

## EL FORMAT XML OBRE LA POSSIBILITAT D'ACCÉS A CATÀLEGS BIBLIOGRÀFICS

Adela d'Alòs-Moner (Ajuntament de Barcelona) Montse Comaposada (Ajuntament de Barcelona). Jordi Serrano (UOC), Begonya Aguilera (UOC),

### Resum

L'estandardització de formats i de protocols ha estat sempre un aspecte molt important en el sector de les biblioteques. Un exemple clar el tenim en el protocol Z39,50 que es dissenya pensant en la necessitat de poder-se comunicar entre sistemes diferents de gestió automatitzada.

El món Internet, a través de llenguatge HTML trenca amb un món tancat d'informació dipositada en sistemes locals per fer-lo accessible a través de la xarxa amb qualsevol navegador.

Per a les biblioteques i els centres de documentació l'HTML té, però, una dificultat, i és la de no poder codificar documents estructurats.

Quan l'any 1998 apareix l'XML s'obren noves possibilitats per a la recerca integrada a través de catàlegs diferents.

El projecte europeu GALA (Global Access to Local Applications and Services) [www.galanet.org](http://www.galanet.org), liderat per l'Ajuntament de Barcelona, i en el qual participa com a soci local la Universitat Oberta de Catalunya, integra, a través de la utilització del format XML, catàlegs bibliogràfics de diferents països i basats en diferents plataformes.

Aquest accés a informació bibliogràfica es pot fer a través de qualsevol navegador però també -i aquest és sens dubte un aspecte molt novedós- a través de la telefonia mòbil.

### Comunicació

L'Ajuntament de Barcelona és el coordinador del projecte europeu **GALA** (**G**lobal **A**ccess to **L**ocal **A**pplications) [www.galanet.org](http://www.galanet.org), en el qual s'han estat duent a terme experiències d'integració de bases de dades de diferents ciutats europees.

El projecte –que començà al gener de 1998 i finalitzarà al desembre d'aquest any- té com a objectiu crear una **Comunitat Digital Virtual**, que ha de permetre als ciutadans accedir a **informació d'interès general** de diferents ciutats, a partir **d'una única interfície de consulta**.

GALA proporciona informació i serveis relacionats amb temes generals d'interès sobre les cinc ciutats que hi participen (Barcelona, Bolònia, Colònia, Goteborg i el Sud-Est de Londres): equipaments, hotels, l'agenda de la ciutat, transport, medi ambient, salut i biblioteques.

El que fa de GALA un sistema únic és que aquest accés es duu a terme a través d'una única interfície de consulta disponible en sis llengües. Les plantilles de visualització també es troben en sis llengües, tot i que la informació està en la llengua original de cada país.

Les informacions que es visualitzen sobre els diferents aspectes de la ciutat són les més bàsiques i el sistema, aparentment, sembla molt senzill. En realitat l'arquitectura sobre la que es basa és bastant complexa ja que:

- **les aplicacions on resideixen les dades estan desenvolupades en arquitectures diferents** i es mantenen localment. No hi ha, per tant, rèpliques de les bases de dades.
- s'ha hagut de definir i desenvolupar un mòdul ("GALA engine") capaç **d'integrar les diferents aplicacions per traduir-les a un format comú.**

Aquesta integració de dades ha estat possible gràcies al **format XML que permet codificar documents d'una manera senzilla**. Fins ara, dels formats disponibles, l'HTML no ho permetia i, si bé amb SGML sí que era possible, la seva complexitat ho feia inviable.

**Per a dur a terme la integració de totes aquestes aplicacions s'ha definit, doncs, una arquitectura tècnica basada en el format XML (Extensible Markup Language).**

L'XML és un estàndard "de facto" des de febrer d'aquest any quan les principals empreses del sector decideixen acceptar-lo i el W3C (World Wide Web Consortium) el recomana.

Per entendre una mica més el format, es pot dir que un document XML és un fitxer ASCII que conté text i "tags" o etiquetes, que identifiquen estructures en el text.

El format XML és descriptiu per si mateix ja que les meta-dades estan integrades dins de la descripció.

Per exemple:

```
<xml versio="1.0" ?>
<publicació_tipus=llibre>
<tipus>
<títol>The tall ship</títol>
<autor>Clark, Hyla M</autor>
<data>1998.</data>
<matèries>
  <matèria>Revolució Americana, 1776-1976</matèria>
  <matèria>Navegació</matèria>
<isbn>89-9393337383-2</isbn>
<biblioteques>
  <biblioteca_localització>
    <localització="A1">
      Bib. de les Corts
```

```

</biblioteca>
<biblioteca_localització>
<localització="C2">
Bib. de Sant Pau
</biblioteca>
</biblioteques>
</publicació>

```

En el cas concret de la integració de catàlegs de biblioteques, un cop consensuats els camps comuns de recerca i de visualització, la tasca tècnica s'ha centrat en:

- Definir un model de dades, amb els camps pels qual es vol fer la recerca i els que es volen visualitzar. Aquest model de dades s'ha definit en format DTD (Document Type Definition). Els DTDs que són els que descriuen els documents XML, especifiquen el format del document i identifiquen les etiquetes (o "tags") que han d'aparèixer-hi. Els DTD són indispensables perquè el sistema entengui els paràmetres quan es fa una recerca a les bases de dades.
- Desenvolupar el que s'anomena el "GALA engine" que té com a funció consultar les bases de dades mostrar els resultats en un format comú unificat. En el cas de GALA, el "GALA engine" ha estat desenvolupat per Londres, que és la ciutat que coordina aquest sector dins del projecte, i en la definició hi ha participat el BLCMP Library System.
- Desenvolupar els anomenats "Applications Agents" que processen la consulta cap al "GALA engine". En concret, en el projecte GALA, la Universitat Oberta de Catalunya (UOC) ha estat el soci local de l'Ajuntament de Barcelona que ha desenvolupat l'"Application Agent" per integrar les seves dades a nivell transeuropeu. D'aquesta manera es poden consultar – d'una manera totalment transparent, a través d'una única interfície de recerca i de visualització- catàlegs sustentats sobre sistemes diferents, com Dynix i VTLS.

Així, en el nostre exemple concret el DTD podria ser:

```

<?dtd versio="1.0">
<!ELEMENT publicació (títol, autor, edició, matèries, isbn, biblioteques)>
<!ATTLIST publicació_ tipus (REVISTA / LLIBRE #REQUIRED)
<! ELEMENT títol (#PCDATA)>
<! ELEMENT autor (#PCDATA)>
<! ELEMENT data (#PCDATA)>
<! ELEMENT matèries (MATERIES*)>
<! ELEMENT matèria (#PCDATA)>
<! ELEMENT isbn (#PCDATA)>
<! ELEMENT biblioteques (BIBLIOTECA*)>
<! ELEMENT biblioteca (#PCDATA)>
<! ATTLIST biblioteca localització CDATA #REQUIRED>

```

Aquesta és una mostra simple del DTD. Si es vol poder buscar o poder visualitzar més camps, tan sols és necessari indicar les etiquetes que es volen fer aparèixer.

Sens dubte, de tots els sectors integrats en GALA, el de les biblioteques ha estat un dels més complexos. Això ha estat degut a que en les altres aplicacions -hotels, agenda de la ciutat, transports, hospitals, etc.-) cal escollir una ciutat i, un cop aquesta ha estat seleccionada, es quan es pot portar a terme la recerca. Aquesta recerca s'executa sobre única base de dades de la ciutat en concret d'on es vol obtenir la informació. No hi ha, doncs, recerques conjuntes a bases de dades diferents, com és en el cas de les biblioteques.

En el cas de les biblioteques, la recerca es fa cridant a diversos catàlegs al mateix temps. És a dir, el sistema permet obrir diverses bases de dades simultàniament i mostrar una recerca combinada dels resultats.

Amb el que s'ha estat desenvolupant fins ara amb GALA, s'ha demostrat que **és possible accedir a través d'Internet a catàlegs integrats de biblioteques, suportats sobre arquitectures i sistemes d'automatització diferents.**

Una extensió del projecte vol **desenvolupar també un protocol de tipus WAP (Wireless Application Protocol) que ha de permetre accedir a tota la informació integrada a través d'un únic punt d'accés i utilitzant la telefonia mòbil.**

WAP es un estàndard "de facto" desenvolupat per les principals empreses de telefonia mòbil d'Europa, que facilita l'accés a Internet per telefonia mòbil. Moltes aplicacions ja s'han començat a desenvolupar en altres sectors com banca, turisme (reserves per hotels i restaurants), salut pública (avisos en situació d'emergència, contaminació ambiental...), i es preveu que moltes d'elles estiguin plenament operatives en pocs mesos.

Aplicat al sector de les biblioteques, el protocol WAP pot permetre consultar, a través del telèfon mòbil, diferents catàlegs –sempre creats i mantinguts en entorns diferents-, fer reserves de llibres, consultar els horaris i serveis de la biblioteca i, fins i tot, utilitzant tecnologia "push", ser avisat pel sistema quan el llibre o el document que es vol en préstec estigui disponible.

És evident que els fabricants de software de biblioteques tenen uns nous reptes que afrontar per poder respondre, amb la suficient agilitat i rapidesa, a noves prestacions i nous serveis que ben aviat estaran disponibles en altres sectors. De no ser així, un sector que ha estat pioner en molts aspectes en innovació tecnològica, pot quedar fàcilment "by passat" per les innovacions que es puguin dur a terme en altres àmbits.

### **Que cal fer per a poder dur a terme la integració de catàlegs?**

El que cal és relativament senzill. Tan sols és necessari:

- que les biblioteques tinguin les bases de dades bibliogràfiques visibles des d'Internet.

- desenvolupar un mòdul que transformi la informació bibliogràfica en el format XML especificat.

En qualsevol cas s'haurà definit prèviament un model de dades d'integració, comú per a totes les biblioteques.

L'arquitectura distribuïda sobre la qual s'ha basat GALA, al mateix temps que permet oferir un punt de consulta únic i integrat a través de navegador, té, com a principals avantatges:

- que cada centre pugui utilitzar el programa que més s'ajusta a les seves necessitats;
- que les dades resideixin i es mantinguin localment, evitant, d'aquesta manera "mirrors" o bases de dades replicades;

Si hem de senyalar algun desavantatge podem dir que el principal rau en el fet que no hi ha un catàleg unificat amb un control d'autoritats, d'autors o de matèries. Això pot comportar que la mateixa publicació es trobi catalogada de maneres diferents i es pugui produir una pèrdua de precisió en la recerca.

D'altra banda, en el cas d'una aplicació trans-europea –com és el cas de GALA- s'ha de tenir en compte també les diferències lingüístiques que fan inviable la recerca conjunta de catàlegs per determinats camps, com és el de les matèries.

En qualsevol cas, **l'experiència duta a terme amb GALA obre moltes possibilitats a la recerca integrada de catàlegs de biblioteques desenvolupats i mantinguts en entorns i arquitectures completament diferents.** A tall d'exemple, dins del mateix Ajuntament de Barcelona es vol utilitzar aquest model per integrar en un únic format de recerca i de visualització catàlegs suportats en entorns diferents (gestors integrats de biblioteques, bases de dades documentals o relacionals).

El demostratiu de GALA (Global Access to Local Applications) està disponible a l'adreça: [www.galanet.org](http://www.galanet.org).

## **Bibliografia i adreces URL**

### **Bibliografia**

Bosak, John; Bray, Tim. "XML and the Second-Generation Web". *Scientific American*,. (May 1999), p. 79-83

"The Extensible Markup Language (XML)". *ETHOS Technology Briefings Series 1: Developments Shaping Internets and Intranets - Briefing TB 1.13*, (Feb. 1998)

Randall, Neil. "XSL: How stylish can you get?". *PC Magazine*, (April 6, 1999), p. 217-219

### **Adreces URL**

Frequently Asked Questions about the Extensible Markup Language: The FAQ. Ver 1,5 <<http://www.ucc.ie/xml>>

Msdn online: XML Extensible Markup Language.  
<<http://msdn.microsoft.com/xml>>

W3C Architecture domain: Extensible Markup Language (XML)  
<<http://www.w3.org/XML>>

Agost de 1999