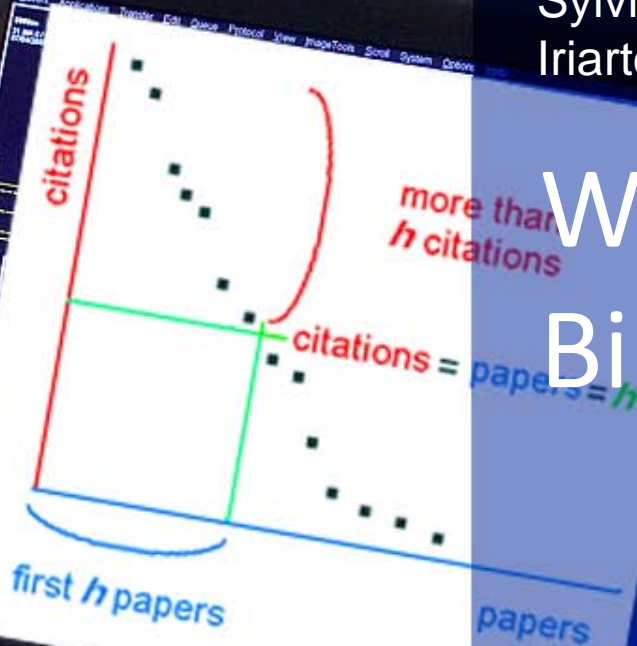


Sylvie Godel, Isabelle de Kaenel, Pablo Iriarte

# Web Services for Bibliometrics



Unil

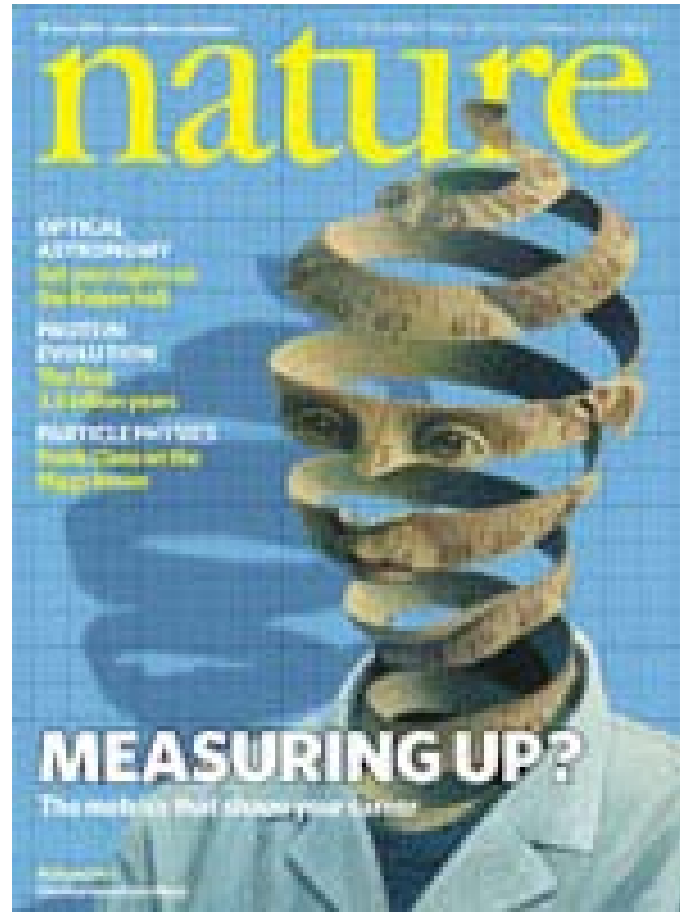
UNIL | Université de Lausanne  
Faculté de biologie et de médecine



# Plan

- Definitions
- Experiences in Lausanne
- The Institutional repository (IR) main role
- Citation-enhanced databases
- From IR to Bibliometric analysis
- Conclusion

# Bibliometrics at the Nature front page of this week!



*Bibliometrics is the analysis of information pertaining to a journal's publications and the citations that link them, which can be aggregated at the level of authors, institutes, countries, journals, and subjects.*

*Web services are applications that can be accessed via Hypertext Transfer Protocol and delivers information in a machine readable form like XML*

# The web service experience

- In Lausanne, the web services are used to fill-in entry forms with identifiers
  - ILL orders web form
  - SERVAL Institutional repository

The image shows a web interface for filling in a document order. It consists of two main parts: a dropdown menu and a main form.

The dropdown menu on the left lists the following options: Article, Pubmed/pmid, DOI, Pubmed/pmid (highlighted), Rero/roid, Rero/isbn, and Web of Science/wosid.

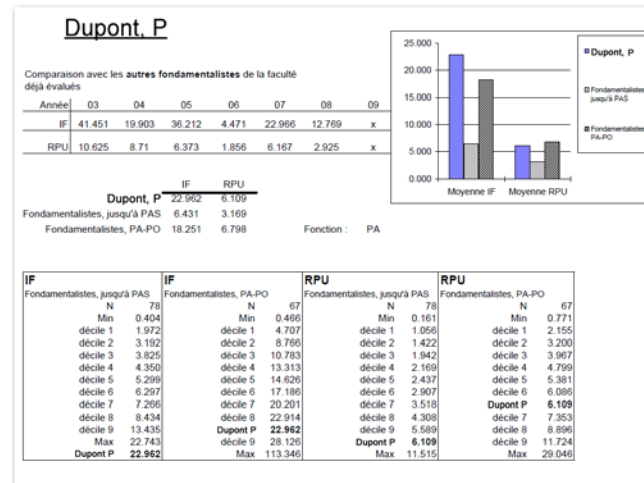
The main form is titled "Remplir la commande à partir du" and includes an "OK" button. It contains the following fields:

- Type de document : Article (dropdown)
- Titre du périodique ou du livre \* : [input field]
- Année \* : [input field] Val. \* : [input field] (No) : [input field] Suppl. : [input field] Pages \* : [input field]
- Titre de l'article ou du chapitre : [input field]
- Auteur(s) : [input field]
- Edition (pour les livres) : [input field] ISSN / ISBN : [input field] UID : [input field]
- Remarques : [text area]

At the bottom of the form are two buttons: "Enregistrer" and "Effacer".

# The bibliometrics experience

- Assistance to the faculty research evaluation unit
- Assessment of research activity at the author level with Impact Factors (IF)



# The institutional repository (IR) main role

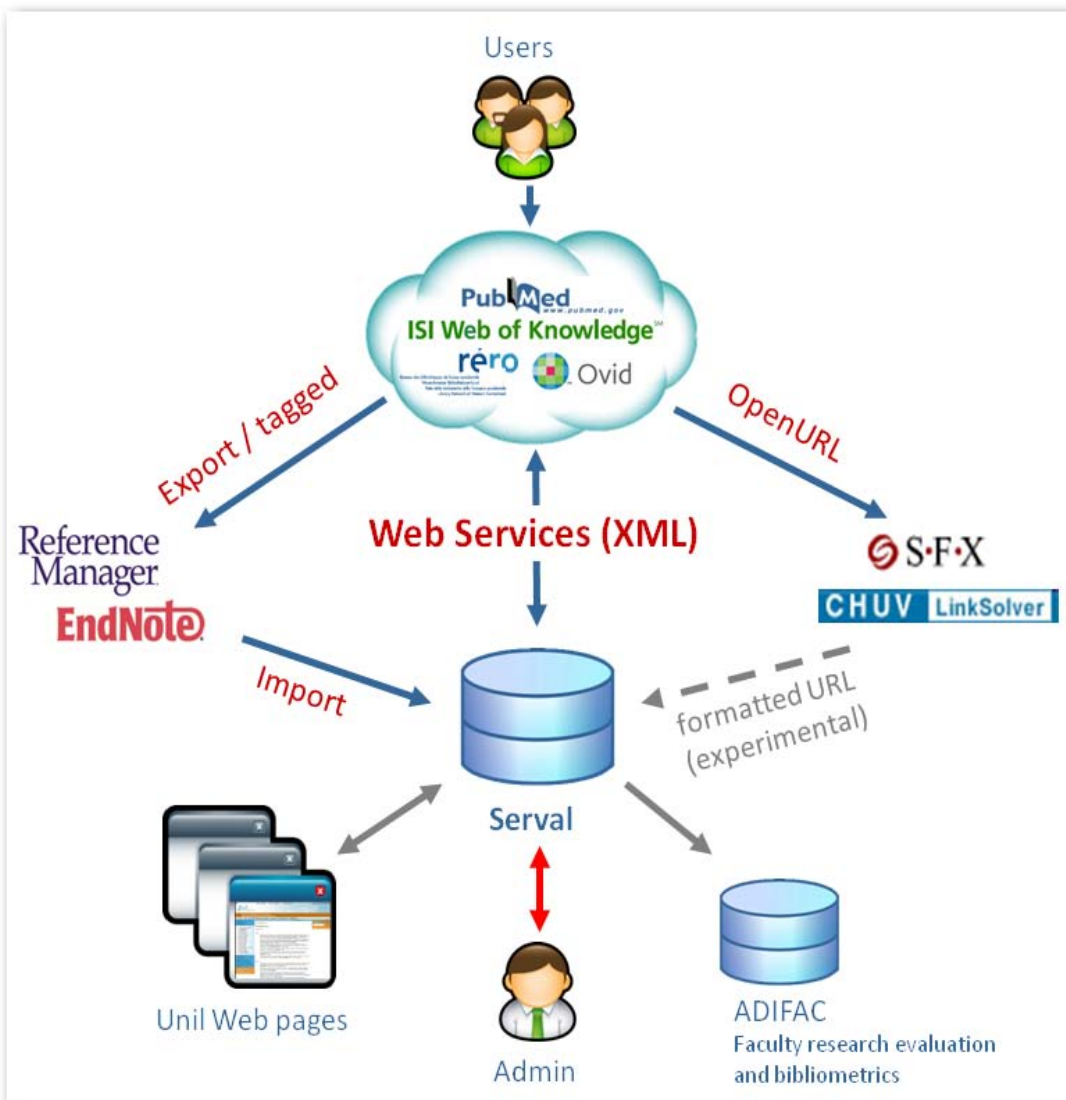
## – Pivot

- One single database
- Author's / units publication lists on the Web
- Metadata re-used by administrative tool (ADIFAC)

## – Quality Control

- Authoritative metadata imported from trusted sources (PubMed, WoS, CrossRef, etc.)
- Data transfers through bibliographic management software
- Introduce unique Identifiers (PMID, DOI, UT, etc.)
- Avoids homonyms and duplicates





**The metadata acquisition techniques tested in SERVAL**

# Citation-enhanced databases

	<b>Size</b>	<b>Citations coverage</b>	<b>Web Service</b>	<b>Free version</b>
<b>Web of Science</b>	46 M	1900 ->	Yes (SOAP / POST)	No
<b>SCOPUS</b>	40 M	1996 ->	Yes (JSON)	Yes (limited)
<b>PubMed Central</b>	2 M	2000 ->	Yes (OAI)	Yes

# From IR to bibliometric analysis

- Web services can be used to
  - Enrich the author's publication list
  - Produce bibliometric analysis on the fly

**serval**  
serveur académique lausannois

+

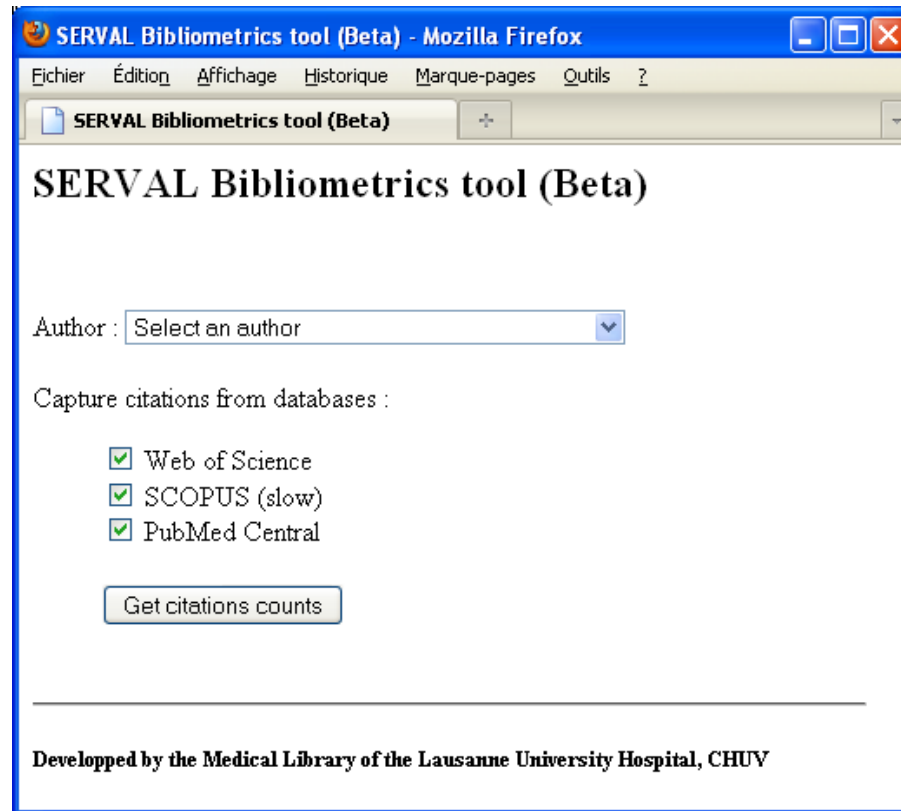
**SCOPUS**

=



Web of Science		Scopus		Pubmed Central	
Publ. (n)	% of total	Publ. (n)	% of total	Publ. (n)	% of total
1	100.00%	1	100.00%	1	100.00%
2	200.00%	2	200.00%	2	200.00%
3	300.00%	3	300.00%	3	300.00%
4	400.00%	4	400.00%	4	400.00%
5	500.00%	5	500.00%	5	500.00%
6	600.00%	6	600.00%	6	600.00%
7	700.00%	7	700.00%	7	700.00%
8	800.00%	8	800.00%	8	800.00%
9	900.00%	9	900.00%	9	900.00%
10	1000.00%	10	1000.00%	10	1000.00%

# Testing the prototype



The screenshot shows a web browser window titled "SERVAL Bibliometrics tool (Beta) - Mozilla Firefox". The browser's menu bar includes "Fichier", "Édition", "Affichage", "Historique", "Marque-pages", "Outils", and "?". The address bar shows the page title "SERVAL Bibliometrics tool (Beta)". The main content area displays the title "SERVAL Bibliometrics tool (Beta)" in a large, bold font. Below the title, there is a form with the following elements:

- An "Author:" label followed by a dropdown menu containing the text "Select an author".
- The text "Capture citations from databases:" followed by three checked checkboxes:
  - Web of Science
  - SCOPUS (slow)
  - PubMed Central
- A button labeled "Get citations counts".

At the bottom of the page, a horizontal line separates the form from the footer text: "Developed by the Medical Library of the Lausanne University Hospital, CHUV".

<http://www.bium.ch/bibliometrics/>

# Behind the screen: WoS

<http://www2.unil.ch/openillink/openlinker/bibliometrics/ut.php?pmid=6091913>

Response :

```
<?xml version="1.0" encoding="UTF-8" ?>
<response xmlns="http://www.isinet.com/xrpc41" src="app.id=SERVAL_Web_Services,env.id=Bibliometrics,partner.email=pablo.iriarte@chuv.ch">
<fn name="LinksAMR.retrieve" rc="OK">
<map>
<map name="cite_id">
<map name="WOS">
<val name="relatedRecordsURL">
<![CDATA[http://gateway.isiknowledge.com/gateway/Gateway.cgi?GWVersion=2&SrcApp=PARTNER_APP&SrcAuth=LinksAMR&KeyUT=A1984TS61800024&
DestLinkType=RelatedRecords&DestApp=ALL_WOS&UsrCustomerID=2109fe80815e8621515050a12879ac1a]]></val>
<val name="sourceURL">
<![CDATA[http://gateway.isiknowledge.com/gateway/Gateway.cgi?GWVersion=2&SrcApp=PARTNER_APP&SrcAuth=LinksAMR&KeyUT=A1984TS61800024&
DestLinkType=FullRecord&DestApp=ALL_WOS&UsrCustomerID=2109fe80815e8621515050a12879ac1a]]></val>
<val name="timesCited">897</val>
<val name="ut">A1984TS61800024</val>
<val
name="citingArticlesURL"><![CDATA[http://gateway.isiknowledge.com/gateway/Gateway.cgi?GWVersion=2&SrcApp=PARTNER_APP&SrcAuth=LinksAMR&K
eyUT=A1984TS61800024&DestLinkType=CitingArticles&DestApp=ALL_WOS&UsrCustomerID=2109fe80815e8621515050a12879ac1a]]></val>
</map>
</map>
</map>
</fn>
</response>
```

# Behind the screen: SCOPUS

Request:

```
http://www.scopus.com/scsearchapi/search.url?devId=[the developer ID]&search=PMID%286091913%29&callback=t
```

Response:

```
t({"PartOK":{"TotalResults":"1","ReturnedResults":"1","Position":"0","Results":[
  {"title":"Organization of the higher-order chromatin loop: Specific DNA attachment sites on nuclear scaffold",
    "firstauth":"Mirkovitch, J.",
    "citedbycount":"391",
    "pubdate":"1984",
    "sourcetitle":"Cell",
    "inwardurl":"http://www.scopus.com/inward/record.url?eid=2-s2.0-0021675784&partnerID=65&md5=4b9314f61428df989c6b434005a7dc36
  }
]})
```

# Behind the screen: PMC

Request:

```
http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?db=pubmed&cmd=neighbor&linkname=pubmed_pmc_refs&id=6091913
```

Response:

```
<?xml version="1.0"?>
<!DOCTYPE eLinkResult PUBLIC "-//NLM//DTD eLinkResult, 10 August 2009//EN" "http://www.ncbi.nlm.nih.gov/entrez/query/DTD/eLink_090910.dtd">
<eLinkResult>
  <LinkSet>
    <DbFrom>pubmed</DbFrom>
    <IdList>
      <Id>6091913</Id>
    </IdList>
    <LinkSetDb>
      <DbTo>pmc</DbTo>
      <LinkName>pubmed_pmc_refs</LinkName>
      <Link>
        <Id>2774341</Id>
      </Link>
      <Link>
        <Id>2683591</Id>
      </Link>
      [...]
    </LinkSetDb>
  </LinkSet>
</eLinkResult>
```

# Match and mix

- Web services used to match
  - the IR records with the citedness scores
    - using unique identifiers to query the sources
    - parsing XML/JSON responses to extract citations counts
- Table display of the mixed results

Coordonnées		Enseignements		Publications		Bibliométrie	
Number of publications in SERVAL: 31							
Metrics	Web of Science®	SCOPUS®	PubMed Central	Max.			
Publications retrieved	24 of 31 (77.42%)	28 of 31 (90.32%)	25 of 31 (80.65%)	29 of 31 (93.55%)			
Sum of the Times Cited	2103	912	510	2124			
Average Citations per Item	87.63	32.57	20.40	73.24			
Number of never cited publications	0 of 24 (0.00%)	0 of 28 (0.00%)	5 of 25 (20.00%)	0 of 29 (0.00%)			
h-index	16	13	10	16			
g-index	45	30	22	46			
g/h ratio	2.81	2.31	2.20	2.88			

Details :						Cited by				Average citations per year			
Ref.	SERVAL ID	Year	DOI	PMID	UT	WoS	Scopus	PMC	Max.	WoS	Scopus	PMC	Max.
1	BJB_CC4FC6004183	1984		6091913	A1984T561800024	897	391	225	897	33.22	14.48	8.33	33.22
2	BJB_FA0C288516BC	1991		1901265	A1991FE95000023	296	70	70	296	14.80		3.50	14.80
3	BJB_F4DB20204181	1988		3132558	A1988M424800010	125	58	33	125	5.43	2.52	1.43	5.43
4	BJB_3C16F94535C1	1988		3132557	A1988M424800009	122	54	25	122	5.30	2.35	1.09	5.30
5	BJB_BB24B8484DC5	1989		2726767	A1989U940700041	96	24	25	96	4.36	1.09	1.14	4.36
6	BJB_F74E14052B12	1994		7523862	A1994PM6400026	81	81	24	81	4.76	4.76	1.41	4.76
7	BJB_55255A09885C	1993		8375375	A1993LW53600001	80	61	10	80	4.44	3.39	0.56	4.44
8	BJB_8AD0CC74260F	1986		3098982	A1986D342500010	63	20	15	63	2.52	0.80	0.60	2.52
9	BJB_3FC72A7C888A	1992		1729591	A1992GW05300001	59	35	22	59	3.11	1.84	1.16	3.11
10	BJB_27C6980EFC77	1987		2894682	A1987L347900012	57	23	10	57	2.38	0.96	0.42	2.38



# Conclusion

- The IR metadata can be enriched with citation information using Web Services
- Citation information moves and is not easy to integrate in the faculty assessment process
- The best of all possible worlds: comparison of citations for each record
  - ➔ **The highest score is selected**
- Web services allows the calculation of a « **mixed data h index** »

Muito obrigado  
Merci!  
Gracias  
Thank you!