



Fifteenth International Conference on Grey Literature
The Grey Audit:
A Field Assessment in Grey Literature

Grey Literature in European Commission Projects



Sara Goggi, Gabriella Pardelli
CNR-ILC
Silvia Giannini, Stefania Biagioni
CNR-ISTI





Introduction

The latest recommendations issued by the European Commission go towards the revision of their policy on dissemination and preservation of scientific information in order to promote the access to the results of the community-funded research



The context

- Websites of European projects as essential vehicle for the diffusion of scientific information
- Need to improve both the websites structures and contents
- EU **Best Practice Guidelines**
- *The Italian CNR and its rule in European projects*

Starting from this context, this survey

- checks the current practices carried out by the partners of European projects
- focuses on the analysis of the documentation/GL literature and its features

Objectives and Methods

- Objective
 - to detect all types of documents visible from projects websites
- CORDIS (*Community and Research and Development Information Service*) directory as starting point
- **The query criteria**
 - ✓ Search for: Projects - *Consiglio Nazionale delle Ricerche*
 - ✓ *Refinements* > Programme: *FP7* > Coordinator Country>*Italy* > Participant Country > *Italy*

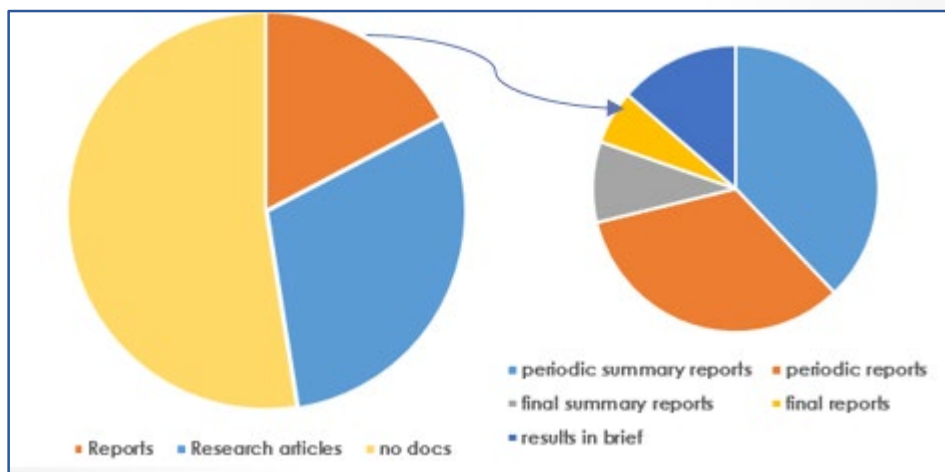
Methods and Material

- 226 projects matching the search criteria
- Data collection between June and August 2013
- Information gathered from:
 - ✓ CORDIS directory contents
 - ✓ Project site structure and contents
 - ✓ Projects production
 - ✓ Projects documentation analysis:
 - a) type of products, b) format, c) availability/access
- Levels
 - explorative (what does CORDIS reveal about projects? does the project have a web site? which kind of documents does it produce? does it produce and present GL?)
 - semantic (categorization)
 - projectual (dataset definition)
 - results analysis

CORDIS websites results

Projects subset		
Status	N.	%
accepted	14	6,2
completed	99	43,8
execution	113	50
Total	226	

Projects subset		
<u>Projects websites</u>	N.	%
<u>yes</u>	140	62
<u>no</u>	86	38
<u>Projects websites by status</u>		
<u>accepted</u>	1	0,8
<u>completed</u>	60	42,8
<u>execution</u>	79	56,4



Results: projects websites analysis

Documents categorization criteria

a) Guidelines sections, b) Logic association

Categories	n. Projects	%	Formats (%)			OA (%)
			pdf	html	other	
<i>Courses material</i>	12	8,5	35,3	37,4	2,3	74
<i>Deliverables & Reports</i>	74	52,8	55,3	10,0	5,6	60,8
<i>eNewsletters</i>	31	22,1	89,8	15,3	0	98
<i>Images</i>	48	34,2	0,08	1,90	98,80	100
<i>Informative material</i>	138	98,5	4,3	95,6	1,3	100
<i>Media press</i>	17	12,1	29,02	50,3	11,2	88,5
<i>Projects meetings material</i>	50	35,7	51,2	44,5	2,5	96,5
<i>Normative documents</i>	2	1,4	57,2	28,5	14,3	100
<i>Patents</i>	1	0,7	na	na	na	na
<i>Research/scientific articles</i>	109	77,9	37,5	6,6	0,1	35,7
<i>Technical documents</i>	96	68,5	7,2	79,8	14,5	100
<i>Tertiary documents</i>	20	14,2	3,1	80	16,3	100
<i>Theses</i>	4	2,8	85,7	0	0	85,7

Guidelines sections

- Overview
- Consortium
- Management struct.
- Case Study
- Deliverables
- Publications
- Events
- Media center
- Glossary

Results: projects websites analysis

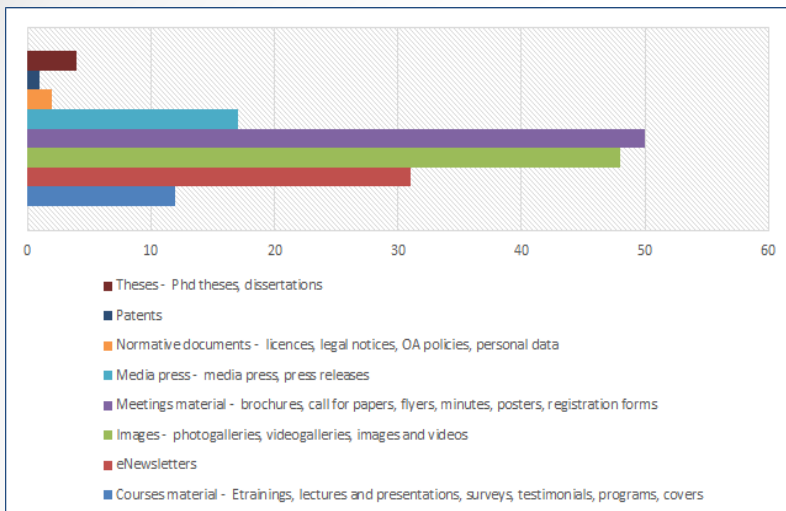
Categories	n. Projects	%
Courses material	12	8,5
Deliverables & Reports	74	52,8
eNewsletters	31	22,1
Images	48	34,2
Informative material	138	98,5
Media press	17	12,1
Projects meetings material	50	35,7
Normative documents	2	1,4
Patents	1	0,7
Research/scientific articles	109	77,9
Technical documents	96	68,5
Tertiary documents	20	14,2
Theses	4	2,8

Categories

Categories	n. Projects	%
Courses material	12	8,5
eNewsletter	31	22,1
Image material	48	34,2
Meetings material	50	35,7
Media press	17	12,1
Normative documents	2	1,4
Patent	1	0,7
Theses	4	2,8

Individual categories

Results: projects websites analysis

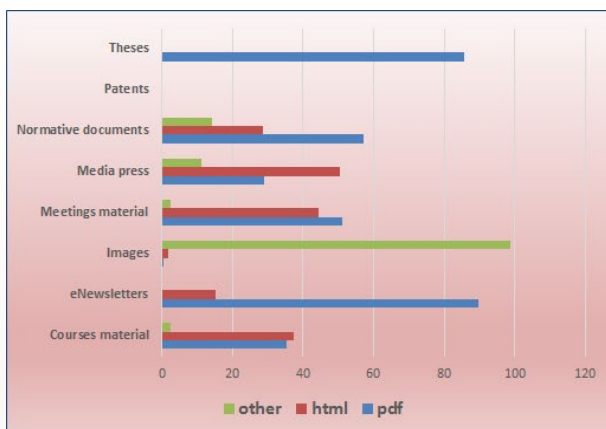


Individual categories (types/subtypes)

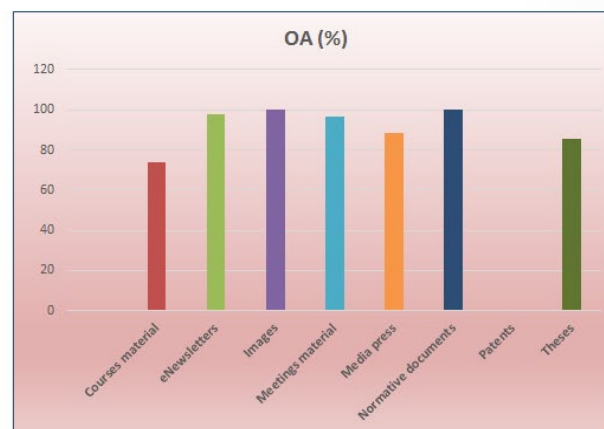
Html and pdf are the prevalent formats for some categories

Conversely:
Images => multimedia material
jpeg, tiff, png...

Almost all types/subtypes can be enjoyed in open access



Individual categories (format and accessibility)



Results analysis: Informative material

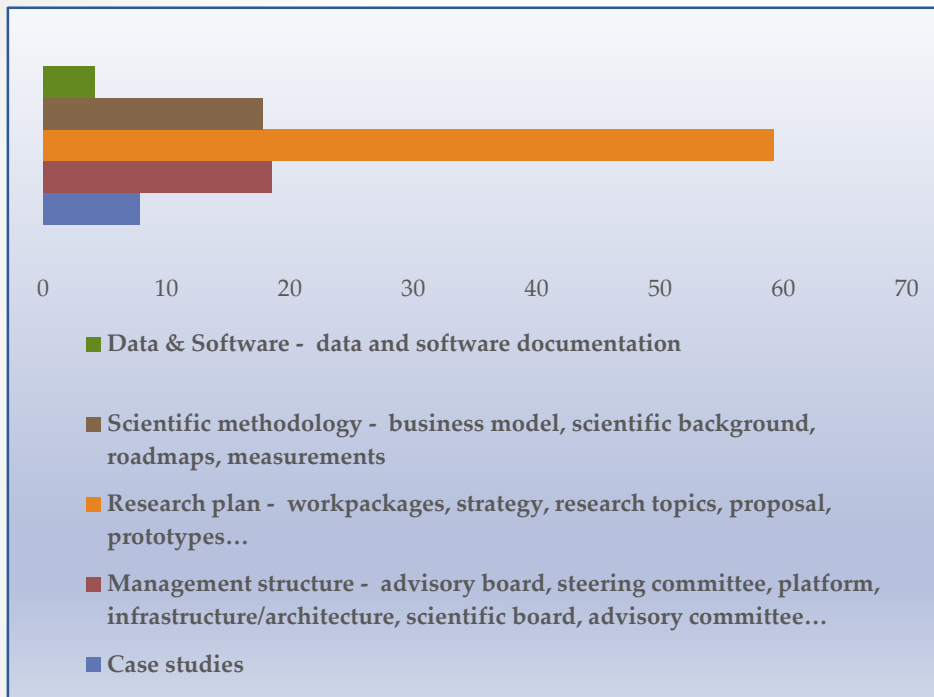


Informative material (types/subtypes)

- *Project overview and Consortium contents = the biggest of the whole category*
- *The information extent varies from project to project*
- *Greater compliance to EU guidelines for Projects overview and Consortium areas*

Related links: how many projects offer the link to similar projects?

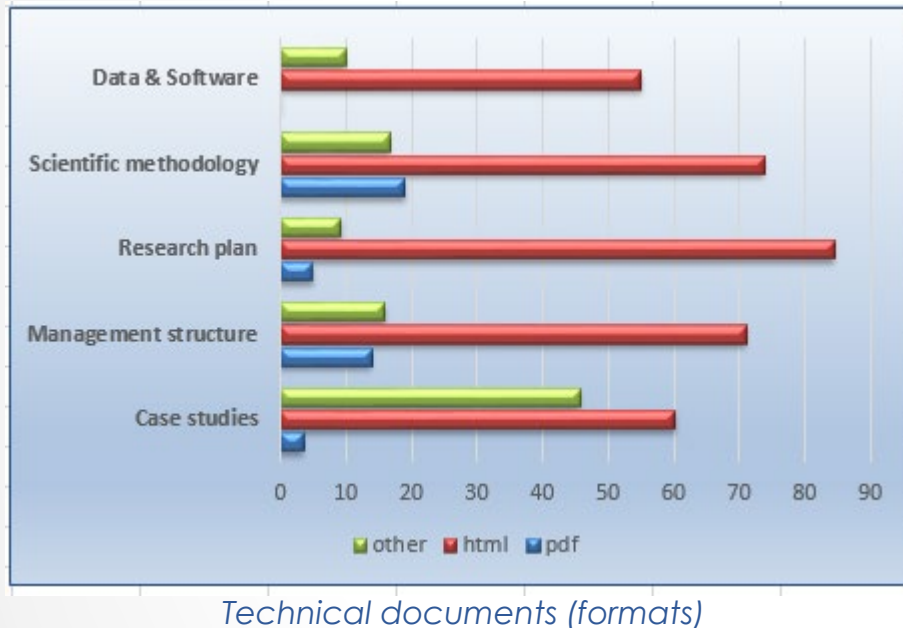
Results analysis: Technical documents



Technical documents (types/subtypes)

- *Scientific methodology and Management structures are the largest categories*
- *47% of projects provides a description of workpackages and/or management structure*
- *23 projects provide the graphical representation of their Research Plan*
- *Case studies closely related to geographic studies*

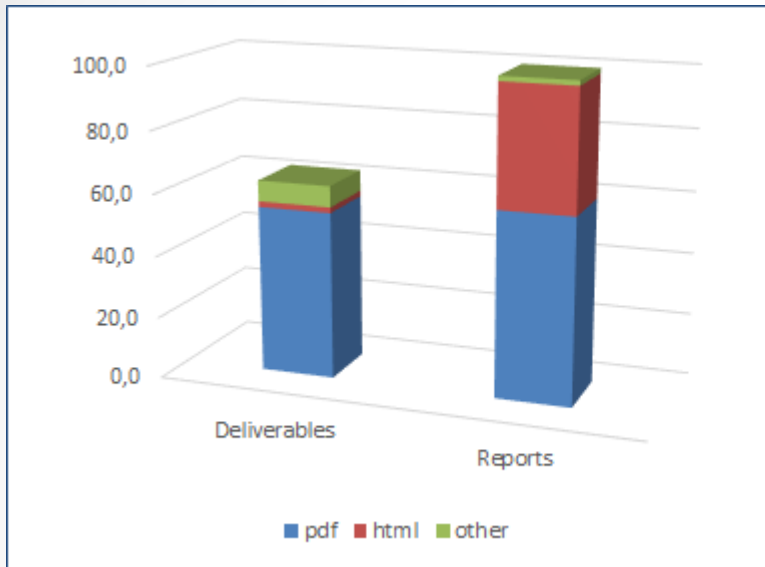
Results analysis: Technical documents



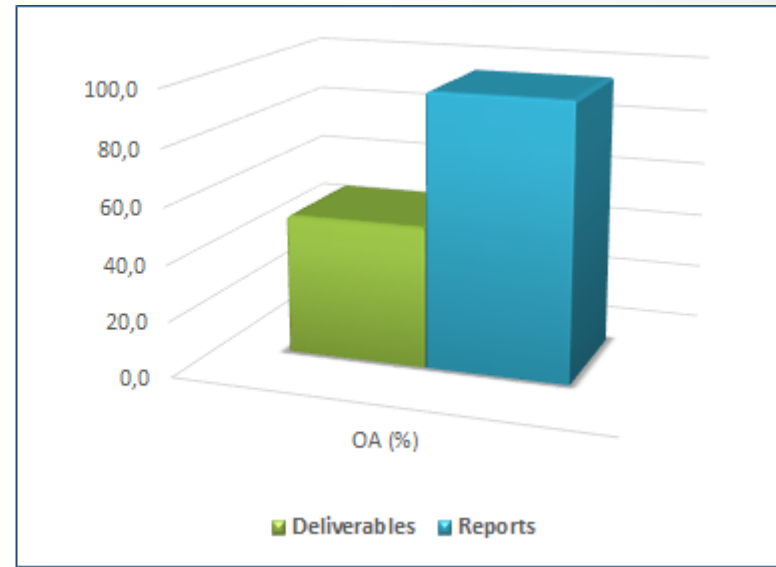
- *Technical documents available through descriptive html pages*
- *Case study => simple links to the studied sites, or to images and photographs of specific subjects*
- *Data & Software don't present any document in pdf format about the developed software*

All documents converged in this category are usable in open access.

Results analysis: Deliverables & Reports



Technical documents (format)



Technical documents (accessibility)

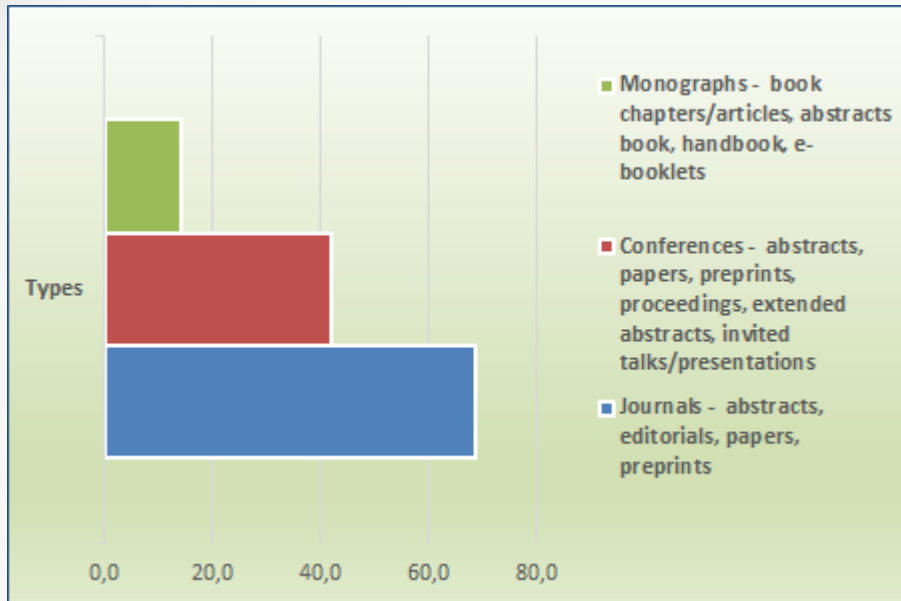
Reports => periodic and final reports / results in brief

54% of the projects presents deliverables

46% doesn't present deliverables?

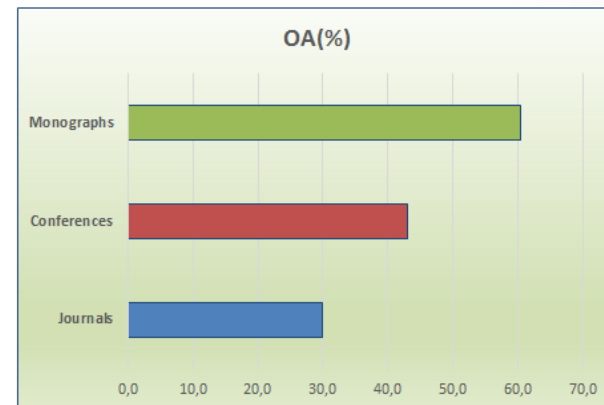
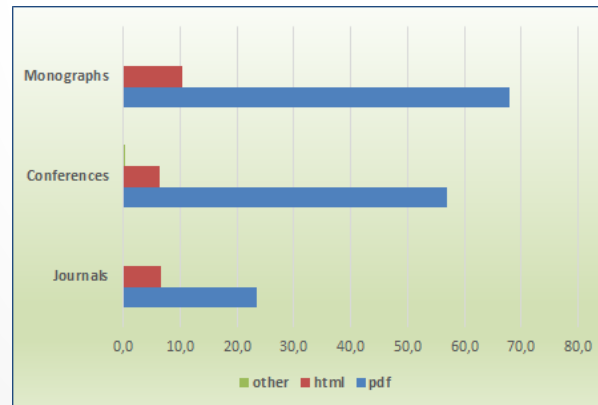
Does restricted area contain deliverables?

Results analysis: Research/scientific articles



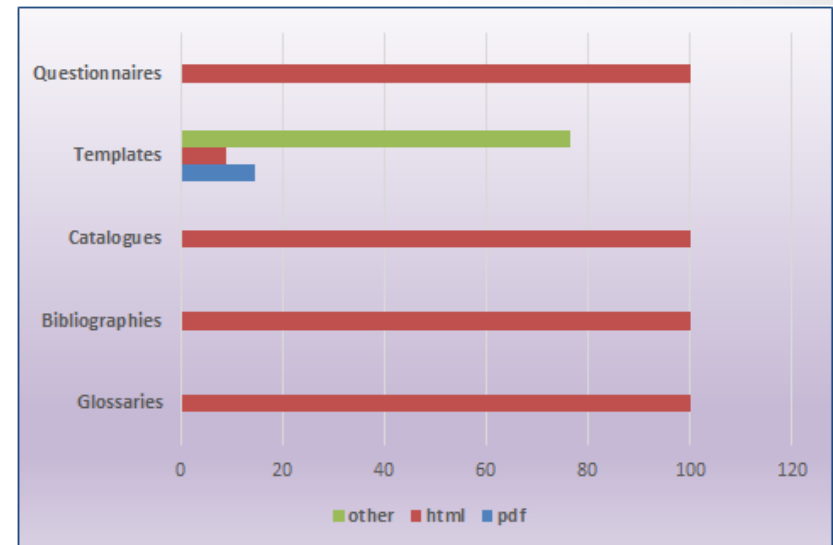
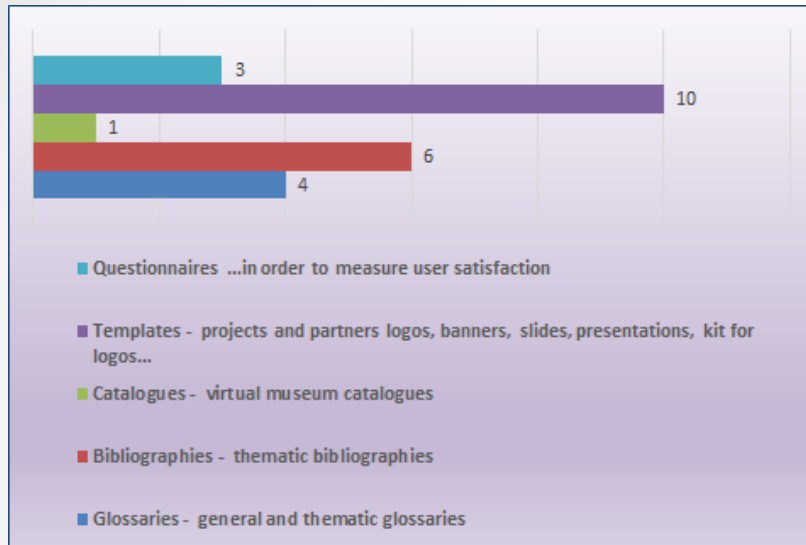
- Highest percentage related to research articles published in journals, followed by publications in conferences and, finally, in monographs.
- More intensive use of pdf and availability of full-text for monographs, followed by Conferences and Journals.

Research/scientific articles (types/subtypes)



Research/scientific articles (format and accessibility)

Results analysis: Tertiary documents



Tertiary documents (types/subtypes/formats)

Tertiary documents as secondary sources... in the context

HTML over all other formats

All documents are fully available in open access

Concluding Remarks



The analysis of projects websites detects:

- a wide heterogeneity on the choice of terms which label the documentation: *documentation, publications, library, results, dissemination* are some of the terms used by the projects to collect and make available their documents
- Grey Literature is the majority of documentation produced by projects but is still lacking a infrastructure which merges, describes and makes it more easily available
- projects websites do not usually bring in the keywords identifying the topics dealt with. Moreover, they do not always show targeted links to related projects
- CORDIS is the only source for establishing links between projects belonging to the same thematic areas but these links are very generic

Future plans

The study of the terminology adopted by the EU within the different frameworks and in projects websites and a deeper analysis of the relationships between projects, websites and research areas could be possible directions for further development and future studies.



Fifteenth International Conference on Grey Literature
**The Grey Audit:
A Field Assessment in Grey Literature**

Thank you!



Sara Goggi, Gabriella Pardelli
CNR-ILC
Silvia Giannini, Stefania Biagioni
CNR-ISTI

