

A Poll on HEP Information Systems

Jointly organised by
CERN/DESY/FNAL/SLAC

VERY preliminary results...
...just a taste of the answers

library.cern.ch/poll.html

Salvatore Mele
CERN

SLAC - May 10th 2007

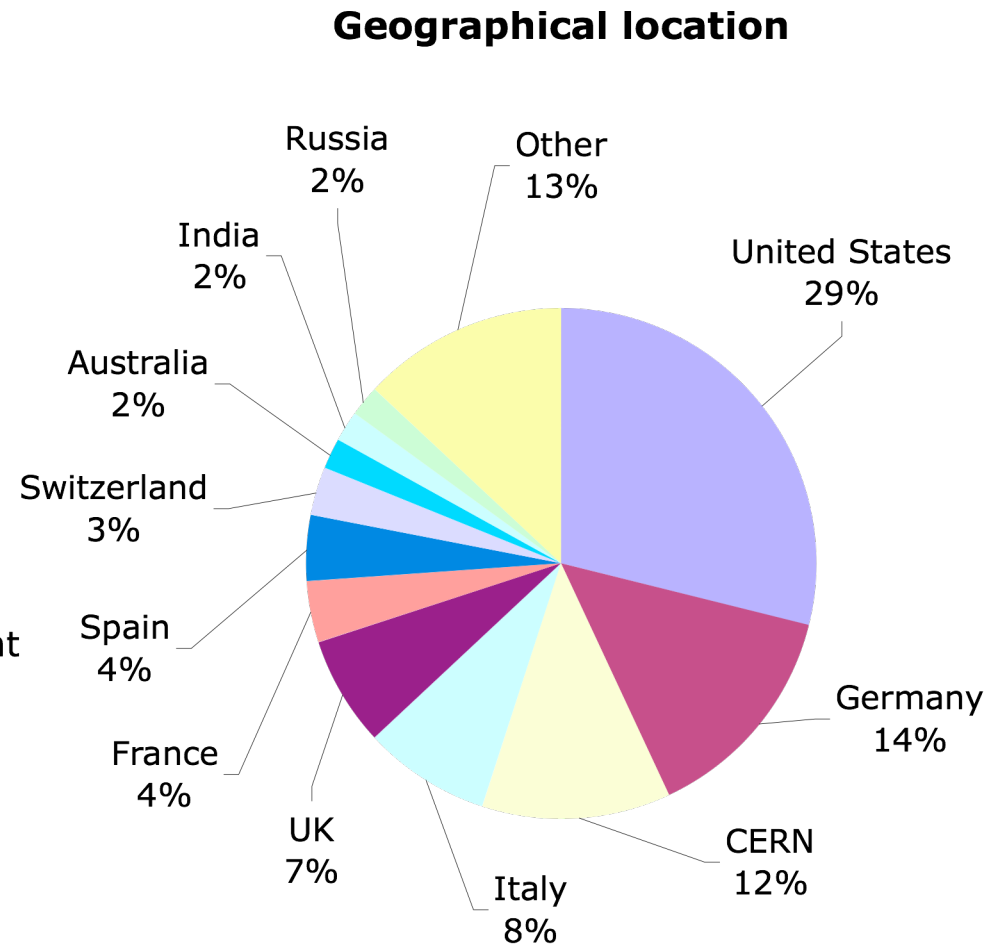
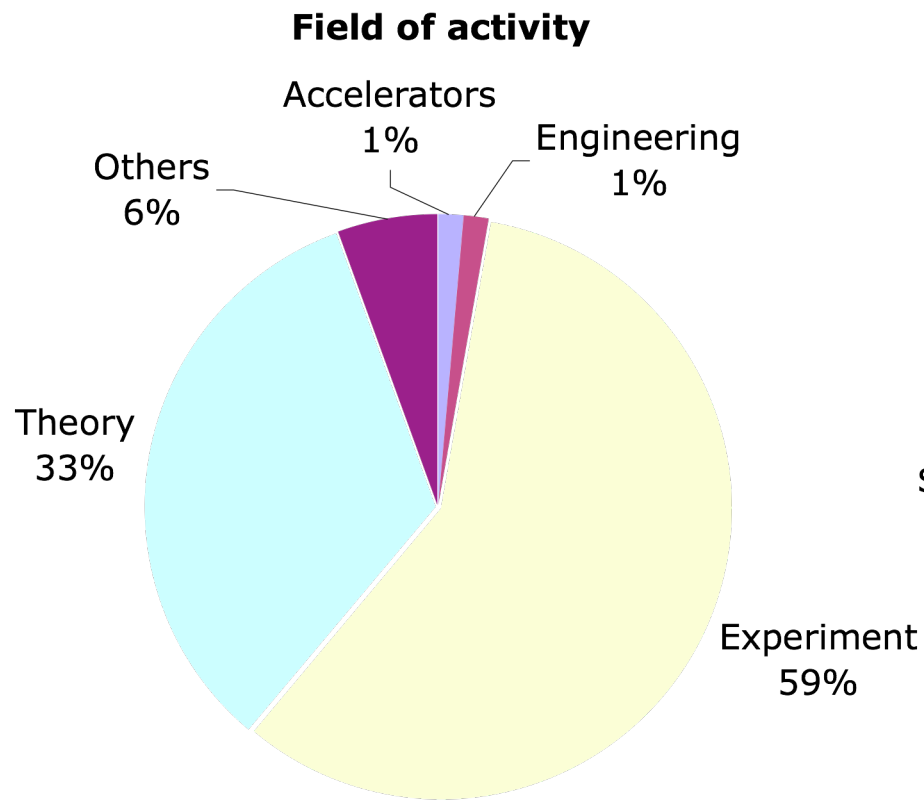
Targets of the poll

- **Survey user perception of present HEP information systems**
- **Assess user requirements and preferences**
- **Learn nitty-gritty details for short-term (easy and feasible) improvements of current systems**
- **Look for the killer application(s) of the next years**
- **Be inspired: develop “HEPIS” as a user-driven platform**

Deployment of the poll

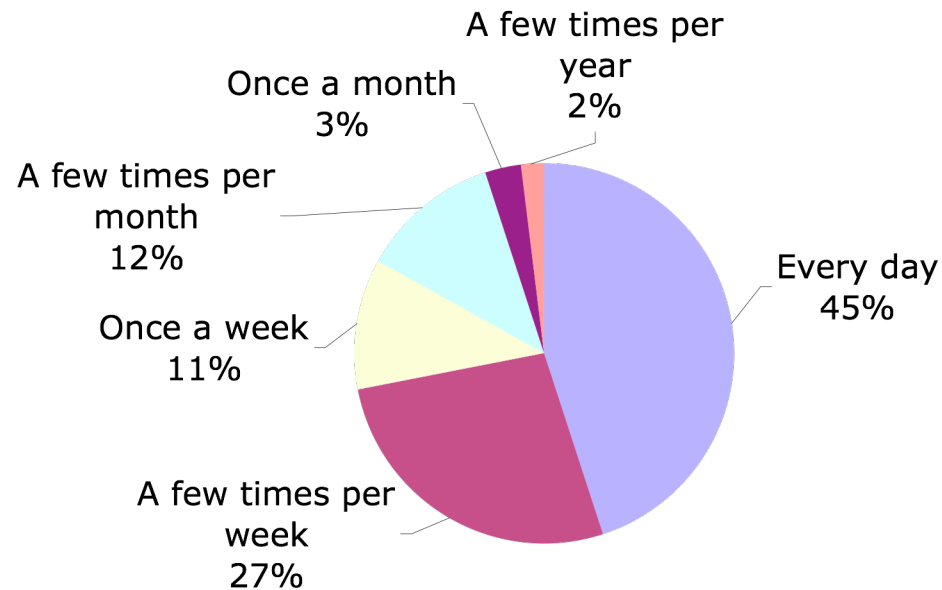
- Launched on April 30th
- Link from the CDS and SPIRES entry pages
- Advertised on CERN experiments mailing lists, CERN TH division, several hypernews, other communication channels at DESY/FNAL/SLAC
- About 500 answers by May 8th 0100 UT

Who answered?

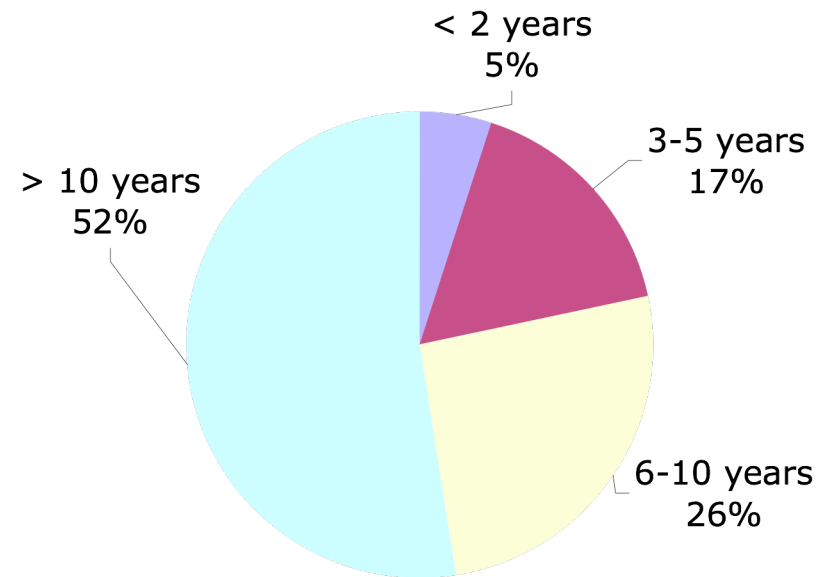


Highly-qualified answers

Frequency of use of HEP Information Services



Experience with HEP Information Services



Concerned users:

- o 43% wish to receive the results via e-mail
- o 89% answered to two or more of the nine “free-text” questions

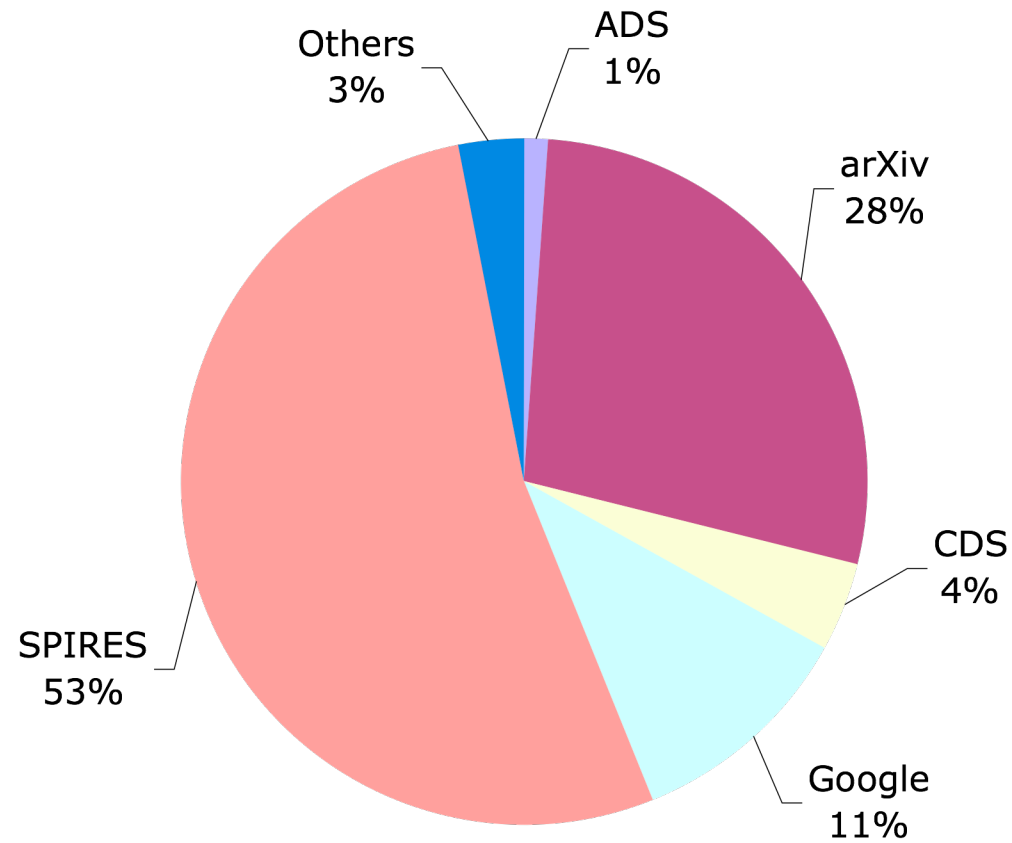
Free text answers: what's on a user mind?



TagCrowd
ALPHA

A survey of present systems

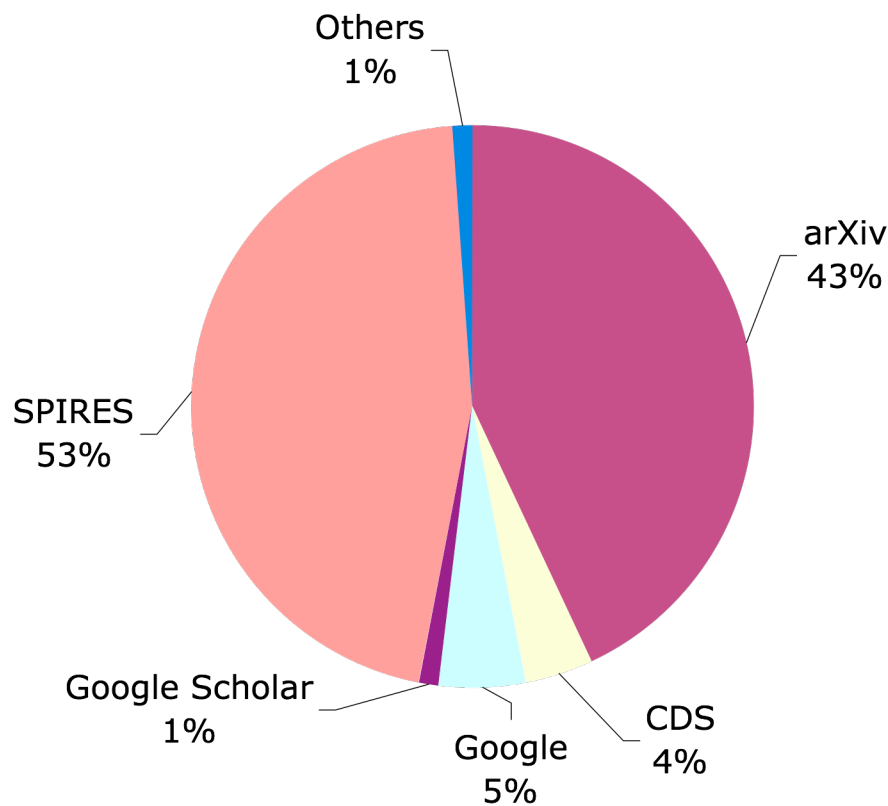
Which system do you use the most?



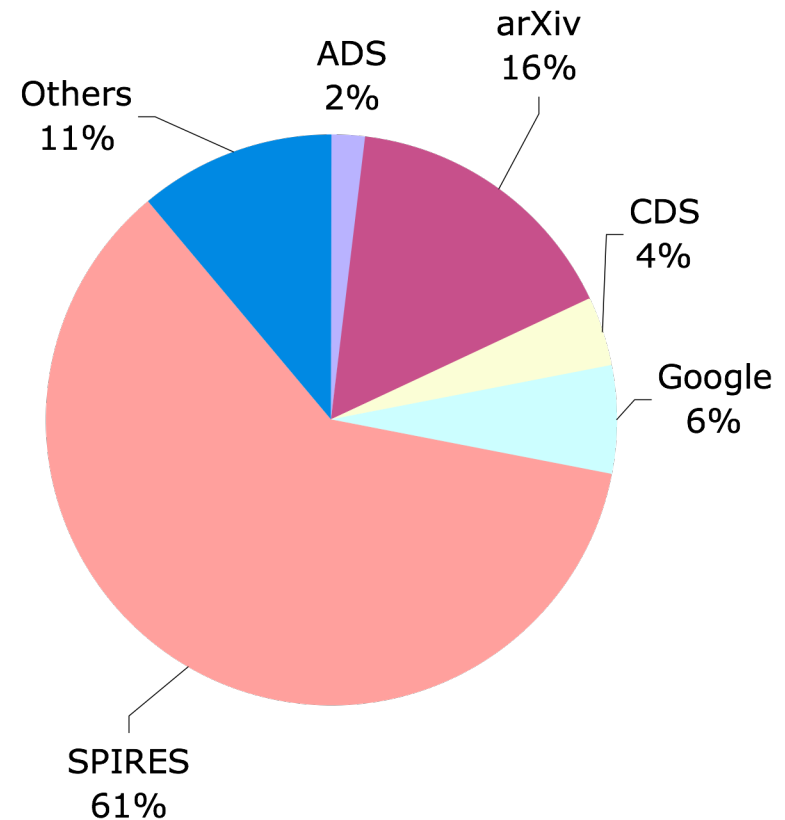
Known-item search

Which system do you use the most...

...to find preprints of which you know the reference?



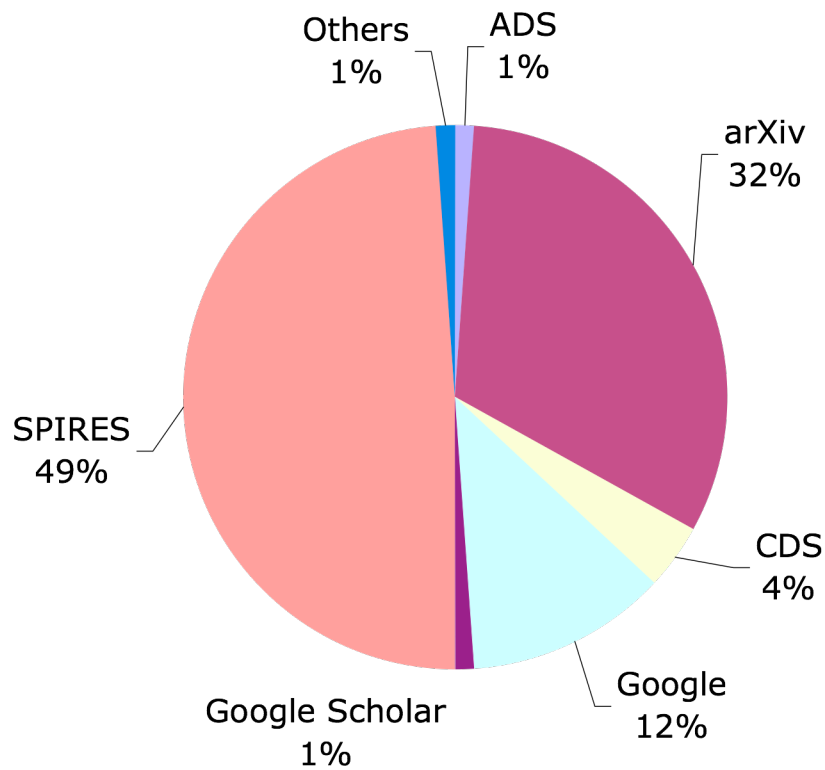
...to find articles of which you know the reference?



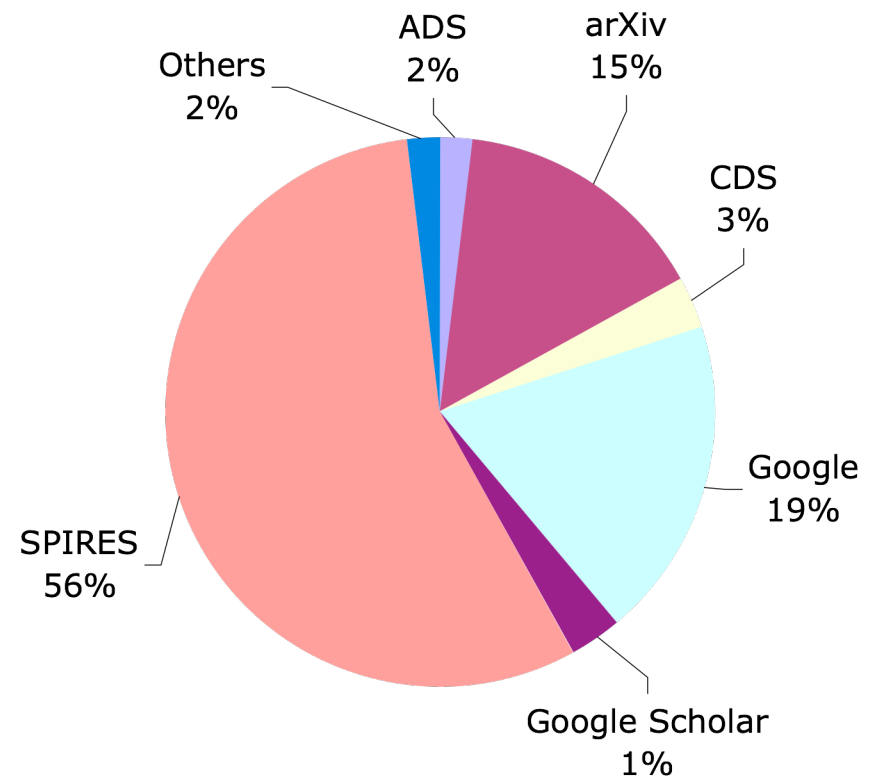
Subject search

Which system do you use the most...

...to find preprints on a given subject?



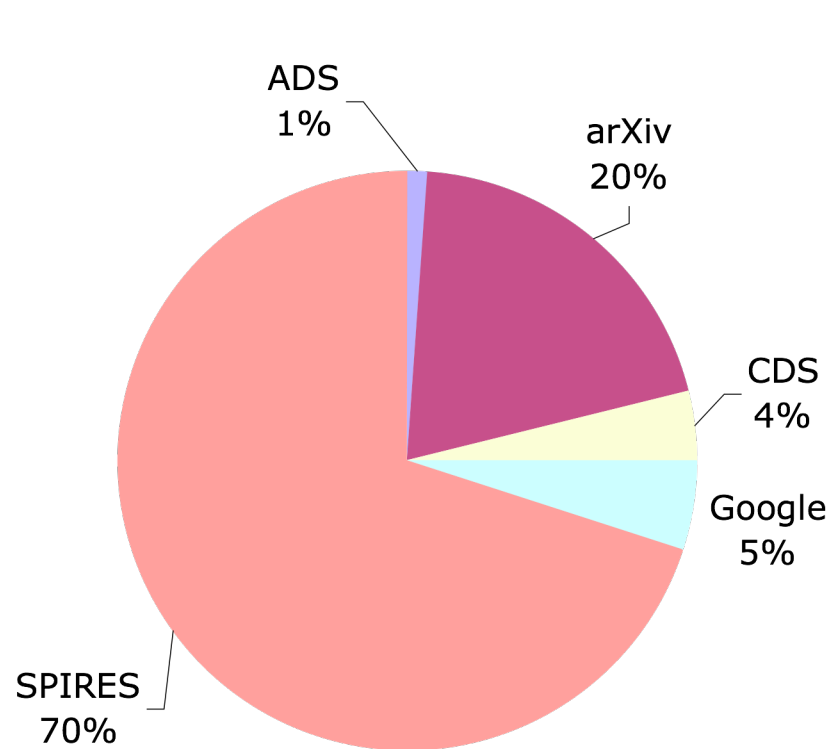
...to find articles on a given subject?



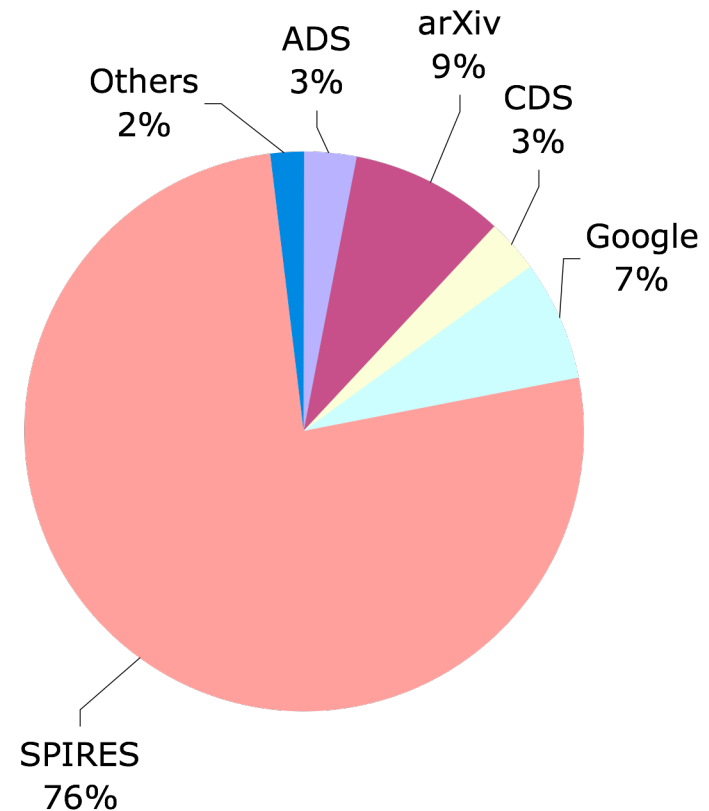
Author search

Which system do you use the most...

...to find preprints by a given author?



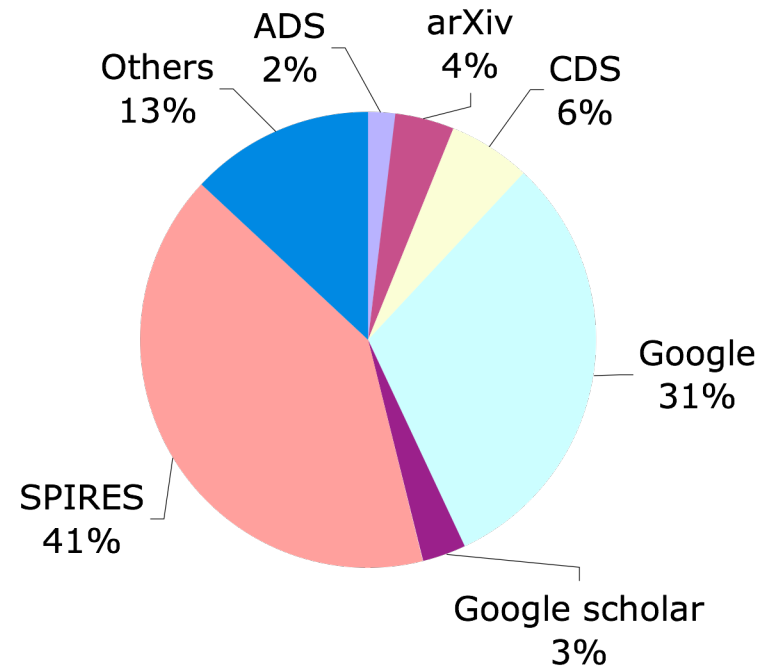
...to find articles by a given author?



“Greyer” literature

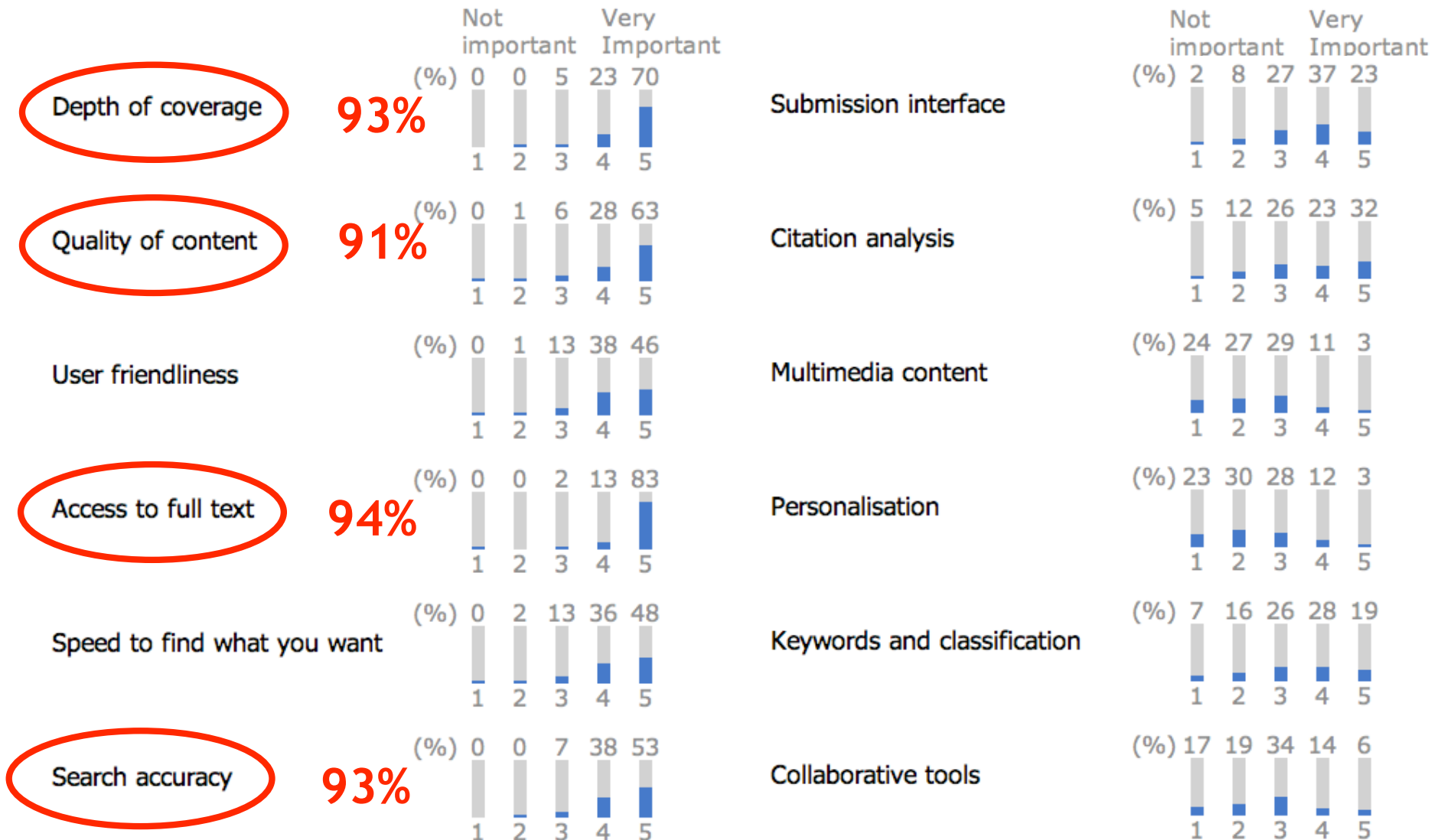
Which system do you use the most...

...to find theses?

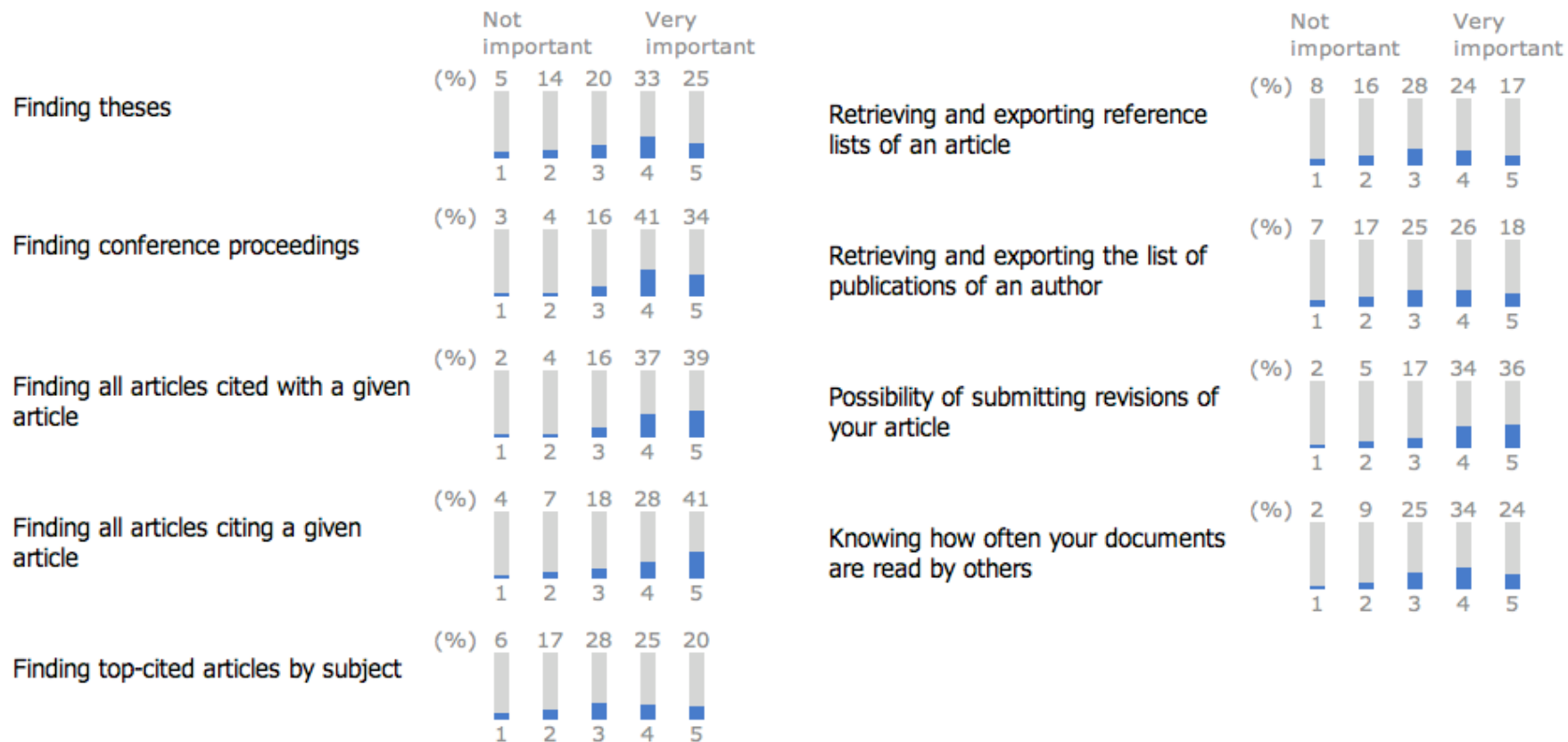


User preferences

How important are these features of an information system?

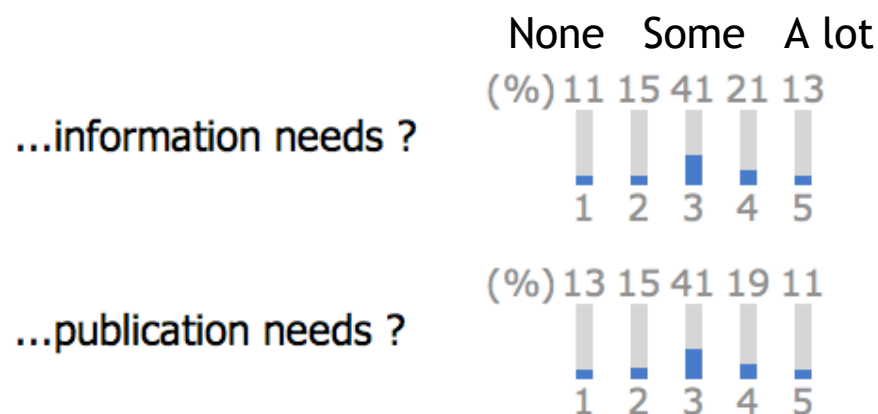


How important are these features of an information system?



A look into the (near) future

Let's give a look to the future. In the next five years, how much change would you expect in your... expect in your...



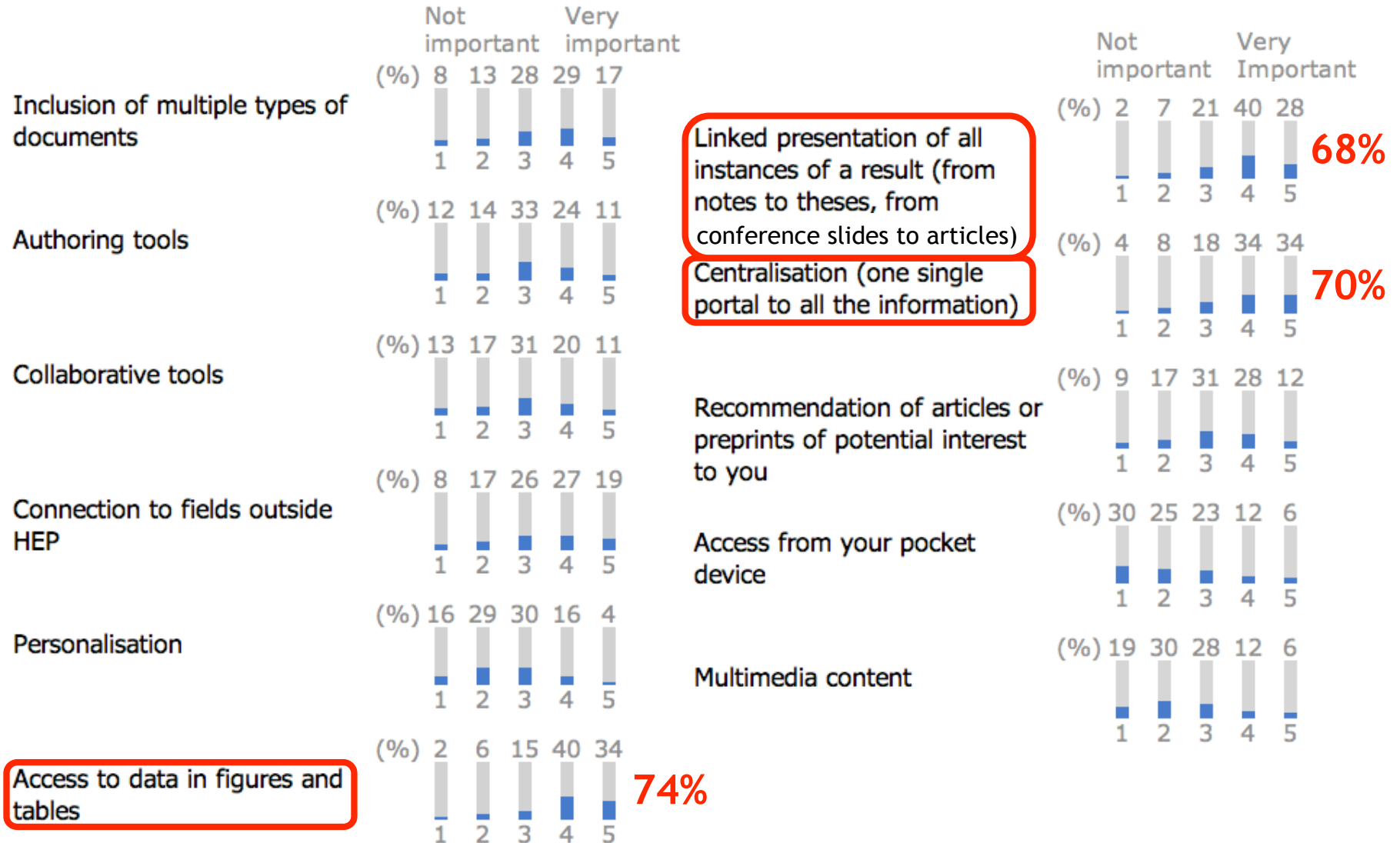
More than 70% of user expect at least some change

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
- 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

Imagine your ideal scientific information system in five years from now. Which additional features will be important?



Any features you would dream of?

Cover the contents of all famous journals going back to their starting dates and a database which contains something like flow diagrams showing how certain articles initiated further research. (Papers citing the original ones etc)

Follow the 'paper trail'. Citations/References linked to be a single click away inside of the article and open access from anywhere and include peer-reviewed journals...

A weekly alert of all the new preprints and publications in certain predefined (but rather general) fields, sent by email, in a compact form (html), with the title, first few authors, with a hidden link to a page with the abstract and a link to the PDF file, similar arXiv today

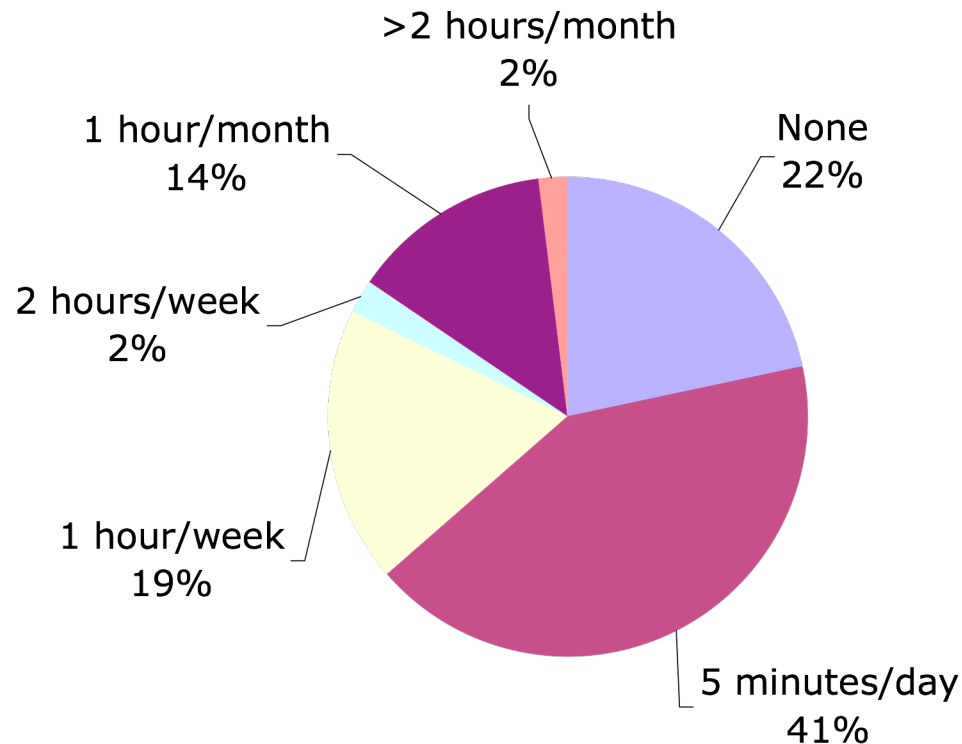
Access to code fragments which could accompany a plot or equation. This would make it easier to re-use results without making many errors.

Instead of a paper centric system it would be nice to have an idea/topic centric information system. Following the ideas of wikipedia/myspace, each idea in HEP would have its own page, with, for example: a brief outline of the idea, a list of the most recent review papers, a list of the latest cutting edge research in that area, a summary of experimental results, a list of people working in that area with links to their personal pages.

A more clever system of searching for a paper that is "connected" by title, citations, references, ... to a given paper

Web2.0 applications?

If a simple web interface would show you an article and offer a set of categories to which it could belong, how much time would you spend in this tagging system to give a service to the community?



Preliminary conclusion

Users know what they want: listen!

Outlook

- The poll is still running, plan in one or two weeks more, maybe with some e-mail reminders
- Many more (juicy!) questions I did not quote today (“what do you like the most/least of”)
- Will put together a white paper, relevant for discussions on a future common platform