



# A short survey of bibliometric practices at the University of Vienna

Wolfgang Mayer

([wolf.mayer@univie.ac.at](mailto:wolf.mayer@univie.ac.at))

Hungarian Academy of Sciences, 02.11.2009



## Schedule:

- I. Historical overview and organizational embedding
- II. Tasks of the Bibliometrics Department
  - A Regular bibliometric analyses and services
  - B Consultation of rectorate, departments and staff
  - C Development partnerships ("early adopter"), product analyses
  - D National and international cooperation
  - E Problem-specific expertises
  - F Structuring, enrichment and coordination of (bibliometric) data
- III. Bibliometric analyses within evaluation of faculties
- IV. Procedures for faculty positions



## I. Bibliometrics at University of Vienna: Historical overview and organizational embedding

### History:

2006: Formation of working group "Scientometrics" (without budget);  
Members of University Library invite Quality Assurance,  
Research Activity Documentation (RAD), Vice Rectorate  
Research and Development, Research Services and  
International Development to discuss and collaborate

### Goals:

1. Coordination and concentration of competencies
2. Networking with international institutions
3. Effects of synergy by combining internal and external data sources

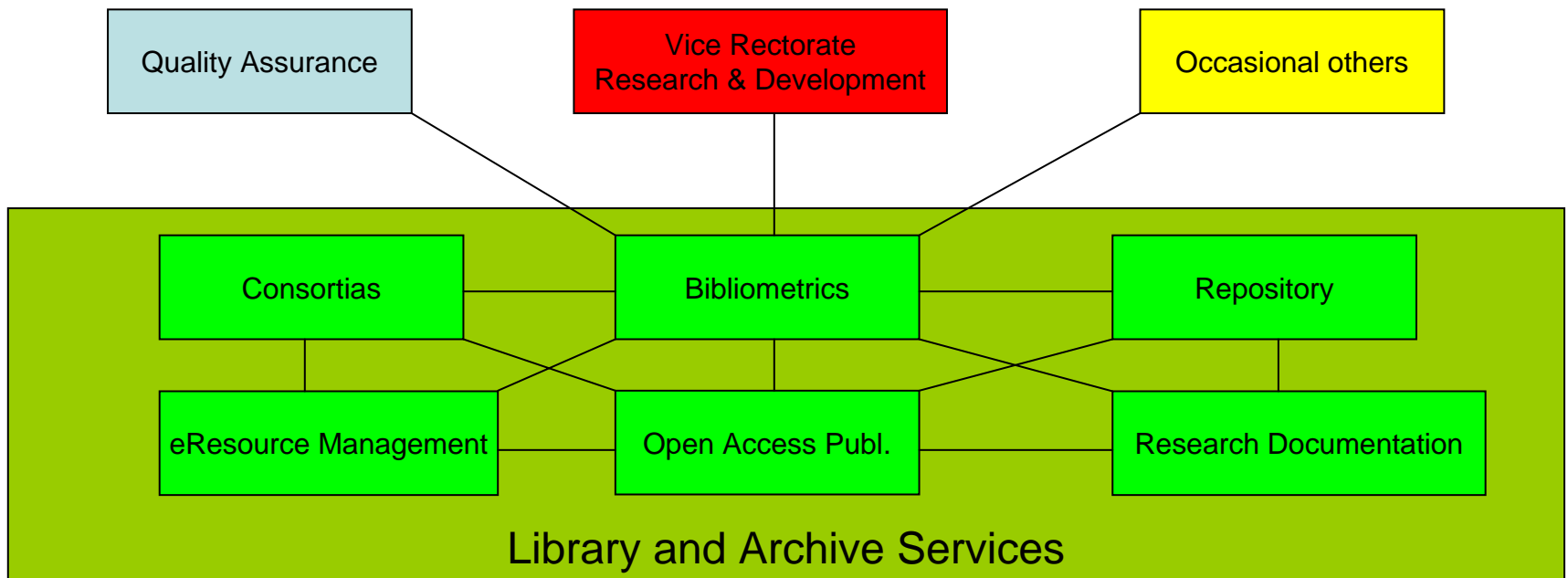
Activities: Discussing and clarifying common goals, individual competencies and quality of existing (though scattered) data. Writing first special reports and expertises for rectorate.



## I. Bibliometrics at University of Vienna: Historical overview and organizational embedding

### Current structure:

Since 2008: Bibliometrics Department implemented as part of Library and Archive Services





## I. Why positioning bibliometrics at (scientific) libraries?

- Librarians should be information specialists
- Librarians should be used to work with metadata, search engines and databases
- Libraries do already collect necessary content (publications) and license required tools (databases)
- Libraries should be independent and interdisciplinary institutions
- Libraries should be free of scientific bias



## **II. Tasks of the Bibliometrics Department**

- A, Regular bibliometric analyses and services
- B, Consultation of rectorate, departments and staff
- C, Development partnerships ("early adopter"), product analyses
- D, National and international cooperation
- E, Problem-specific expertises
- F, Structuring, enrichment and coordination of (bibliometric) data



## II. A. Regular bibliometric analyses

1. Periodic and cyclic evaluation of publication output of the 15 faculties and 3 centres (approx. 4 year cycle)

Quality Assurance Department prepares Performance Agreements (target concept) between rectorate and faculty. Discussion if and to what degree bibliometric analysis makes sense.

Important: Bibliometric report is NOT the result of evaluation but a discussion paper presented <sup>(1)</sup> to the dean in advance, <sup>(2)</sup> to the selected peer in revised form, and <sup>(3)</sup> revised and annotated to rectorate.

[Workflow: see III]



## II. A. Regular bibliometric analyses

### 2. Procedures for faculty positions

Service offer to the faculty administration to analyze applications for vacant chairs (service formerly provided in part by Quality Assurance).

Critical discussion about standards and criteria wanted by faculty.

Complying with results of this discussion all publications or presented publication lists of applicants are analyzed.

[Workflow: see IV]





## II. B. Consultation

### Activities:

- Basic trainings in bibliometric and other scientific data bases (WoS, Scopus, abstract and indexing services)
- Bibliometric consulting for optimizing and personalizing publication output (Distinct Author Feedback, ResearcherID et al.)
- Using bibliometric services and resources for evaluation (incl. JCR, ESI, SciMago...)
- Training faculty staff in analysing publication lists
- Preparing statements to new developments (rankings...)
- Designing bibliometrics service and information website (Start: January 2010)



## II. C. Product development

### Activities:

- Library decides (within budgetary boundaries) licensing of e-resources, tools et al.
- eResource department is preparing and substantiating these decisions
- Negotiating development / early testing status, especially in case of sophisticated (and cost-intensive) products
- Trim product to match institutional interests  
(correcting/enhancing data, improving functionalities...)
- Creating reports for policy maker and provider



## II. D. National and international cooperation

### Activities:

- Organizing an country-wide platform for the different departments dealing with bibliometrics, research evaluation, quality assurance...
- Inviting international specialists (providers, renowned scientists) to discussions and workshops
- Creating channels for (legal) data exchange for expertises, benchmarks...
- Organizing international conferences (S&TI 2009, ISSI 2013)
- Platform "Bibliometrics & Quality Assurance" with German universities of excellence
- Summer School Bibliometrics with German institutions



## II. E. Problem specific expertises

### Examples:

- Comparing Social Science faculties of Vienna / Zurich / Oslo
- Publications of Computing Sciences using alternative resources (CiteSeerX...): faculty or institutes at other faculties
- Citation loss of Vienna University due to inaccurate affiliation information (Austrian Academy of Science, IQOQI, IMP...)
- Evaluation of SCOAP<sup>3</sup> commitment für University of Vienna
- Citation frequency in Open Access initiatives (BioMed Central...)
- Retrospective analysis of Vienna University employment policy (what happened to unselected candidates )
- Evaluation of University Library's collection development



## II. F. Structuring data

### Goal:

Correct, up-to-date, consistent and compatible data in RAD, repository and external resources.

Efficient and uninterrupted workflows for entering and exchanging data.

### Activities:

- Enriching local RAD with external data (ISSNs, peer review, IF)
- Consultation in concept of author / institution identifiers
- ResearcherID upload
- Organizing contact with providers for future data migration interfaces
- Bibliographic / bibliometric control of data (entered by authors)
- Consultation in designing / modifying metadata set
- Controlling / ensuring data completeness in repositories and RAD



### **III. Bibliometric analyses within evaluation of faculties (and international comparisons)**

#### Workflow:

- Data collection
- Data evaluation
- Presentation of results
- Problems and discussion

Data should give a picture of a faculty's publication output to enable policy maker to identify strenghts and possible weaknesses.



## III. 1. Data collection

### Complex of problems:

- Neither authors nor affiliations standardized
- Incomplete and ambiguous author/affiliation relations
- "Distinct Author Sets" only of limited help  
(algorithm currently still flawed but much improved)
- Initials instead of first names (at least until 2005)
- Cognitive disambiguation necessary in case of homonyms
- Subject areas only at journal level (solution: indexing backbone?)
- Range of publications seldom constrained to certain "subject areas"
- Problem affiliation history



## III. 1. Data collection

### Process:

- Formulating a scaleable, reproduceable and refreshable search string
- Data harvesting in two steps (institution and author based)
- Affiliation disambiguation via cognitive and manual control
- Modelling logistics for further analyses





## III. 2. Data evaluation

Publication specific data is analyzed following four different ways of interpretation:

- a) Activity (Productivity)
- b) Visibility
- c) Impact (Citation)
- d) Other analyses



## III. 2. a. Activity (Productivity)

Analysis of the publication output. According to acknowledged bibliometric standards the following indicators for (typically) the last 5 years are examined:

1. Count of publications (overall and in chronological sequence)
2. Ratio of different document types (article, review, editorial, letter, proceedings, et al.)
3. Ratio of language of publication
4. Subject areas of publications
5. Number of co-authors/affiliations (normal and fractional counting)



## III. 2. b. Visibility

Papers in "higher ranked" publications have higher visibility potential. Accordingly the "quality" (attributes) of publications should be measured and depicted.

This illustrates the general publication strategy of an institution.

### Counting:

1. "Peer reviewed" publications
2. Impact Factors (focus on "Field Average Impact Factors")
3. Included in various special indexes / lists (eg. ERIH, ACM...)



## III. 2. b. Visibility of publications via Relative Impact Factor

### Indicator:

Sum of "Relative Impact Factors" (RIF) of journals you publish in.  
RIF is division of journal's Impact Factors (IF) by the Median or Aggregate Impact Factor of the respective subject area(s).  
Factor per publication can be also be calculated inter- or multidisciplinary.

## III. 2. b. Visibility of publications via Relative Impact Factor (example)

Publication	Journal	IF Journal	Subject Area Journal	Aggregated IF Subject Area	RIF
1	A	2.000	X	4.000	0.500
2	A	2.000	X	4.000	0.500
3	B	1.000	X	4.000	0.250
4	C	4.000	Y	2.000	2.000
5	D	3.000	Y	2.000	1.500
Sum					4.750
Per publication					0.950



## III. 2. c. Impact (Citation)

Citation analysis based evaluation of an institution's "impact" in the scientific community caused by publications.

### Counting:

1. Number of gained citations (overall and in chronological sequence)
2. "Origin" of citations (self citations, citing authors, institutions et al.)
3. Characteristics of most cited publications
4. Number and characteristics of non-cited publications
5. h-Index, g-Index et al.
6. Crown Indicator / RCR (Relative Citation Rate)
7. Ratio of highly cited publications using Baselines and Percentiles



## III. 2. c. Impact: Crown Indicator

Definition (CWTS Field Normalized Citation Score):

- Comparing the number of citations to publications of an institution inside a specified period with the worldwide average of citations to publications with similar document type, period and subject area.
- Harmonization of citation counts is achieved through sum of citations and "field citation scores".
- 0.9 means publications of this institution are cited 10% below, 1.2 means 20% beyond the average.

Computation:

1. Add all citations to publications of a group
2. Add worldwide average of citations consistent with selected papers considering <sup>(1)</sup> document type, <sup>(2)</sup> publication year and <sup>(3)</sup> subject area; [publication in case of RCR]
3. Divide sum of citations by sum of worldwide average

### III. 2. c. Impact: Baselines - Percentiles

- Essential Science Indicators (analytical tool from WoK)
- Count of citations necessary to belong to „Top x" percent
- Depends on publication year and subject area
- Spreadsheet divided into subject areas
- Also computable for smaller or virtual fields

#### Example: Physics

Physics	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	All Years
0.01 %	1899	1065	1077	929	776	613	462	331	210	80	25	756
0.10 %	442	402	382	345	284	229	186	128	79	37	12	259
1.00 %	129	125	118	107	89	75	65	47	30	14	5	80
10.00 %	31	31	30	26	24	21	18	14	9	4	2	19
20.00 %	17	17	17	15	14	12	11	9	6	3	1	10
50.00 %	5	5	5	5	5	4	4	3	2	1	0	3





## III. 2. d. Other analyses

1. Identification of collaboration patterns
2. Landscape of disciplines / interdisciplinarity
3. Co-citation analysis  
(common sources of a research field, citing traditions, relations, et al.)
4. Answers to institution specific questions regarding strategies, general conditions and structures



### III. Example: Bibliometric maps of research field

- Oslo: strong interdisciplinary network; focal points in many subject areas
- Zurich: strong focus on disciplines; emphasis on political science
- Vienna: tighter network than Zurich; more but less distinctive focal points

#### Zurich

Disciplines: 23  
Relations: two: 11 / three: 1

Name	Elements	Color
Political Science	5	
Geography	3	
Sociology	3	
Geriatrics and Gerontology	2	
Mathematics, Social Sciences	2	
Others (8)	1	



#### Vienna

Disciplines: 31  
Relations: two: 22 / three: 3

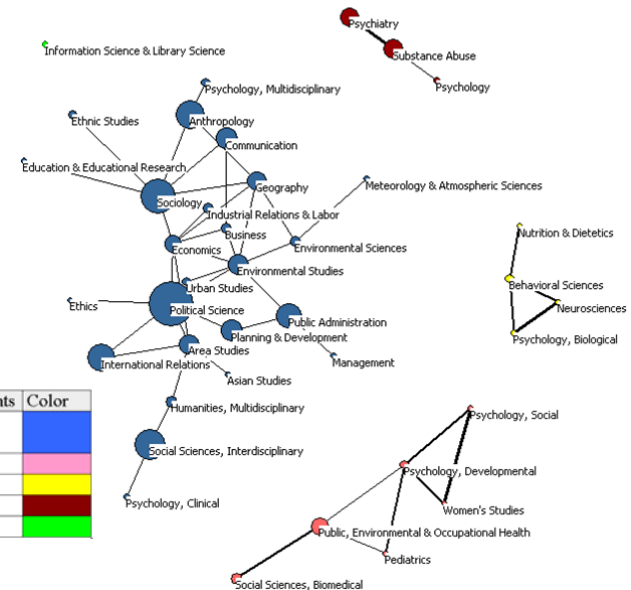
Name	Elements	Color
Public, Environmental Health	14	
Political Science	4	
Biotechnology	2	
Demography and Law	2	
Gerontology	2	
Others (7)	1	



#### Oslo

Disciplines: 39  
Relations: two: 48 / three: 11

Name	Elements	Color
Political Science, Sociology, Social Sciences	25	
Public, Environmental Health	6	
Behavioral Science	4	
Substance Abuse	3	
Others (1)	1	

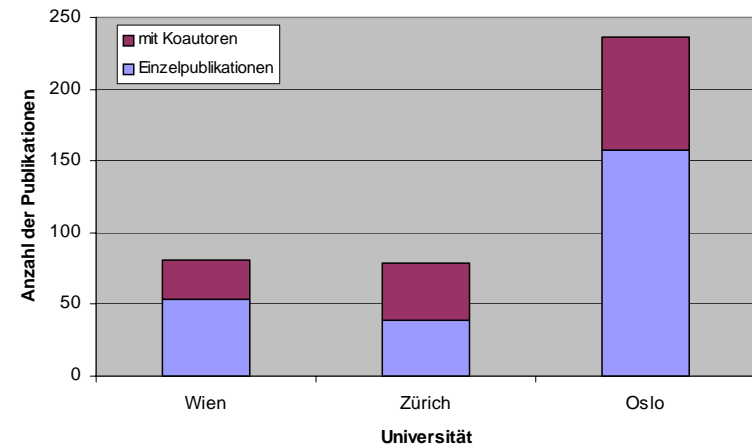




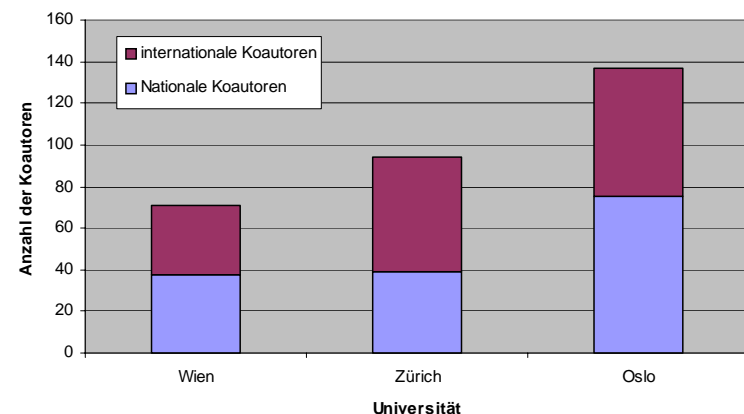
### III. Example: National and international collaboration / co-authorship

- Single author publications:  
Vienna: 67%  
Oslo: 67%  
Zurich: 49%
- Number of co-authors per article:  
Zurich: 1,2  
Vienna: 0,9  
Oslo: 0,6
- Ratio of international co-authors:  
Zurich: 59%  
Vienna: 46%  
Oslo: 45%

Einzel- und Koautorenschaften



Koautoren nach Nationalität

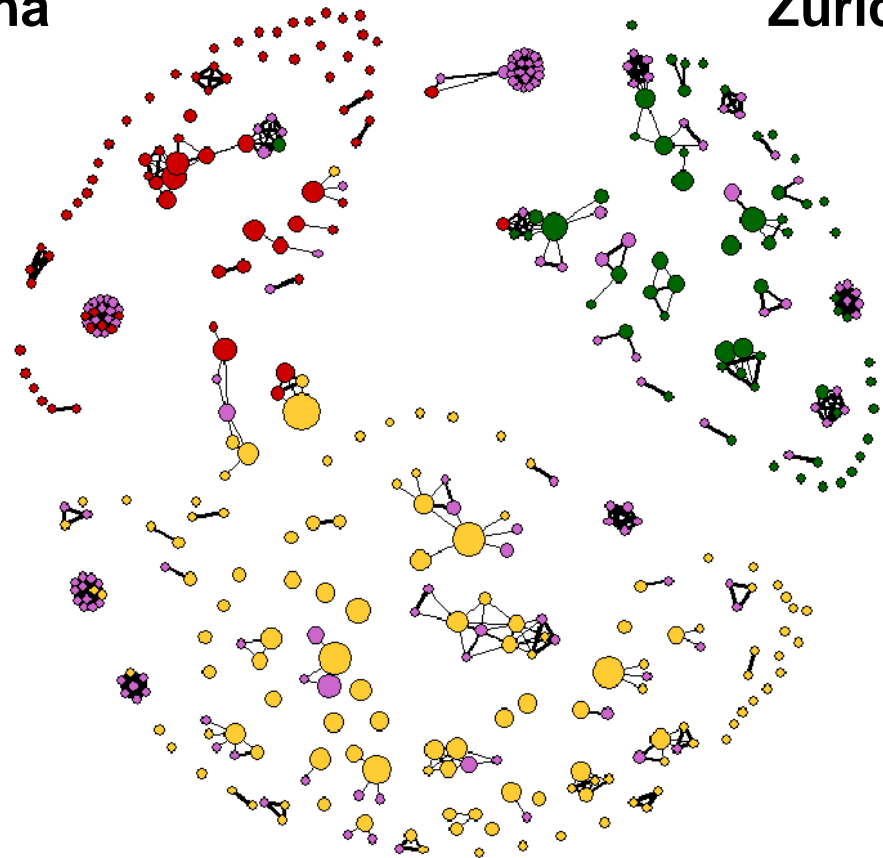


### III. Example: Visualization of co-authors

Name	Elements	Color
international	143	purple
Oslo	125	yellow
Wien	76	red
Zürich	62	green

Vienna

Zurich



Oslo



### **III. Questions related to university specific strategies, general conditions and structures**

1. How are formulation and implementation of exploratory focus reciprocated in publication behavior (co-publications, interdisciplinarity, ...)?
2. Which influence do age pattern and internal structure have on productivity and visibility?
3. Effects of specific publication strategies and incentive systems?
4. Are radical breaks or reforms identifiable (changes in internal structure, change of staff, et al.)?
5. Are changes of publication behavior specific to university?  
Are there parallels to global changes?
6. Loss of visibility due to insufficient / false given affiliation?  
Results on Rankings?



## IV. Procedures for faculty positions (strategies of Vienna University)

### Workflow (Template):

*(0. Joint concept of significant analyses)*

1. Data collection in WoS (resp. other databases)  
Time and citation windows: since first graduation;  
Recency and other micro analyses (eg. last 5 years)
2. Bibliometric analysis (computation of different indicators)
3. Check of excellence: belonging to "Top 5" in selected categories (see below)
4. Presentation and discussion of results



## IV. Procedures for faculty positions: data collection

1. Identification of candidates in WoS  
(index, "Distinct Author Sets", et al.)
2. Export of all likely data to Excel, Access, Endnote files
3. Comparison with publication list presented by candidate
4. a, WoS data not in publication list: further analysis of  
affiliations, co-authors, subject areas  
b, Publications not found in WoS: causation
5. Examination by means of original document



## IV. Procedures for faculty positions: Bibliometric Analysis

Computation of following indicators (expansion planned):

1. Number of papers (overall and in chronological sequence)  
Ratio of different document types
2. Gained citations with / without self citations  
Gained citations per publication with / without self citations  
Number of self citations  
Number of non-cited documents
3. h-Index with / without self citations
4. Sum Impact Factors  
Sum Impact Factors per paper  
Aggregate / Median IF of most paper's subject area





## IV. Procedures for faculty positions: Excellence

Belonging to "Top 5" in selected categories:

1. Number of papers (articles + reviews)
2. Gained citations with / without self citations
3. Gained citations per paper with / without self citations
4. h-Index with / without self citations
5. Sum Impact Factors
6. Sum Impact Factors per paper

In due consideration of scientists not represented adequately in such a manner (eg. young scientists...).



**Thank you very much for your attention**

Wolfgang Mayer  
Vienna University Library  
Bibliometrics Department | eResource Management  
Dr. Karl Lueger-Ring 1  
A 1010 Vienna  
AUSTRIA