

How CL Researchers Should be Evaluated: Make your Voice be Heard!

CICLING 2018, Hanoi, VN
March 19–23, 2018



What is this special event about?

Open, informal chat on how
research in CL / NLP
is currently being evaluated,
and how we all think
it should be



What is this special event about?

- Are you currently being evaluated for hiring or for promotion, and feel that the criteria routinely used should be improved?
- Bibliometrics is now widely used for evaluation purposes, but are the measures being used reasonable?
- Can they be “gamed”?
- How should new forms of publishing (e.g., arxiv) be taken into account?
- Is bibliometrics **too** dominant when evaluating CL researchers?
- CL is at the crossroads of computer science and linguistics, two fields characterized by different publication patterns; how do we feel CL researchers should be evaluated?

I would like to eventually produce a brief report that summarizes the opinions that will emerge.
Let us make our voices be heard!

The Informatics Europe Report

- 2017 Report (draft): “**Informatics Research Evaluation**”, by the Research Evaluation Working Group of **Informatics Europe** (the European Association of Academic and Industrial Research Institutes in the field of Information and Computer Sciences)
- Focusing on **computer science** and its specificities
- Focusing on the evaluation of **individuals**
- Issuing **recommendations for evaluators**



The Informatics Europe Report

- 1 Informatics is an original discipline combining mathematics, science, and engineering. Researcher evaluation must **adapt to its specificity**.



The Informatics Europe Report

- ① Informatics is an original discipline combining mathematics, science, and engineering. Researcher evaluation must **adapt to its specificity**.
- ② A distinctive feature of publication in Informatics is the **importance of selective conferences**. Journals have complementary advantages but do not necessarily carry more prestige. Coupled conferences and journals, where the papers of a conference are published directly in a journal, are a growing trend that promises to bridge the dichotomy between conferences and journals.



The Informatics Europe Report

- ① Informatics is an original discipline combining mathematics, science, and engineering. Researcher evaluation must **adapt to its specificity**.
- ② A distinctive feature of publication in Informatics is the **importance of selective conferences**. Journals have complementary advantages but do not necessarily carry more prestige. Coupled conferences and journals, where the papers of a conference are published directly in a journal, are a growing trend that promises to bridge the dichotomy between conferences and journals.
- ③ **Open archives** and overlay journals are recent innovations in the Informatics publication culture that offer improved tracking in evaluation.

The Informatics Europe Report

- ① To assess impact, **artifacts such as software can be as important as publications**. The evaluation of such artifacts, which is now performed by many conferences (often in the form of software competition), should be encouraged and accepted as a standard component of research assessment. Another important indicator of impact are advances that lead to commercial exploitation or **adoption by industry or standards bodies**.



The Informatics Europe Report

- ① To assess impact, **artifacts such as software can be as important as publications**. The evaluation of such artifacts, which is now performed by many conferences (often in the form of software competition), should be encouraged and accepted as a standard component of research assessment. Another important indicator of impact are advances that lead to commercial exploitation or **adoption by industry or standards bodies**.
- ② Open science and its research evaluation practices are highly relevant to Informatics. Informatics has played a key enabling role in the open science revolution and should remain at its forefront.



The Informatics Europe Report

- ① To assess impact, **artifacts such as software can be as important as publications**. The evaluation of such artifacts, which is now performed by many conferences (often in the form of software competition), should be encouraged and accepted as a standard component of research assessment. Another important indicator of impact are advances that lead to commercial exploitation or **adoption by industry or standards bodies**.
- ② Open science and its research evaluation practices are highly relevant to Informatics. Informatics has played a key enabling role in the open science revolution and should remain at its forefront.
- ③ **Numerical measurements** (such as citation and publication counts) must never be used as the sole evaluation instrument. They **must be filtered through human interpretation**, specifically to avoid errors, and complemented by peer review and assessment of outputs other than publications. In particular, **numerical measurements must not be used to compare researchers across scientific disciplines**, including across subfields of Informatics.

The Informatics Europe Report

- ① The order in which a publication in Informatics lists authors is generally not significant. In the absence of specific indications, it should not serve as a factor in the evaluation of researchers.



The Informatics Europe Report

- ① The order in which a publication in Informatics lists authors is generally not significant. In the absence of specific indications, it should not serve as a factor in the evaluation of researchers.
- ② In assessing publications and citations, the ISI Web of Science is inadequate for most of Informatics and must not be used. Alternatives include Google Scholar, CORE, DBLP, and (potentially) ACM's Digital Library.



The Informatics Europe Report

- ① The order in which a publication in Informatics lists authors is generally not significant. In the absence of specific indications, it should not serve as a factor in the evaluation of researchers.
- ② In assessing publications and citations, the ISI Web of Science is inadequate for most of Informatics and must not be used. Alternatives include Google Scholar, CORE, DBLP, and (potentially) ACM's Digital Library.
- ③ Any evaluation criterion, especially quantitative, must be based on clear, published criteria. Furthermore, assessment criteria must themselves undergo assessment and revision.

Comments?

