



CERN Scientific Information Service

- The context
- Presentation of the Service
- How do they search and use information?
- The project Inspire: resource integration
- Open Access Publishing: the challenge

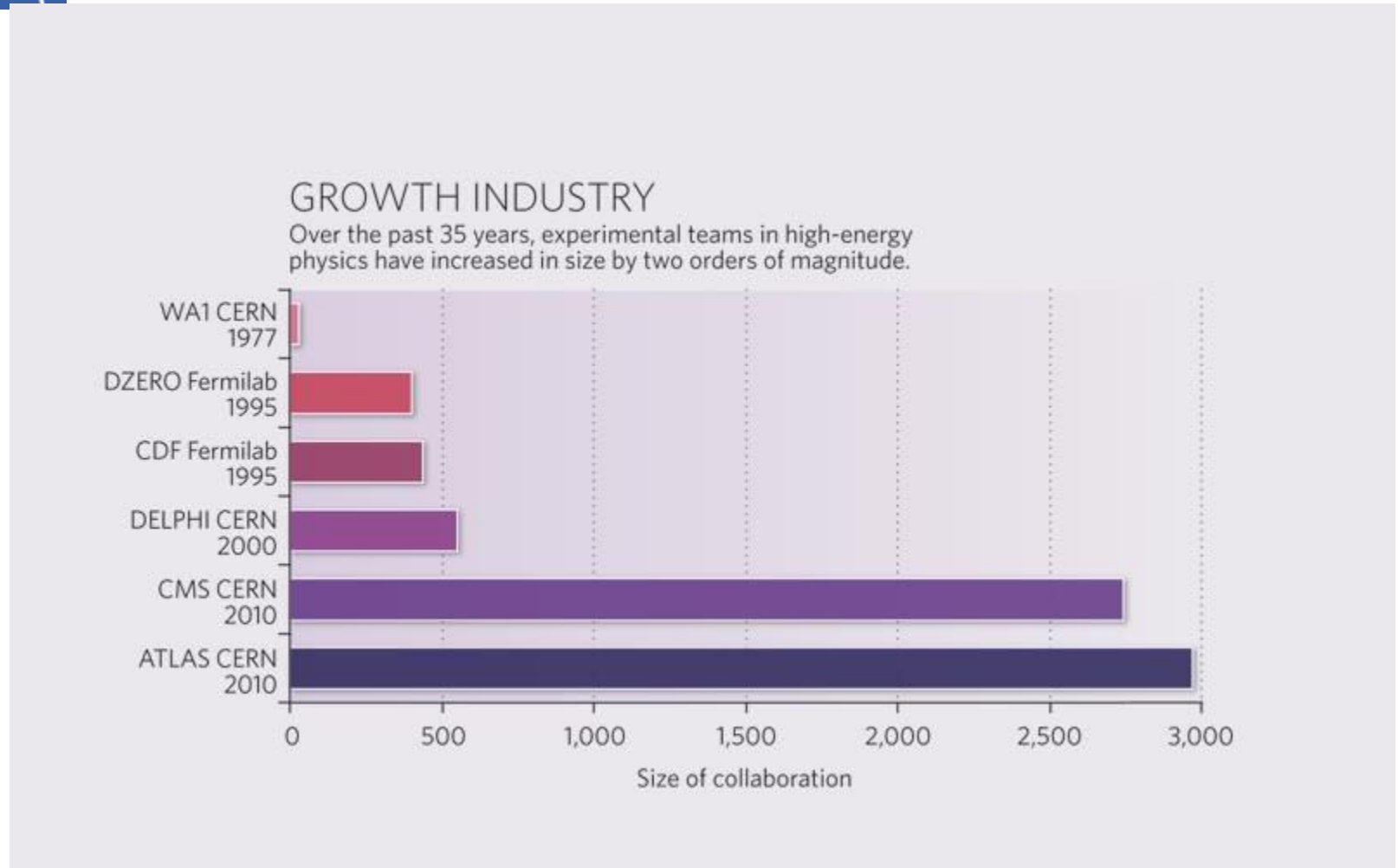


CERN: the context

- About 2,500 staff: physicists, engineers, technicians, administrative support...
- About 8,000 visiting scientists, from 580 institutions and 85 countries
- The collaborations: a huge human experiment. The ATLAS collaboration: 2,900 scientists from 172 institutions and 37 countries
- LHC in operation: first preliminary results published



CERN: the context



Tullio Basaglia, CERN GS-SI

“Physics: The Large Human Collider”, *Nature*, **464**, 482-484 (2010)



CERN: the users, their needs, their expectations

- Geographically dispersed, nomadic community
- Consumer and producer of information coincide to a large extent
- Key role of “grey literature” (reports, conference contributions)
- Heavy usage of informal communication channels: arXiv.org
- Open, immediate access to the article in full text
- Information in electronic format is privileged: e-books, e-journals,
- Commercial online bibliographic databases are hardly ever used, Google plays a role, but it does not dominate



Information Production, Storage and Search @ CERN

– **A wide range of information and documentation is available at CERN and more is produced every day:**

- Scientific documents
- Engineering documents
- Administrative documents
- Web-pages
- Minutes of meetings
- etc.

– **Depending on their nature, they are stored in different document collection systems such as:**

- CERN Document Server (CDS),
- Electronic Document Handling (EDH),
- Engineering Data Management Service (EDMS),
- websites and databases.

Thanks to A. Wagner, presentation at CERN Computing Seminar 23.5.07



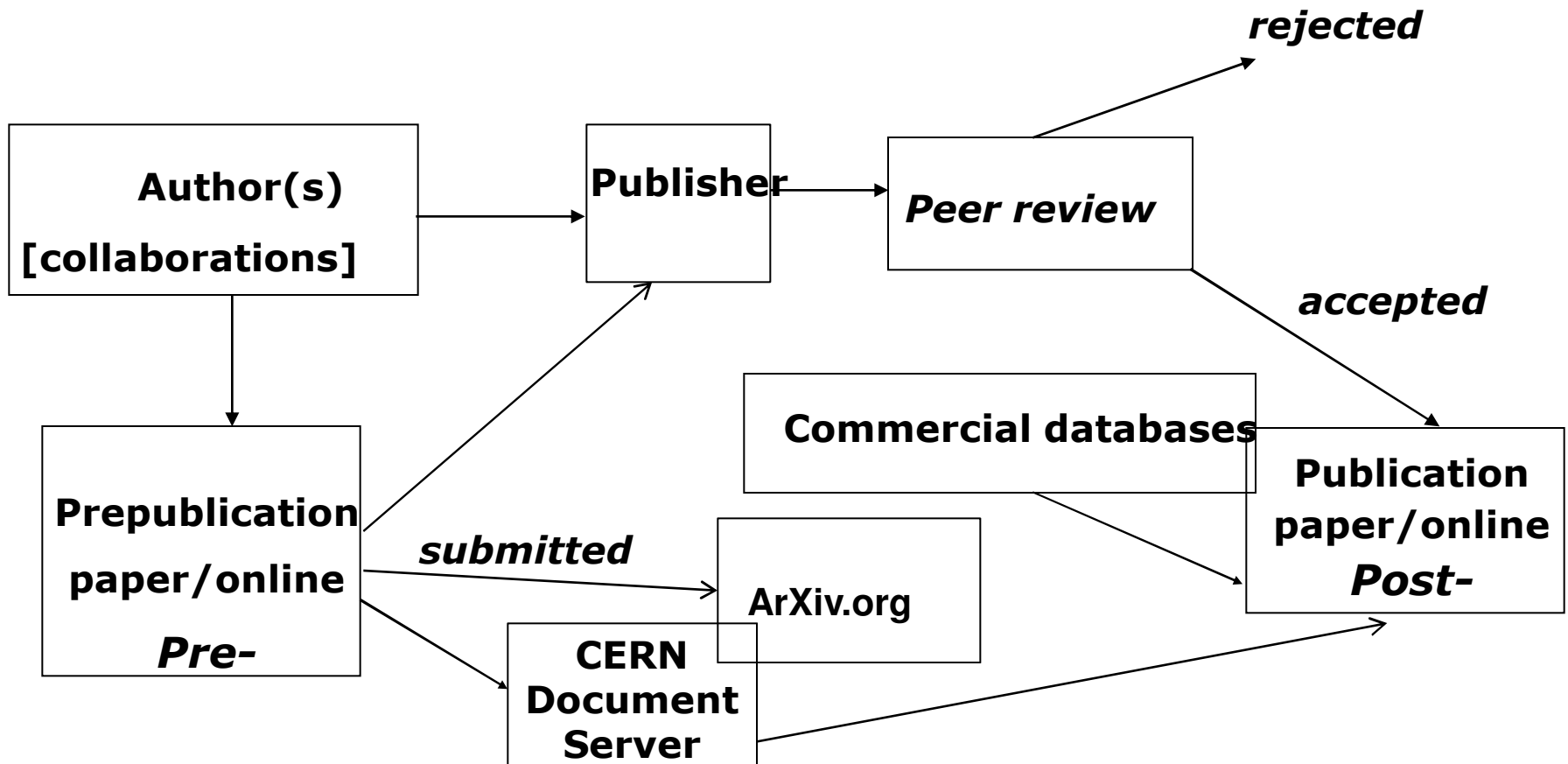
Some Figures

Collection	Number of Documents
Admin documents	~ 2,500,000
CERN Web Intranet	~ 900,000
EDMS	~ 1,000,000
CDS	~ 1,000,000
Agenda/Indico	~120,000
Xfind (knowledgebase)	< 50,000
Total	~ 5 - 6 million documents

Thanks to A. Wagner, presentation at CERN Computing Seminar 23.5.07

Tullio Basaglia, CERN GS-SI

Scientific communication



Scientific communication: formal and informal channels



- “Mass distribution of unedited, unreferred and often unproofread preprints, which has recently been proposed, would put journals out of business or transform them into repositories”

“Is Journal Publication Obsolescent?”, Physics Today, May 1966



The Scientific Information Service

- Library services
- Open Access: the project
- The Archives

The site: <http://library.cern.ch>



Knowledge Preservation and Archiving

- CERN Document Server (CDS): CERN's institutional repository

The screenshot shows a Windows Internet Explorer browser window displaying the CERN Document Server (CDS) homepage. The browser's address bar shows the URL <http://cdsweb.cern.ch/>. The browser's title bar reads "CERN Document Server: Home - Windows Internet Explorer provided by CERN". The browser's menu bar includes "File", "Edit", "View", "Favorites", "Tools", and "Help". The browser's toolbar includes a search icon, a home icon, a print icon, a page icon, and a tools icon. The browser's status bar shows "CERN Document Server: Home".

The CDS homepage features a navigation menu with the following items: [Search](#), [Submit](#), [Convert](#), [Agenda](#), [Webcast](#), [Bulletin](#), and [Library](#). The user is logged in as "guest" and can click on [login](#).

The main content area displays the text: "Over 800,000 bibliographic records, including 360,000 fulltext documents, of interest to people working in particle physics and related areas. Covers preprints, articles, books, journals, photographs, and much more."

Below this text is a search bar with the text "Search 893,941 records for:". The search bar includes a text input field, a dropdown menu set to "any field", a [Search](#) button, and a [Browse](#) button. Below the search bar are links for [Search Tips](#) and [Advanced Search](#).

At the bottom of the page, there are two sections: "Narrow by collection:" and "Focus on:". The "Narrow by collection:" section has a checked checkbox next to [Articles & Preprints](#) (731,008), with sub-links for [Published Articles](#) (284,821) and [Preprints](#) (388,626). The "Focus on:" section has a link for [CERN Articles & Preprints](#) (93,869), with sub-links for [CERN Published Articles](#) (47,862) and [CERN Preprints](#) (14,817).



CERN Document Server (CDS)

- CERN's institutional repository
- ~ **1,000,000** bibliographic records, including more than **400,000** full text documents,
- Covers articles, books, journals, photographs, and much more.
- Powered by CDS Invenio, produced by CDS Software Consortium



CERN Document Server

Software Consortium Products

- **CDS Invenio**: runs electronic preprint servers, online digital library catalogues or document repositories on the web. It complies with the Open Archives Initiative (OAI) and uses MARC21 as its underlying bibliographic standard.
- **CDS Indico**: allows to schedule conferences, from single talks to complex meetings with sessions and contributions. It also includes an advanced user delegation mechanism, allow paper reviewing, archival of conference information and electronic proceedings.
- Both are available under the GNU General Public License.



The Scientific Information Service

- <http://cdsweb.cern.ch> : using the basket functionality one can create, modify and share basket(s). Users can request the loan of documents from their desktop.
- Online and on-shelf access to literature: books and [e-books](#) (~10,000 titles) , journals and e-journals (~1,500 titles)



The Scientific Information Service

- More than a book repository: the standards' collection. Purchased on request, obtained generally in PDF format, accessible from the CERN Intranet.
- Looking for information on standards? The [Perinorm](#) database is available (username/password required)
- The Bookshop in the Central Library is at disposal of anyone at CERN. A stock of selected titles in Physics, Mathematics, Engineering, IT is available.
- If a reader is looking for an item not in stock at the Bookshop (also out-of-print books), we can purchase it for her/him. Contact email is Bookshop@cern.ch



The Scientific Information Service

- What if a reader cannot find what s/he is looking for? The Interlibrary Loan service can help:
<http://library.web.cern.ch/library/Library/ill.html>
 - Free of charge (standards and dissertations excepted)
 - quick and reliable
 - for articles, possibility to order independently (no intermediation)
 - high success rate
- Looking for a bibliography on a specific topic? Use [online databases](#)
- Questions, problems, feedback: library.desk@cern.ch
- We listen to user needs: they can propose a purchase for the library collection



The Scientific Information Service

- We rely on readers ...
- All libraries are open 24/24 h, 7/7, 365/365. We operate openly but on trust
- The desk of the Central Library is staffed: Monday-Friday, 8h30-19h00
- Sharing information
 - Authors should submit their documents (reports / conference papers / articles etc) to CERN Document Server
 - Donating books
- Suggest purchases for the Library collection
- Helping us shaping Library collections that meet users ' needs

Tullio Basaglia, CERN GS-SI



The Scientific Information Service

- According to O.C. no.6 , authors should submit their documents to CERN Document Server.
- Self-archiving articles (preprints/reports/conference papers etc) is possible in CDS: <http://cdsweb.cern.ch/> , using the submit button

Questions?



Open Access Publishing I the Present Landscape

- Investigation done on 5016 articles submitted to arXiv:hep in 2005 and published in peer-reviewed journals:
- **90%** of articles are in theory and by less than 3 authors
- **83%** of articles published in 6 leading journals
- **87%** of articles published by four publishers
- **57%** of articles by not-for-profit (not-for-loss) publishers
- S.Mele *et al.* JHEP 12(2006)S01 arXiv:cs.DL/0611130

Thanks to S. Mele, presentation on the SCOAP3 Model, 14.3.07



Open Access Publishing II

the Future: SCOAP3

- **SCOAP3 = Sponsoring Consortium for Open Access Publishing in Particle Physics. Its aims:**
- Establish Open Access in HEP publishing in a transparent way for authors.
- Convert existing high-quality peer-reviewed journals to Open Access, in a sustainable way.
- Operate along the blueprint of large collaborations.
- Generate savings through negotiating power, author awareness and competition among journals.
- Price tag of 10M€ /year to be shared according to the distribution of HEP articles per country.
- The model has high potential but is only viable if every country contributing to HEP is on board!
- SCOAP 3 model could be rapidly generalized to related fields: Nuclear and Astroparticle Physics

Thanks to S. Mele, presentation on the SCOAP3 Model, 14.3.07



Integrating and Building Knowledge Repositories - I

- **PPA (=Particle and Particle Astrophysics) Information Resource Summit took place at SLAC, Stanford, CA between 10 and 12 May 2007. Participants:**
 - SLAC (Stanford, CA, USA)
 - ArXiv
 - DESY (Hamburg, D)
 - CERN
 - Publishers (Springer, APS, Elsevier, IOP)



Integrating and Building Knowledge Repositories - II

- **Aims of a High Energy Physics Integrated Information System:**
 - Integrating the content of different HEP repositories and databases (metadata and full-text)
 - Develop advanced functionalities: automatic keywording, full-text searching, data mining, new metrics of impact
 - Develop Web 2.0 community applications: tools for co-authorship, commenting, alternative peer-review system