

Informatics Research Evaluation

Workgroup:

Carlo Ghezzi

Floriana Esposito

Mauel Hermenegildo

Helene Kirchner

Luke Ong



Evaluating Research



The assessment of research is of **great public interest**:

- for an individual scientist, an assessment can have profound and long-term effects on one's career;
- for a department, it can change prospects for success far into the future;
- for disciplines, a collection of assessments can make the difference between thriving and languishing.

Evaluation can be highly effective in improving research quality and productivity. To achieve the intended effects, research evaluation should follow established principles, benchmarked against appropriate criteria, and sensitive to disciplinary differences.

Informatics Europe Documents

- the 2008 report “Research Evaluation for Computer Science” which developed 10 recommendations, still valid, to use in an assessment process w.r.t. Computer Science;
- the 2013 report on “Department Evaluation”, which proposed an evaluation model based on the self-evaluation methodology;
- a new report on “Informatics Research Evaluation” which focuses on research evaluation performed to assess individual researchers, typically for promotion or hiring.

The Task

To provide Informatics Europe's viewpoint on research evaluation, specific issues have been taken into account:

- the Informatics peculiarities,
- the methods for evaluating the research culture of a discipline which has empirical, methodological and theoretical dimensions,
- the problems concerning the evaluation of impact due to the variability of the population interested into the different subfields,
- the nature itself of bibliometrics and standard impact measurements

The Focus

The main focus is on principles and criteria that should be followed when individual researchers are evaluated for their research activity in the field of Informatics in order to:

- suggest guidelines and best practices to be discussed in the community of Informatics Europe in order to standardize and enrich the variants of the assessment protocols and to propose recommendations to people involved in evaluation committees and funding agencies.
- compare and critically analyze the different main methodologies that national assessment agencies can use when evaluating research in terms of products/single researchers/research groups/institutions.

Outcomes

- ❑ A discussion panel on research evaluation was held as part of the program of the ECSS 2016 in October 26.
- ❑ A first release of a report that provides Informatics Europe's viewpoint on the topic, stressing general principles; it is published in conjunction with ECSS 2017 for discussion.
- ❑ Collection of data about research evaluation efforts in different European countries with the aim of developing a document gathering information about current practices of researchers' evaluation through Europe (and linking it from the short document to make a longer, evolving, on-line report).
- ❑ A publication in ERCIM News in April 2018 in the Section "Research and society" concerning "research evaluation". The section should contain an introduction, a presentation of the I.E. report and 6-10 papers from ERCIM and IE.

What about Informatics ?

- A relatively **young science** which is **rapidly evolving** in close connection with technology.
- An original discipline with roots in mathematics, science, and engineering.
- It is **pervasive** and it results in new interdisciplinary research fields.
- It has a high **societal and economic impact**.
- The outcome of Informatics research is often the creation of **new artifacts**.

Informatics research must be evaluated according to criteria that take into account its specificity.

Quantitative measures of impact are possible, but they may not tell the implied story.

Conferences vs. Journals

The publication culture within Informatics differs from other sciences in the prominent role played by conference publications. Ongoing debate on the value of conference publications:

- ❑ When competing with other disciplines, this publication model needs to be defended (differences across countries).
- ❑ Conference rankings are being established, but are still controversial.
- ❑ Number of conferences has increased dramatically, at the price of overall quality:
 - Too many conferences (and journals) that accept low quality papers
 - Reviewing load has increased, there is less time for reviewing, and reviews are shallow
 - Predatory conferences that accept everything without proper reviews

Conferences vs. Journals

In order to bridge the dichotomy between conferences and journals, new alternatives are now in place that are changing the publication culture:

- ❑ Coupled conferences and journals: this may combine the advantages of timely publication of conferences with the impact tracking of journals
- ❑ Open Archives (like HAL, ArXiv, etc.) give the opportunities to publish first versions and protect intellectual property of new results

How to evaluate the impact of research

- **Bibliometrics** - Numerical impact measurements, such as citation counts, have their place but must never be used as the sole source of evaluation.
- **Artifacts** - To assess impact, artifacts such as software can be as important as publications.
- **Open Science** - It advocates practices such as open access publishing, open data, and open peer review.
- **Awards** - “Best paper award”, “most influential paper award” or “test-of-time award”

Bibliometrics



Are the objective (= quantitative) ways to measure

- the productivity of institutions
- the productivity of a researcher
- the quality of journals suitable?

Ranking all research institutions in a given country may be a necessity for informed political decisions about distribution of public funding.

Very often the criteria used for evaluating the institutions are used (tacitly) in order to evaluate the individual researchers.

this constraints to consider mainly the bibliometric indexes derived from citation counts, often neglecting the content relevance and the quality of the contributions.

IEEE statement (sept. 2013)



- The use of **multiple complementary bibliometric indicators** is fundamentally important to offer an appropriate, comprehensive, and balanced view of each journal in the space of scholarly publications.
- Any **journal-based metric** is not designed to capture qualities of individual papers, and **must therefore not be used as a proxy for single-article quality** or to evaluate individual scientists.
- While bibliometrics may be employed as a source of additional information for quality assessment within a specific area of research, the **primary manner for assessment** of either the scientific quality of a research project or of an individual scientist **should be peer review**

Towards more quality and impact

- The goal of research assessment is primarily to assess quality and impact over quantity

Any policy that tends to favour quantity over quality has potentially disruptive effects and would mislead young researchers with very negative long-term effects.
- Quantitative data and bibliometric indicators must be interpreted in the specific context of the research being evaluated

Human insight is needed to interpret data and discern quality and impact; numbers can only help
- Assessment criteria must themselves undergo assessment and revision

References



"Protocol for research assessment in Informatics, Computer Science and IT Departments and Research Institutes." (2013, Informatics Europe, ed. Manfred Nagl).

"Research Evaluation for Computer Science", Eds. Bertrand Meyer, Christine Choppy, Jan van Leeuwen and Jørgen Staunstrup. (2008, Informatics Europe Report).

"Conferences vs Journal in CS, what to do? Evolutionary ways forward and the ICLP/TPLP Model." Hermenegildo M.V. Position paper for Dagstuhl meeting 12453: Publication Culture in Computing Research. 2012.

"Invisible Work in Standard Bibliometric Evaluation of Computer Science," Jacques Wainer, Cleo Billa, Siome Goldenstein, Communications of the ACM, Vol. 54 No. 5, Pages 141-146, 2011.

"Incentivizing Quality and Impact: Evaluating Scholarship in Hiring, Tenure, and Promotion," by B. Friedman and F.B. Schneider. CRA Best Practice Memo of February 2015.

Evaluation of Research Careers fully acknowledging Open Science Practices — Rewards, Incentives and/or recognition for researchers practicing Open Science. European Commission - Research and Innovation doi:10.2777/75255.
<https://www.acm.org/publications/policies/artifact-review-badging>