Archives on the Web and users expectations

towards a convergence with Digital Libraries

Pierluigi Feliciati, University of Macerata (IT)

Alessandro Alfier, IBC (IT)

FONDAZIONE DEL MONTE 1473
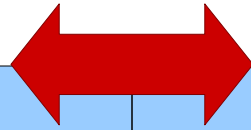
IBC parer

FONDAZIONE CASSA DI RISPARMIO IN BOLOGNA

contents

Why: lack of convergence of Archives on the web towards Digital Libraries (DLs)

What: the quality of Archives on the web: present issues

What: the importance of user studies to evaluate effectiveness of archives on the web

How: *Città per gli Archivi* project: a case study

Why?
The lack of convergence

Digital libraries	Archives on the web
User centric	Material centric
Providing access by subject	Providing access just by "provenance"
Focused on "aboutness"	Focused on "ofness"
Defined by a simple structure (horizontal / net setting)	Defined by a multilevel structure (hierarchical setting)
Focused on item level and digital reproductions	Focused on different, related descriptions
Sensitive to human users and web agents (interoperability)	Sensitive to human users and rarely to web agents
Use of simple standards (i.e. Dublin Core)	Use of complex standards (EAD and EAC)

sacerdotal vs. secular archival service

the displaying of archival descriptions on the Web has a heavily different nature from the traditional presentation: **there is no more mediation**
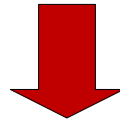
A traditional **sacerdotal approach** (users relate with the archival information only by archivists' mediation)

VS

An emerging **secular approach** (users relate directly with the archival information)

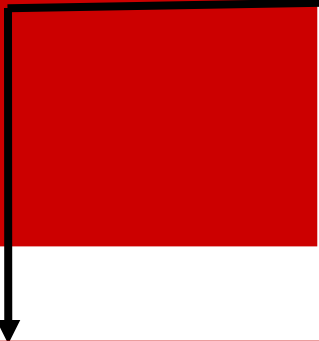Again, no convergence
towards digital libraries methodology

Why: input, output and quality

Input   **Output**

Strict application of archival standards, made for input

Archival information has a technical nature and it's highly encoded

Encoded archival information is "intangible":

it is decoded only by oral communication between archivists and users

Analogical enviroment

Archival information itself assumes the responsibility of its decoding:

it must work as a "**boundary spanner**", mediating between the cognitive and intellectual abilities of users and the nature of the archives
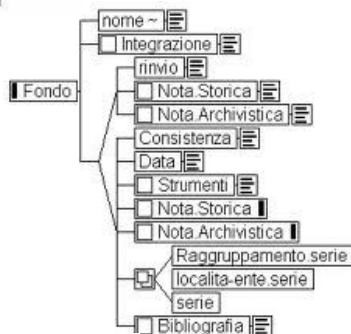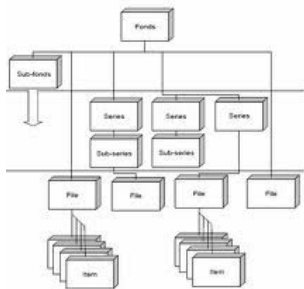
Digital enviroment (Web)

The application of archival standards: a **necessary but not sufficient condition** for the quality of archives on the web

Mediation of archivists in the interstitial space between input and output, to enable the archival information to be easily usable by applying digital libraries best methods

What?

output is not input!

Archival **terminology**:
Language used is too technical

Archival **structure**: hierarchic
browsing guided just by titles

Archival **search**: no-intuitive tools

Archival **contents**:
encoded informations

Archival **coverage**:
what may I find?

What?
quality depends on users

Quality of a digital service: "the capability of the software product to enable specified users to achieve specified goals with effectiveness, productivity, safety and satisfaction in specified contexts of use." (ISO/IEC 9126-1:2001)

Three core elements in a DL: the content it makes available, the services it provides, and the users. Understanding the needs, expectations, and perception of users, whenever these needs are explicit or not, is core to the delivery of a good service.

The North American archivists proposed an **inductive method**: building **good archival displays** – based on the contributions of web design, human-computer interaction studies and cognitive psychology – and checking them through **user studies**.

Programming user study activities inside archival portals' projects appears to be more and more a necessity in order to guarantee that return on investment of projects, whose **first profit should be users satisfaction**.

How?
The CPA case study

Una città per gli archivi (A city for Archives – CPA*)* is a project supported by two bank foundations, with the collaboration of some institutions

Its mission is to improve the **preservation, valorization and full access** for the **most important Archives** related to the **contemporary history** of the city of **Bologna** (Italy)

The **huge project started in 2007**. 200 archives – with 300.000 photographic and graphic materials, sound and audiovisual recordings in addition to paper documents – have been described using ICT tools. 70 archivists were involved. 200.000 digital objects.

A **Web portal** was developed, to enable users to access the described archives and the digital reproductions.

Export of metadata towards SAN, Culturaitalia, Europeana planned.

CPA portal prototype
a snapshot

Language, structure, search

Language: not using, as far as possible, the archival jargon for the labels of archival descriptions and for search tools; adopting terms of common sense and, as far as possible, universally known.

Structure: making the browsing among the hierarchical archival descriptions as intuitive as possible; implementation of tools alternative to hierarchical browsing, such as authority files of access keys + subjects

Search: supporting the immediate **approach of users** towards the documents, not archivists approach towards documentation; implementation of an **information retrieval engine** based on an algorithm of **natural language**, that allows the management of **synonymies** and the ranking of results based on a their **semantic relevance**, inspired to the the general search engines

Content: from the ICT side: re-embedding of metadata into text strings as close as possible to natural language, with the CMS supporting automatic retrieving and exctraction of metadata;
From the archival side: giving preference to structured texts, to facilitate users decoding even in case of long contents; banning the use of archival abbreviations and terms.

Coverage: maximum clearance on project coverage: paper archives, photos ans graphics, sound and audiovisual recordings related to the contemporary history of the city of Bologna (XIX – XX century), created by public, private and religious corporate bodies, persons and families

The **CPA Portal** is presently in a **prototype phase**, so it's not possible to investigate real users experience, but **user studies** will be crucial to guarantee this **service quality**: recommendations for present release finalisation and for future developments

Complexe study methods applied (qualitative data):
- **1 focus group with stakeholders** (content creators + ICT staff + bank foundation staff): **+20 people**
- **5 focus groups with final users** (2 university students groups, 1 high schools students, 1 archivists + cultural heritage + librarians, 1 high school teachers + general public): **+60 people**
- **7 experts evaluation** (3 archivists, 3 experts on digital CH, 1 ICT university professor)

conclusions

- The web environment undermines the **traditional cultural mediation** between archivists and users

- The web **output** must be distinguished from the ***encoded input***, to let it be *decoded* and clear; the archivists should offer user centric diplays of their contents

- To build up good archival diplays the archivists should mediate their descriptive techniques with **human-computer interaction studies** and **cognitive psycology**, checking their prototypes applying user studies

- **User studies** are crucial to guarantee service efficiency: both in the **prototype** phase (to finalise it) and in the **online phase** (to check users' satisfaction and behaviours)

14

Have any questions?

pierluigi.feliciati@unimc.it
aalfier@regione.emilia-romagna.it