Open Is Not Enough
Grey Literature in Institutional Repositories

Joachim Schöpfel (University of Lille 3)
Isabelle Le Bescond (University of Lille 1)
Hélène Prost (INIST-CNRS)
OA and GL

Great diversity of IR

IR collect intellectual output for long-term access

Open is not enough

Different business models

IR are significant channel of digital scientific communication

GL at home in IR
Empirical studies show that empirical studies show the importance of GL with rapid growing. Relevance of IR for GL shows that 82% of IR contain GL. Natural home for GL is IR. IR increase availability, but not quality. Probably large usage of grey items is observed. 20 to 30% of repository content is grey.
GL in IR
significant issues

Community
Quality control
New item formats
Metadata
Integration
Interoperability
Succeeds to Grisemine, the first open archive for GL in France.

A story of trial and error, search for opportunities, benchmarking, exploration and adaptation to a moving context.
2001-2005
Rise and decline of Grisemine

Collect, preserve and disseminate French GL

CinDoc software from Cincom
- Compliant with Dublin Core metadata
- Technically viable

1,300 documents in late 2005, widely consulted

No real institutional recognition
The initial goal was too ambitious
CinDoc software not adjustable for evolution
2006-2010
From Grisemine to IRIS

National decree on ETD processing
Mandatory deposit of e-theses at Lille 1
Recognition of IRIS repository

Migration from CinDoc to DSpace
Connection to the national STAR infrastructure
Heritage collection on history of sciences

625 theses and 711 other documents in IRIS in late 2010
187,315 downloads in 2010

No development of DSpace
No general mandatory policy
2010-2011
Rebirth of IRIS

Clarification of strategic positioning
Audit and decision development

Installation of the ORI-OAI system:
• Compliance with French metadata standards
• Interoperability with STAR and HAL

A third platform for the self-deposit of scientific production
A composite repository with two systems:
• ORI-thèses with EDT and learning objects
• IRIS with the collection of history of sciences

Work in progress
The third platform is under construction
IRIS experience Debriefing

- Institutional support and recognition of the project
- Institutional strategy and policy in the domain of OA
- Human resources with sufficient IT and LIS capacities
- Metadata standard(s) and precise bibliographic description of the content
- Software fitting with local needs and IT environment as well as with national infrastructure and standards
- Added value services for legal aspects and usage statistics
- Knowledge of the scientific community’s information
GL increases the content of IR

**Free availability**, dissemination, visibility and referencing

Relatively **high usage**

Increased **security** and long-term accessibility vs a personal web site

Less **problem with copyright issues**

Solution for the **processing, disseminating and archiving of EDT**

Control on research output and content includes **unpublished documents**

**GL in IR** improves **impact** on the web

Evolution from "collection development" to "**content recruitment**"

**Poor bibliographic control**

**No digital curation** of metadata

**Problem of recognition**

Deposit is **time consuming**

**Missing support** from institution

**Bad quality** of self-deposits

**No priority for the evaluation**

**Opposition to mandatory policy**

Lack of commitment from **institution**
What sort of home for grey literature?

Not one home, but many homes
Four scenarios
1. Publishing grey literature

**Strategy**
- Communication and publishing of scientific papers
- Rapid and direct access to full-text

**Goal**
- To increase the impact of particular document
- To reduce the cost and increase the benefits from the dissemination of the institution’s research and teaching outputs

**Key points**
- Self-deposit of full-text and institutional workflows for EDT are essential for content recruitment while mandatory deposit policy or incentives are not.
- Key elements are a high rate of full-text

**Others**
- Other services may be less crucial but would add value to the site (usage statistics, preservation, publishing)
2. Special items container

Strategy
• A container for all kind of materials produced by faculty

Goal
• Availability and visibility of all kind of materials

Key points
• No clear vision on collection and acquisition
• Quality control through validation or labelling is not an issue

Other
• The most promising perspective may be the linking of the deposits to research data
3. Scientific heritage

**Strategy**
- The past and present scientific production, with GL and published documents and other material

**Goal**
- To increase the impact
- To reduce the cost and increase the benefits from the dissemination

**Key point**
- The definition of an acquisition or content recruitment policy
- Digitization of older copyright cleared material
- Metadata indexing policy

**Other**
- The closest scenario to traditional library collection building
- The local presence of a digitization center
4. Institutional deposit

**Strategy**
- The main interest of these repositories is not collection building but evaluation

**Goal**
- To demonstrate the value of the institution itself
- To facilitate control over scientific production and evaluation procedures

**Key point**
- Peer-reviewed publications will play a major role, grey literature will be less valued or appreciated

**Other**
- Include services (usage statistics, research assessment and monitoring, etc)
- Connection to a CRIS
Map of four scenarios for IR with GL
Minimum requirements for the home of GL

Full-text

Conservation

Quality

Metadata

Openness
Thank you for your attention

joachim.schopfel@univ-lille3.fr
isabelle.le-bescond@univ-lille1.fr
helene.prost@inist.fr