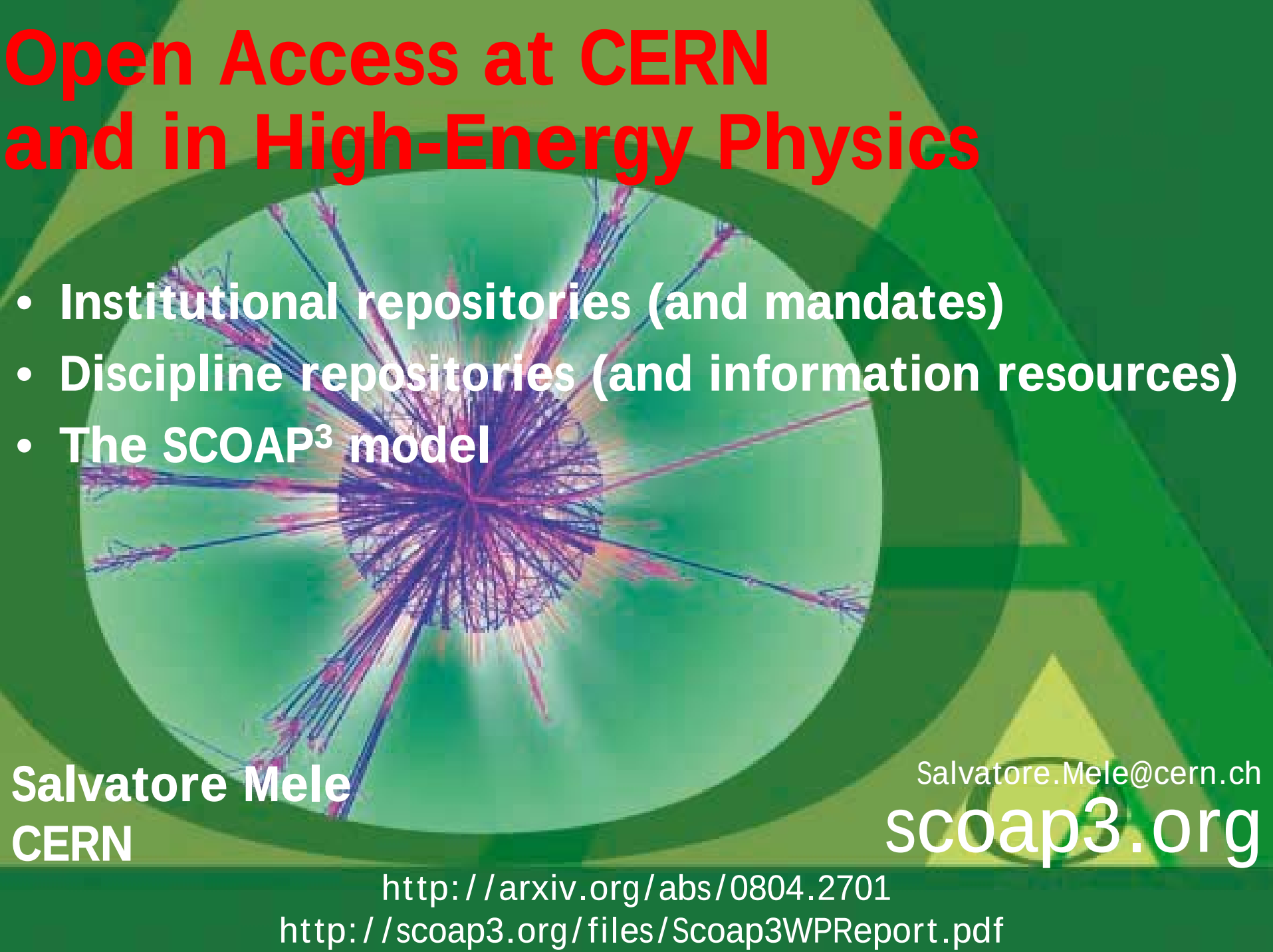
Open Access at CERN and in High-Energy Physics



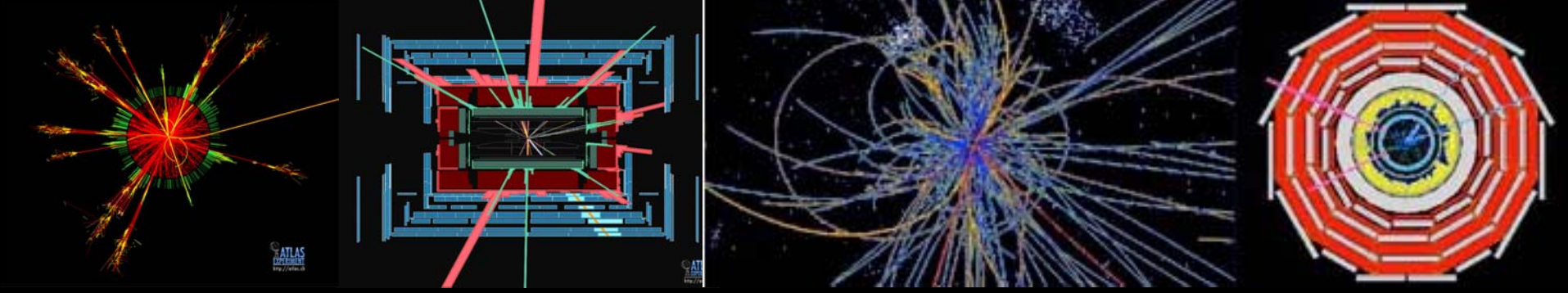
- Institutional repositories (and mandates)
- Discipline repositories (and information resources)
- The SCOAP^3 model

Salvatore Mele
CERN

Salvatore.Mele@cern.ch

scoap3.org

http://arxiv.org/abs/0804.2701
http://scoap3.org/files/Scoap3WPReport.pdf

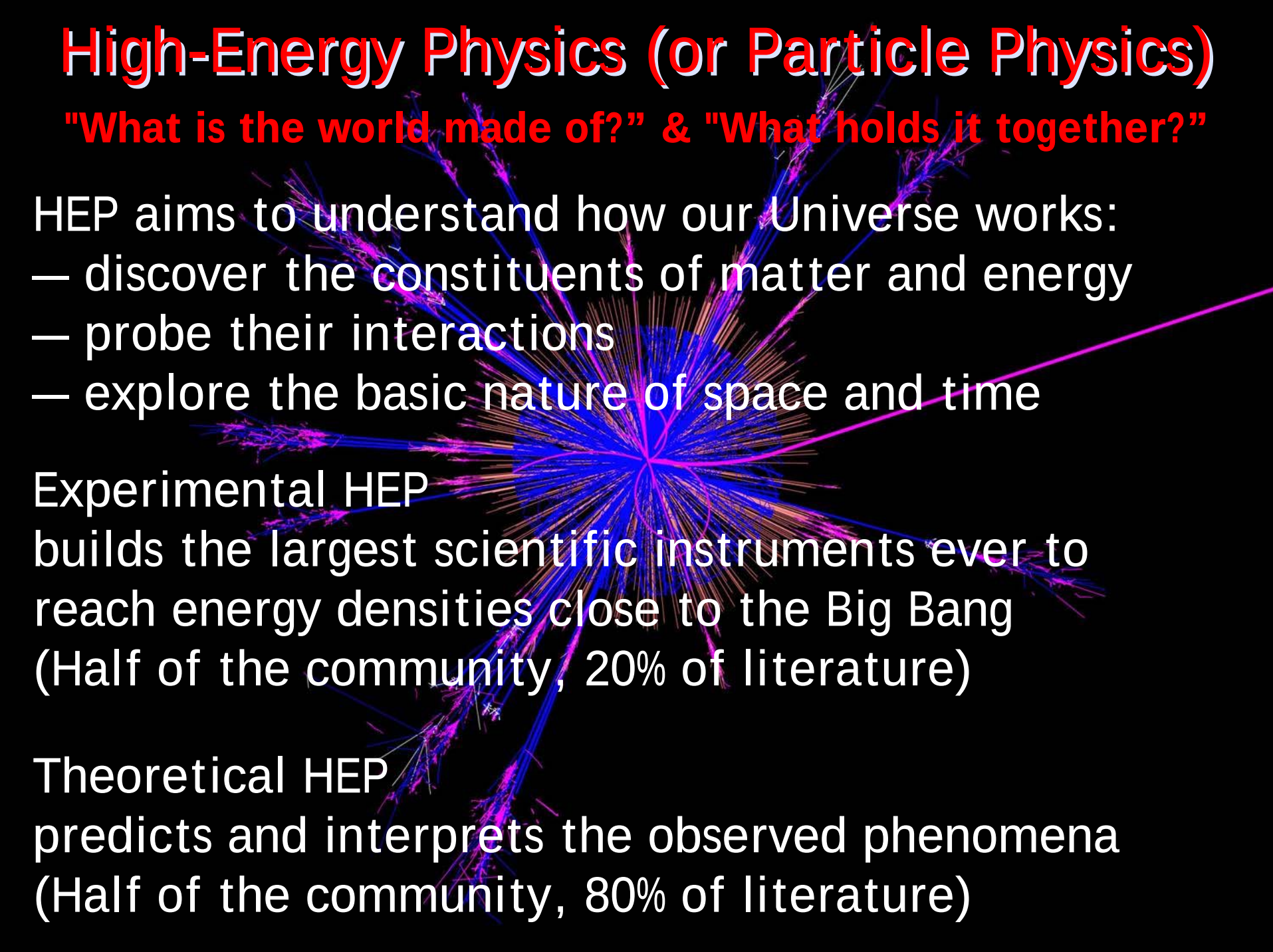


Open Access:

Grant anybody, anywhere and anytime access to the (peer-reviewed) results of (publicly-funded) research

High-Energy Physics (or Particle Physics)

"What is the world made of?" & "What holds it together?"



HEP aims to understand how our Universe works:
— discover the constituents of matter and energy
— probe their interactions
— explore the basic nature of space and time

Experimental HEP
builds the largest scientific instruments ever to reach energy densities close to the Big Bang
(Half of the community, 20% of literature)

Theoretical HEP
predicts and interprets the observed phenomena
(Half of the community, 80% of literature)

CERN: European Organization for Nuclear Research (since 1954)

- The world leading HEP laboratory, Geneva (CH)
- 2500 staff (mostly engineers)
- 9000 users (mostly physicists)
- 3 Nobel prizes (Accelerators, Detectors, Discoveries)
- Invented the web    
- Commissioning the 27-km (6000 M€) LHC accelerator
- Runs a 1-million objects Digital Library

The CERN Convention (1953) contains what is effectively an early Open Access manifesto:

"… the results of its experimental and theoretical work shall be published or otherwise made generally available"

The HEP "preprint culture"

L.Goldschmidt-Clermont, 1965, http://eprints.rclis.org/archive/00000445/02/communication_patterns.pdf
L. Addis, 2002, http://www.slac.stanford.edu/spires/papers/history.html

- In the '60s, golden era of weekly HEP discoveries, HEP scientists could not wait ~1 year for their articles to reach their peers through journals
- *Preprints* as main vehicle of information in HEP
- Researchers (of affluent institutions) mass-mailed preprints to hundreds of (prestigious and therefore affluent) institutions
- *Ante-litteram* (author-pays) Open Access
- HEP libraries classified preprints received worldwide

Enter revolution #1: '70s, IT starts to meet libraries

- SPIRES (1974): e-catalogue of preprint and publications

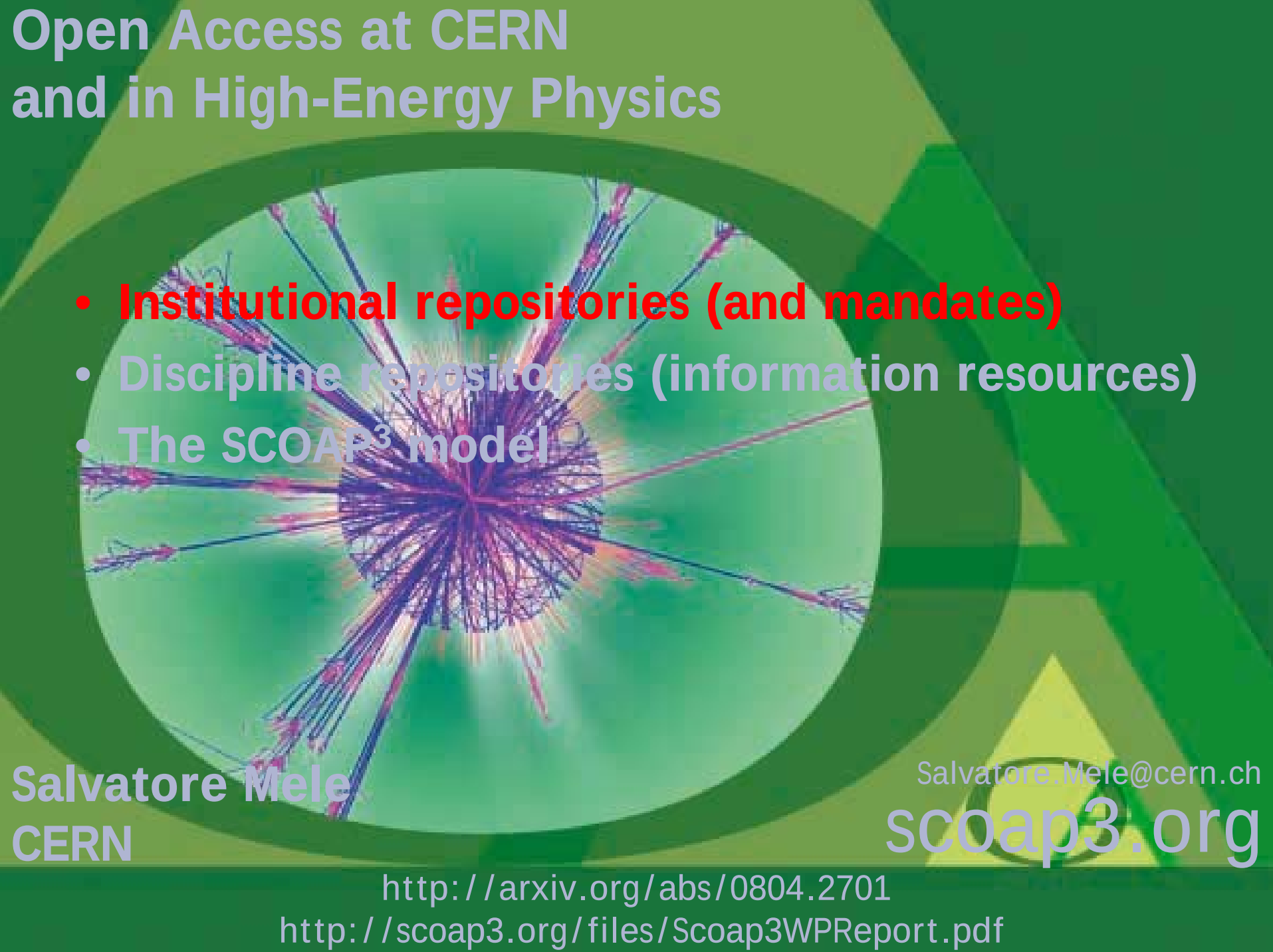
Enter revolution #2: '91, arXiv by Paul Ginsparg at LANL

- HEP preprints and the Internet indissolubly linked

Enter revolution #3: '91, the web by Tim Berners-Lee at CERN

- First U.S. WWW server at SLAC in '91 to access SPIRES
- Summer 1992, SPIRES links to the full-text documents on arXiv

Open Access at CERN and in High-Energy Physics



- **Institutional repositories (and mandates)**
- Discipline repositories (information resources)
- The SCOAP^3 model

Salvatore Mele
CERN

Salvatore.Mele@cern.ch

scoap3.org

http://arxiv.org/abs/0804.2701
http://scoap3.org/files/Scoap3WPReport.pdf

CERN Institutional Repository (deposit is mandated)

Author efforts

A quick reality check on the annual production:

1. ~250 theoretical papers, we capture 0% (!!!)
2. ~50 experimental papers, we capture 90% (!!!)
3. ~500 theses, we capture 10% (world average …)

Library efforts

Strategies to fill the repositories

1. Import from arXiv: 100% coverage for theory
2. Tag CERN authors in publishers feeds, contact research groups, agreements with publishers'
3. targeted e-mails to authors brings the coverage up to 30%, even for theses of 10 years ago

Special projects

98% of all 12'000 papers in theoretical HEP in 55 years of CERN history are now Open Access!

Dig in the archives, lots of scanning, write to authors, seek publishers permission, liaise with sister laboratories with similar holdings, …

Lessons learned

Mandating and advocacy have limits

- "Top-scientists" tend to ignore both "mandating" and "mandated" librarians
- Authors need to see an immediate return from their time invested to submit their work to repositories:
  - the secret of the success of discipline repositories, *i.e.* arXiv (communication, immediacy, visibility, standing)
  - dissemination and preservation of theses
  - opportunity to "bring out of the closet" unpublished material
- We lose OA content from HEP-sub communities where arXiv is less important (instrumentation, accelerators, computer science
- Always remember: discovery, engagement, curation are person-power intensive.

Open Access at CERN and in High-Energy Physics

- Institutional repositories (and mandates)
- Discipline repositories (information resources)
- The SCOAP^3 model

Salvatore Mele
CERN

Salvatore.Mele@cern.ch

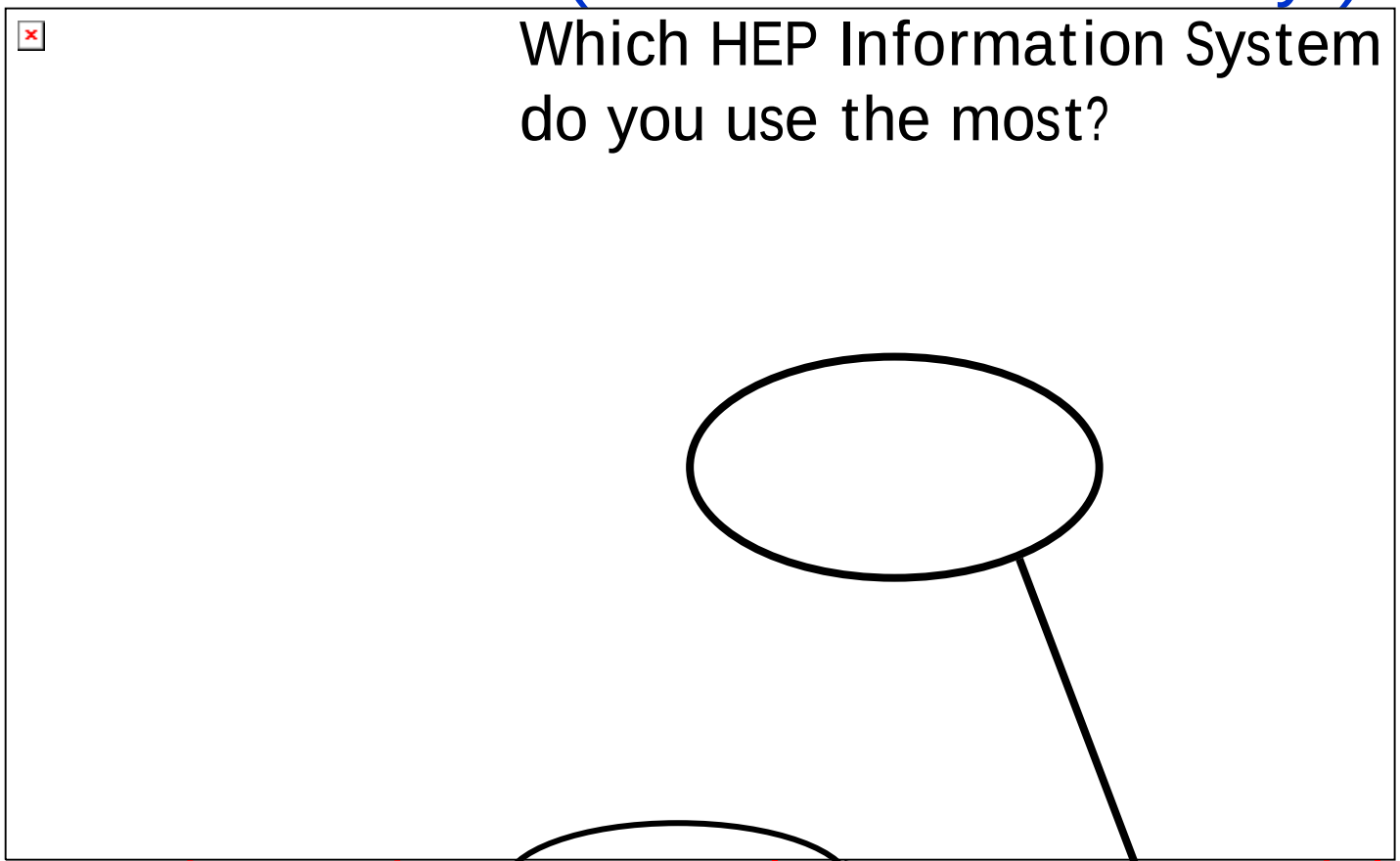
scoap3.org

http://arxiv.org/abs/0804.2701
http://scoap3.org/files/Scoap3WPReport.pdf

Information discovery in HEP

A poll of the HEP community

>2000 answers (10% of the community!)



91 % Community services

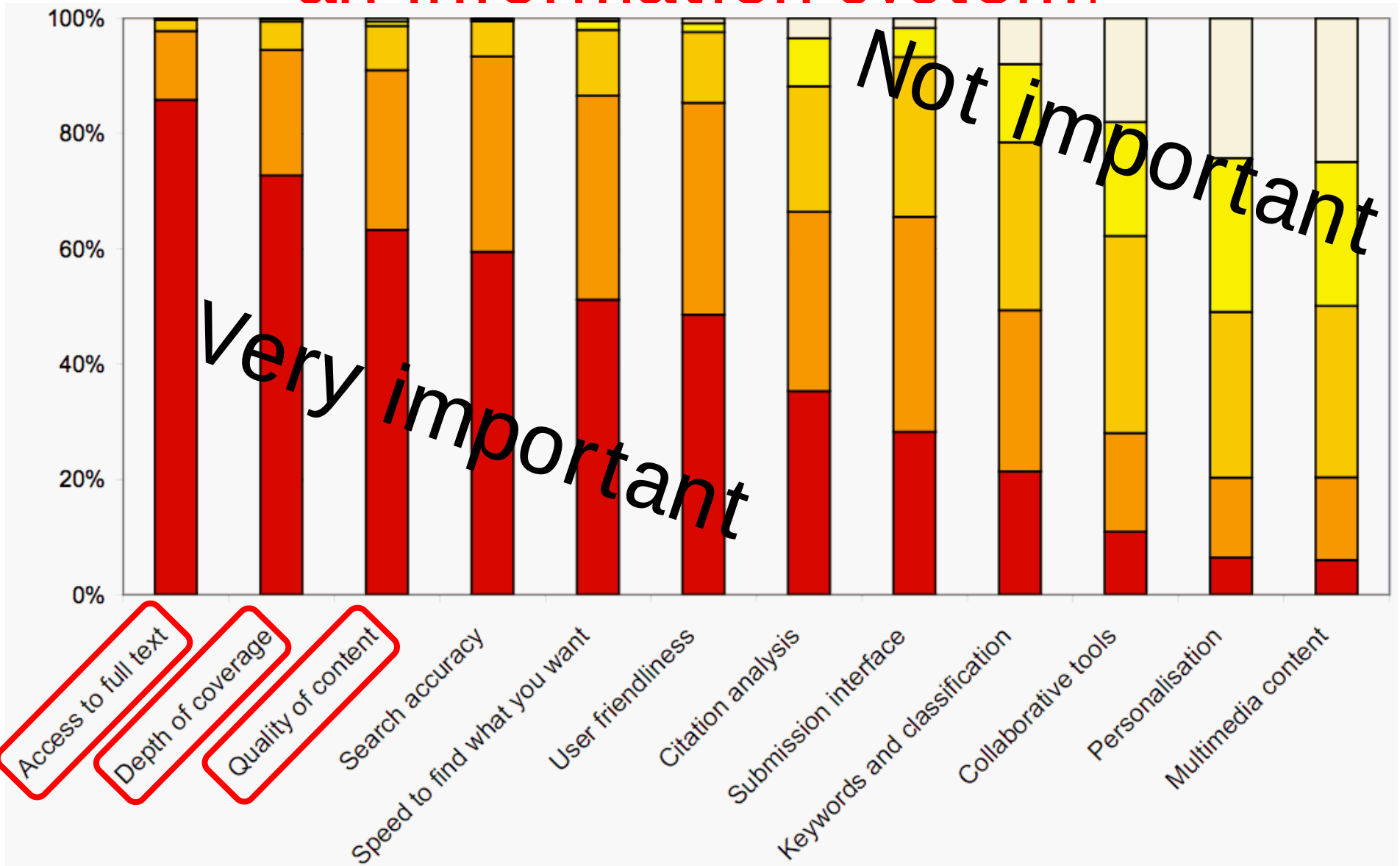
- 40 % Subject repositories
- 51 % Lab-supported databases

9% Google

<0.1% Commercial services

6% for scholars > 6 career years
22% for scholars < 2 career years

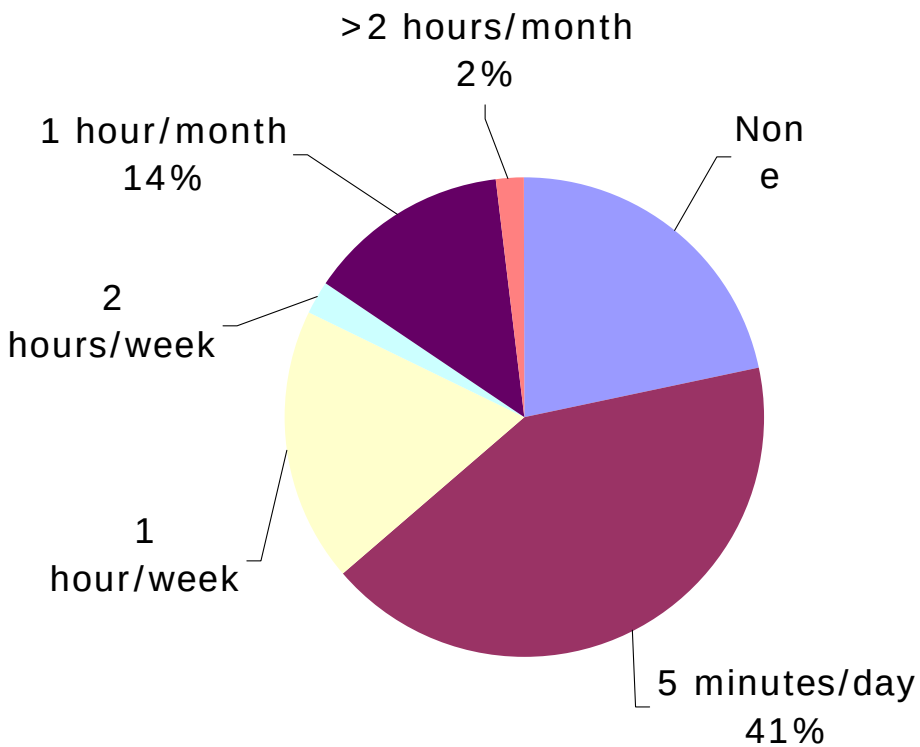
How important are these features of an information system?



Which changes do you expect?
Summary of recurrent and inspiring answers

- Seamless (open) access to older articles
- Improved (full-text search and) access to public experiment notes (grey literature)
- Indexing of conference .ppt slides (interlinked with the corresponding article)
- "Publication" of "ancillary" material:
  - Data in tables & figures; correlation matrices
  - Data (high-level objects)
- (A new kind of) Peer-reviewing overlaid on arXiv
- "Smarter" search tools (related papers)
- Fragments of computer code accompanying equations

Would users invest time in online community service (here content tagging)?



Immense potential to be harnessed

Vision for an integrated digital library for HEP scientific communication

- Keep the principle of user-pulled more than technology-pushed features
- Plug into a complete dataset (40 years of metadata, 17 years of articles)
- Not demonstrators but production systems with large user community

1. Enable text- and data-mining applications
   - recommendation systems, automatic curation, noise-reduction,…
   - search tools for scientific plots
   - open-access database, as a sandbox for third-party applications
2. Explore citation analysis
   - “cited with”, author networks, new tools for information discovery
   - new impact definitions, hybrid metrics
3. Implement Web2.0
   - harness user-tagged content
   - explore new communication channels in integrated workflow
4. Integrate data
   1. mine back data from tables and images
   2. platform to publish beyond paper: multi-dimensional tables, higher level data

Open Access at CERN and in High-Energy Physics

- Institutional repositories (and mandates)
- Discipline repositories (information resources)
- The SCOAP^3 model

Salvatore Mele
CERN

Salvatore.Mele@cern.ch

scoap3.org

http://arxiv.org/abs/0804.2701
http://scoap3.org/files/Scoap3WPReport.pdf

HEP and Open Access

QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

**After preprints, arXiv and the web,
Open Access journals
are the natural evolution of
HEP scholarly communication**

Is it not about vocal librarians!

"We, strongly encourage the usage of electronic publishing methods for our publications and support the principles of Open Access Publishing, which includes granting free access of our publications to all. Furthermore, we encourage all our members to publish papers in easily accessible journals, following the principles of the Open Access Paradigm."

4 experimental groups
5000 scientists
from 54 countries

ATLAS; approved on 23rd February 2007
CMS; approved on 2nd March 2007
ALICE; approved on 9th March 2007
LHCb; approved on 12th March 2007

HEP and its journals

- Journals are on the way to lose (lost?) a century-old role as vehicles of scholarly communication.
- Still, evaluation of institutes and (young) researchers is based on high-quality peer-reviewed journals.
- The main role of journals is to assure high-quality peer-review and act as keepers-of-the-records
- The HEP community needs high-quality journals, our "interface with officialdom"
- Implicitly, the HEP community supports this role by purchasing subscriptions, as it reads off arXiv anyhow
- Subscription prices make the model unsustainable
- As an "all-arXiv discipline" HEP is at high risk to see its journal canceled by large multidisciplinary university libraries (when not already happened)

Open Access experiments

Sponsoring model: institutions funds. No author charges. All content free to read. Successful niche journal (200 art./y), 15 sponsors

Springer Open Choice

APS physics FREE TO READ

SPONSORED ARTICLE
ELSEVIER

Hybrid model: authors can pay journals to make their articles OA. The rest of the journal is under subscriptions. Subscriptions reduced according to the fraction of OA articles.

- Springer in 2004, followed by APS and Elsevier
- Prices range from 975$ to 3,000$
- Little, if any, success: in competition with research funds

Open Access experiments

DEUTSCHE PHYSIKALISCHE GESELLSCHAFT | IOP Institute of Physics

New Journal of Physics

The open-access journal for physics

PMC Physics A

Author-pays: all content of the journal is free to read. After acceptance, authors pay journals for processing fees

- Successful in Life Sciences (BioMedCentral). However sustainability problems are arising (subscriptions to other journals still there & library budgets fixed or reducing)
- *New Journal of Physics* (IOP) since 1998, but attracts only a small HEP fraction, with 20 articles/year
- PhysMathCentral *Physics A*, a new HEP journal
- Model in its infancy in HEP. FAQ: why pay something (peer-review) you can get for free (the library pays subscriptions)

Other Open Access experiments

Institutional membership: for a (small) fee in addition to subscriptions, JHEP and JINST publish OA all articles with at least one author from the institution.

- SLAC, Fermilab, DESY, CERN, and the entire France trying this scheme.

"EPJC is willing to negotiate with funding agencies interested in Open Access to become fully Open Access. In anticipation of such successful negotiations, all experimental papers accepted by The European Physical Journal C will be published Open Access without any fees"

The SCOAP^3 model

Sponsoring Consortium for Open Access Publishing in Particle Physics

A practical approach:

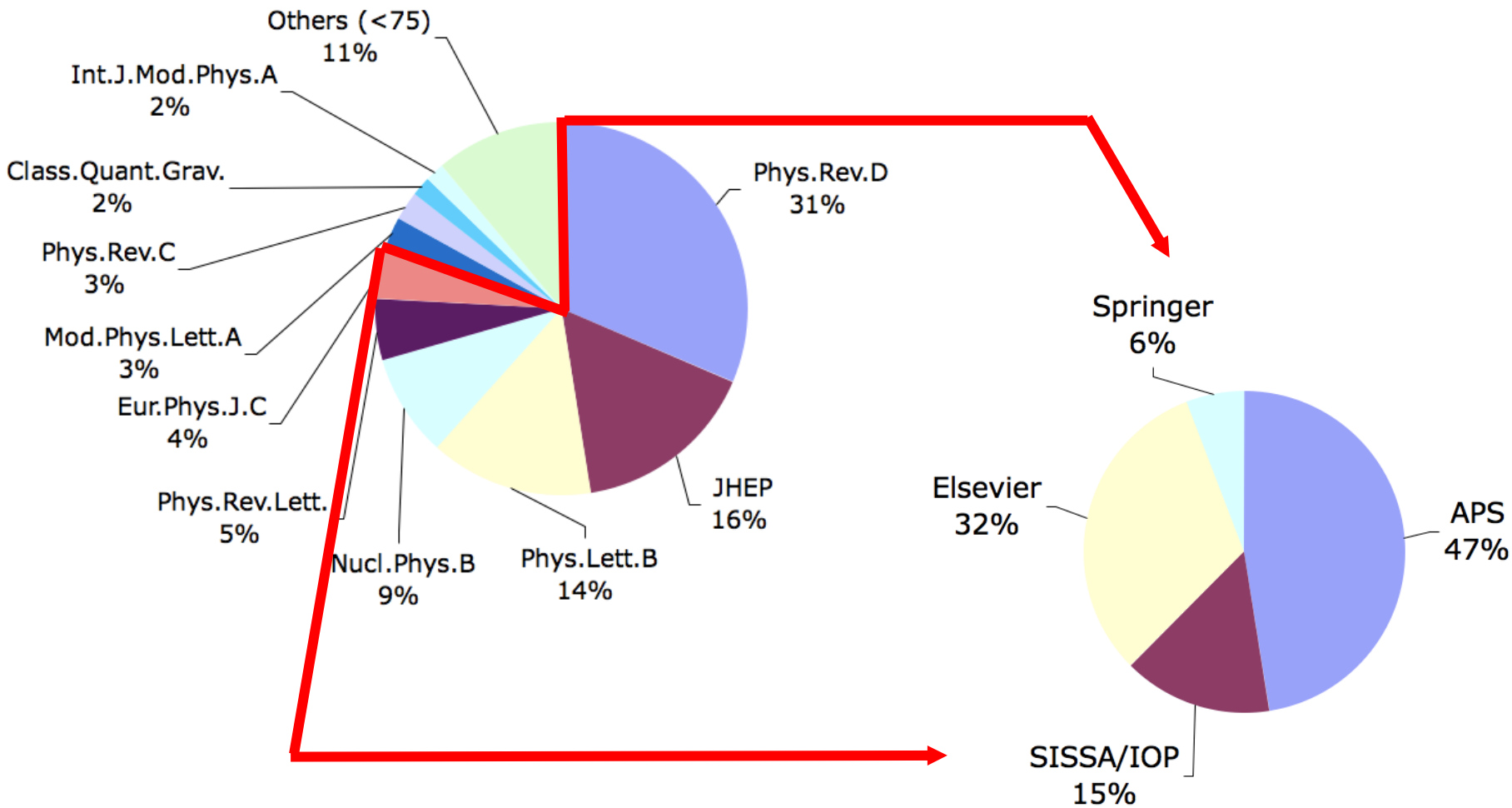
How to publish OA about 5'000 articles/year, produced by a community of about 20'000 scientists?

http://scoap3.org/files/Scoap3ExecutiveSummary.pdf

http://scoap3.org/files/Scoap3WPReport.pdf

The HEP publishing landscape

Source: SPIRES



90% of articles are in theory

80% of articles published in 6 leading journals by 4 publishers

62% of articles by not-for-profit (nor-for-loss) publishers

SCOAP^3

Sponsoring Consoritum for Open Access Publishing in Particle Physics

A consortium sponsors HEP publications and makes them OA by re-directing subscription money.

Today: (funding bodies through) libraries buy journal subscriptions to support the peer-review service and to allow their users to read articles.

Tomorrow: funding bodies and libraries contribute to the SCOAP^3 consortium, which pays centrally for the organization of the peer-review and other editorial services. Articles are free to read for everyone.

- Five "core" journals: PRD, JHEP, PLB, NPB, EPJC
  - Carry a majority of HEP content: aim to convert entirely to Open Access
- Two "broadband" journal: PRL, NIM
  - 10% & 25% HEP: conversion to OA of this fraction

SCOAP^3 is not limited to this initial set of journals but open to all high-quality HEP journals!

Guesstimating the budget envelope

(data and exchange rate of April '07)

- *Physical Review D* (APS) operates on revenues of 2.7M€/year (31% of arXiv:hep)
- *Journal of High Energy Physics* (SISSA/IOP) needs ~1M€/year (19% of arXiv:hep)

HEP Open Access price tag: 10M€/year

- A published PRD article costs APS ~1500€
- 6-8 leading journals publish 5000-7000 art./year

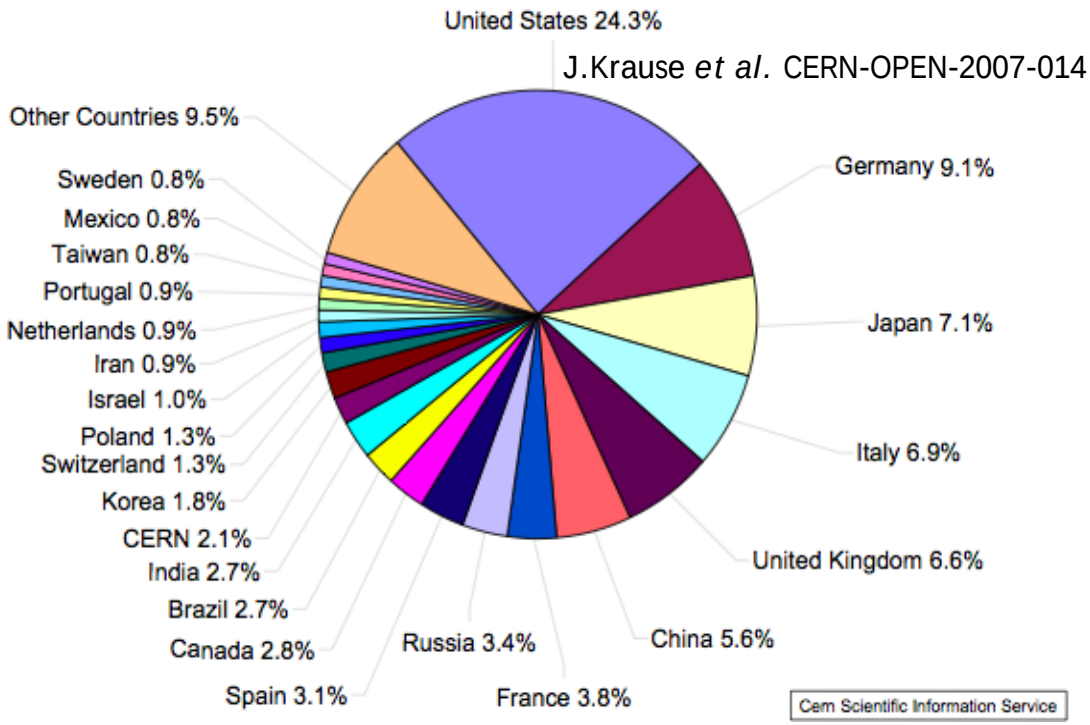
SCOAP^3 exact yearly cost to be known after a tender is sent to publishers asking for the per-article cost of their peer-review and other editorial services.

Outcome of the tender: selection of high-quality/best-price offers. Aim to conversion of entire literature

SCOAP³ financing

SCOAP³ to be funded according to a "fair-share" model based on the fraction of HEP articles per country: the more a country uses the system the larger its share.

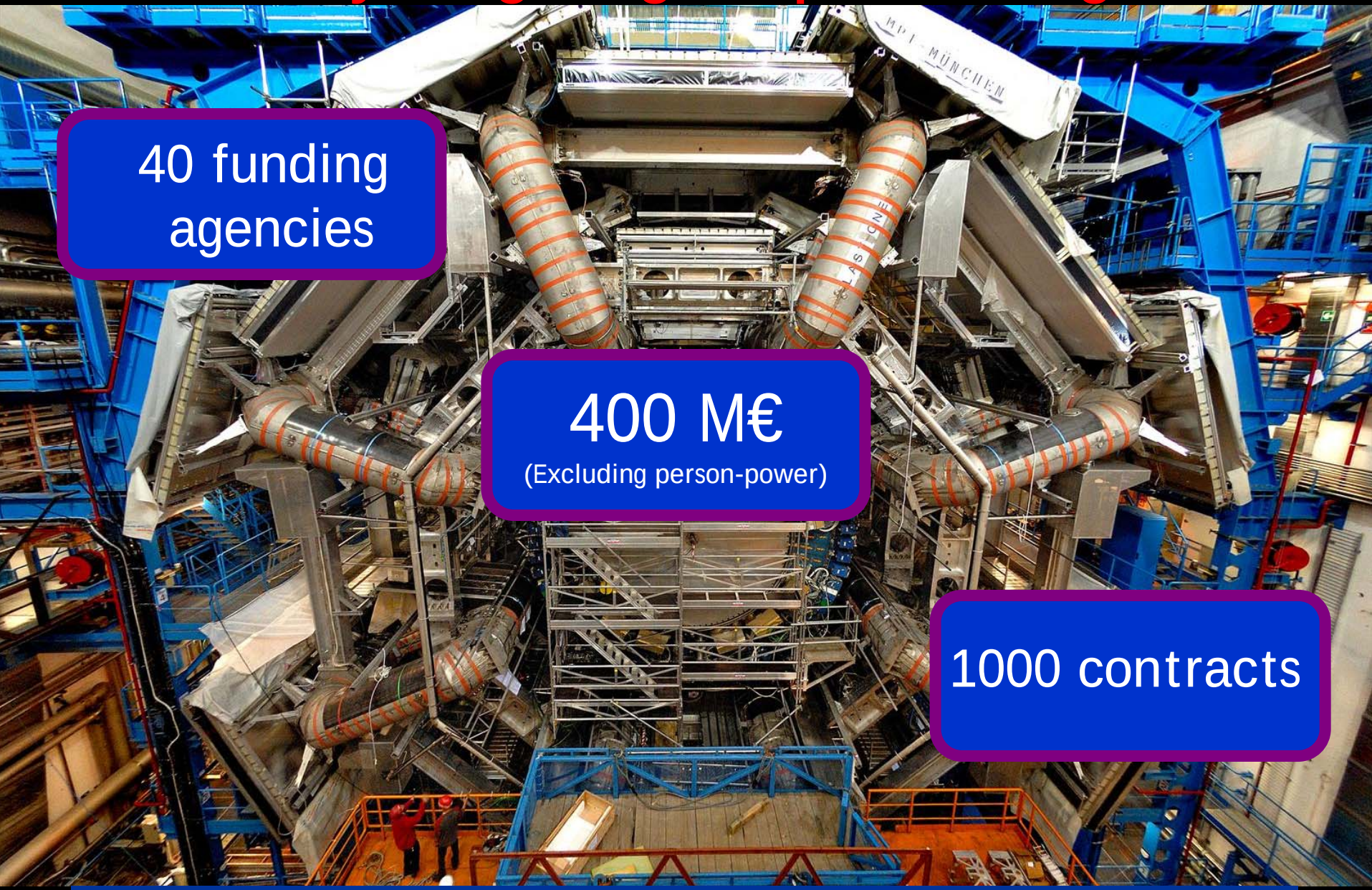
Make a 10% allowance for developing countries who at the beginning might not contribute to the scheme.



The model is viable only if every country is on board!

Allowing only SCOAP³ partners to publish Open Access would replicate the subscription scheme and not solve the problems.

How are you going to put it together?



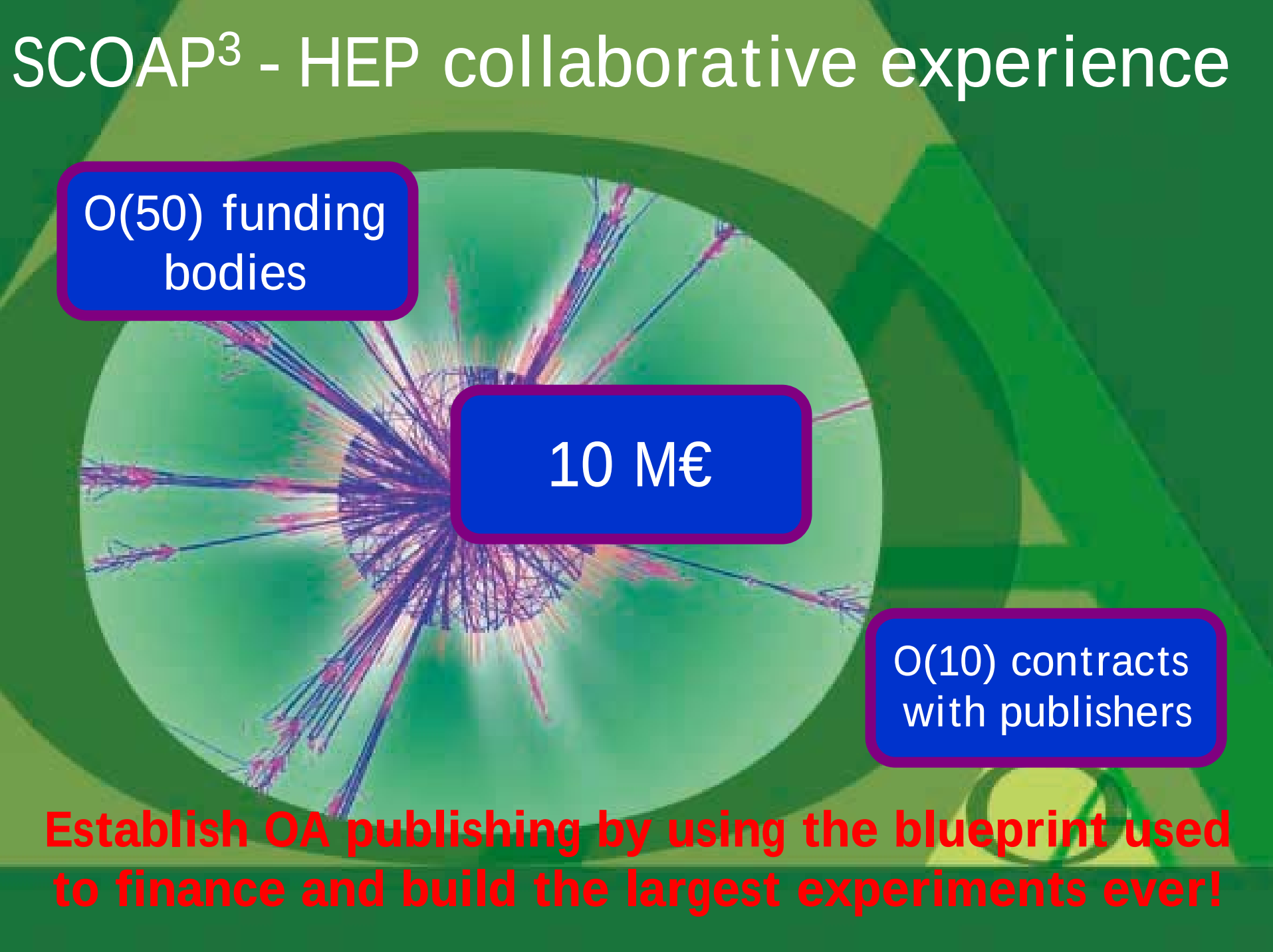
40 funding agencies

400 M€
(Excluding person-power)

1000 contracts

The ATLAS detector ready for discoveries at the LHC!

SCOAP^3 - HEP collaborative experience



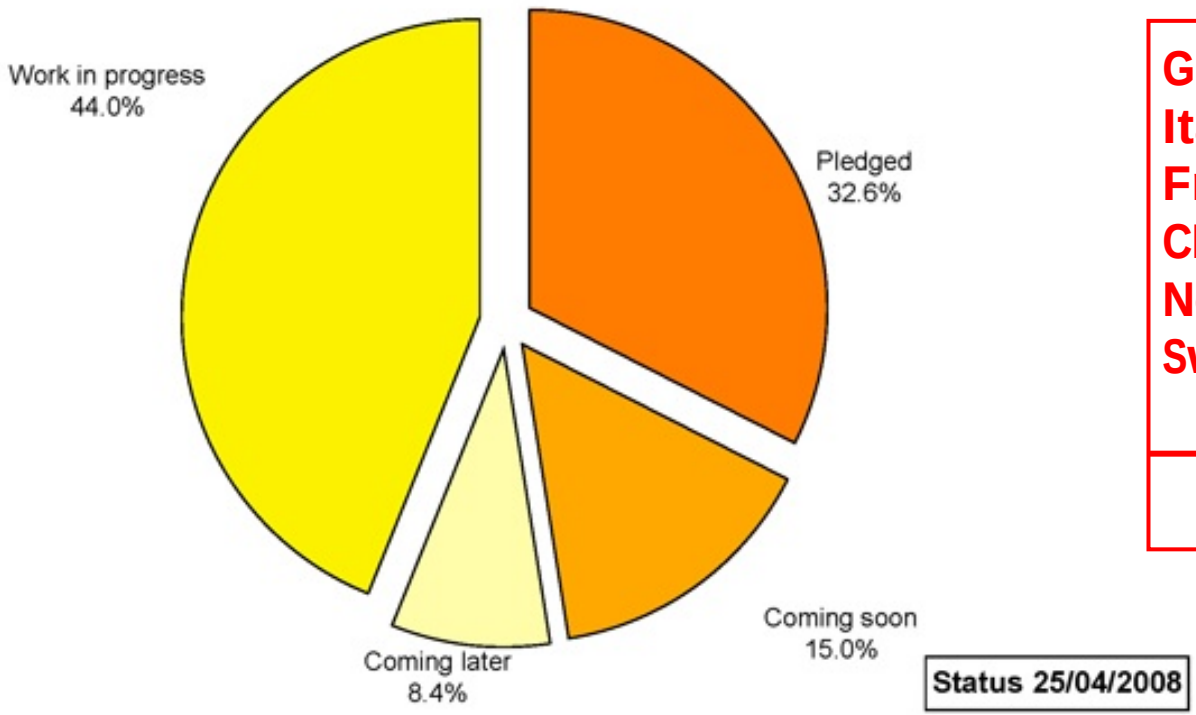
O(50) funding bodies

10 M€

O(10) contracts with publishers

Establish OA publishing by using the blueprint used to finance and build the largest experiments ever!

Progress of fund-raising



Germany	Romania
Italy	Hungary
France	Denmark
CERN	Greece
Norway	Slovak rep.
Sweden	Austria
	Belgium

26 US partners

3.3M€ pledged so far 1.5M€ coming very soon

All European countries internally discussing SCOAP3.
Negotiations/Discussions in progress with:
Canada, China, Japan, Korea, Russia, rest of US, …

SCOAP3 fund-raising and next steps

- Bodies which today buy HEP journals (libraries, consortia, institutes which finance them) estimate their current expenditure on the HEP journals targeted by SCOAP3
- Partners pledge a re-direction of their current expenditure as their contribution to SCOAP3 through an Expression of Interest.
  - A single partner (or a joint venture) partner pledges funds for an entire country, organizing the re-direction of subscriptions in that country
  - In the U.S., decentralized funding means many partners each pledging their present expenditures
- Once a sizeable fraction of budget is pledged
  - SCOAP3 formally established, with international governance
  - SCOAP3 can issue a tender to publishers
- Publishers answer the tender, with agreement on:
  1. Journal licence packages are un-bundled, the OA titles are removed and subscription prices are reduced accordingly
  2. In the case of long-term subscription contracts, publishers will be required to reimburse subscription costs pertaining to OA journals
- SCOAP3 partners adjudicate contracts and commit funds
- Contracts with publisher are signed and funds are transferred to SCOAP3
- Aim to 3-year tendering cycle and funding in sliding windows

"La seule chose promise d’avance à l’échec, c’est celle que l’on ne tente pas.”

Paul-Emile Victor

French Polar Explorer, 1907-1995

Thank you !

Salvatore.Mele@cern.ch

scoap3.org