (IDEAS'2001)*: 252-262. (Cartago, Costa Rica: Instituto Tecnológico de Costa Rica).
- Brisaboa, N. R. / Penabad, M. R. / Places, A. S. / Rodríguez, F. J. (2001b): "Using ontologies for federation of Web accessible databases" en *13th international conference on software engineering & knowledge engineering (SEKE'2001)*: 87-94 (Skokie, IL, USA: Knowledge Systems Institute).
- Brisaboa, N. R. / Places, A. S. / Pérez-Sanjulián, C. F. / Rodríguez, F. J. (2001c): "An Architectural Proposal for a Cross-Language System to Federate Multilingual Digital Libraries", en Russian Academy of Sciences (org): *Digital Libraries: Advanced Methods and Technologies, Digital Collections* (Petrozavodsk: Institute of Applied Mathematical Research).
- Davis, J. (1995): "Creating a Networked Computer Science Technical Report Library", *D-Lib Magazine*, <http://www.dlib.org/>.
- Davis, J. / Lagoze, C. (1996): "The Networked Computer Science Technical Report Library", *D-Lib Magazine*, <http://www.dlib.org/>.
- Fox, E. A. / Sornil, O. (1999): "Digital Libraries" en Baeza-Yates, R. / Ribeiro-Neto, B. (1999): *Modern Information Retrieval* (New York: Adison Wesley/ACM Press).
- Friedlander, A. (1998): "D-Lib Magazine: Research in Digital Libraries", *D-Lib Magazine*, <http://www.dlib.org/>.
- Lesk, M. (1997): *Practical Digital Libraries; Books, Bytes and Bucks*. (San Francisco, CA, USA: Morgan Kaufmann).
- Romero, D. / Vaquero, I. (2001): *Da periferia á rede. Internet en Galicia. Lingua e contidos* (Vigo: Xerais).