

Methodology and Tools for Research: Scientific publishing

Yannick Prié
Polytech Nantes, University of Nantes
Master DMKM, 2015-2016

Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)



- This course "*Methodology and Tools for Research: Scientific Publishing*" by Yannick Prié is licensed CC BY-SA 4.0
- This license covers the general organization of the material, the textual content, the figures, etc. *except where indicated*.
- This license means that you can share and adapt this course, provided you give appropriate credit to the author and distribute your contributions under the same license as the original
 - for more information about this license, see <http://creativecommons.org/licenses/by-sa/4.0/>
- For any comment on this course, do not hesitate to contact me: yannick.prie@univ-nantes.fr or @yprie

Objectives of this course

- Understand the many facets of publishing:
 - Journals, conferences, books
 - Types of publications
 - Publication workflows
 - Economy of publication
- Get an idea on “publication-based” evaluation
 - Impact factor
 - H-index

- Ressources for the course

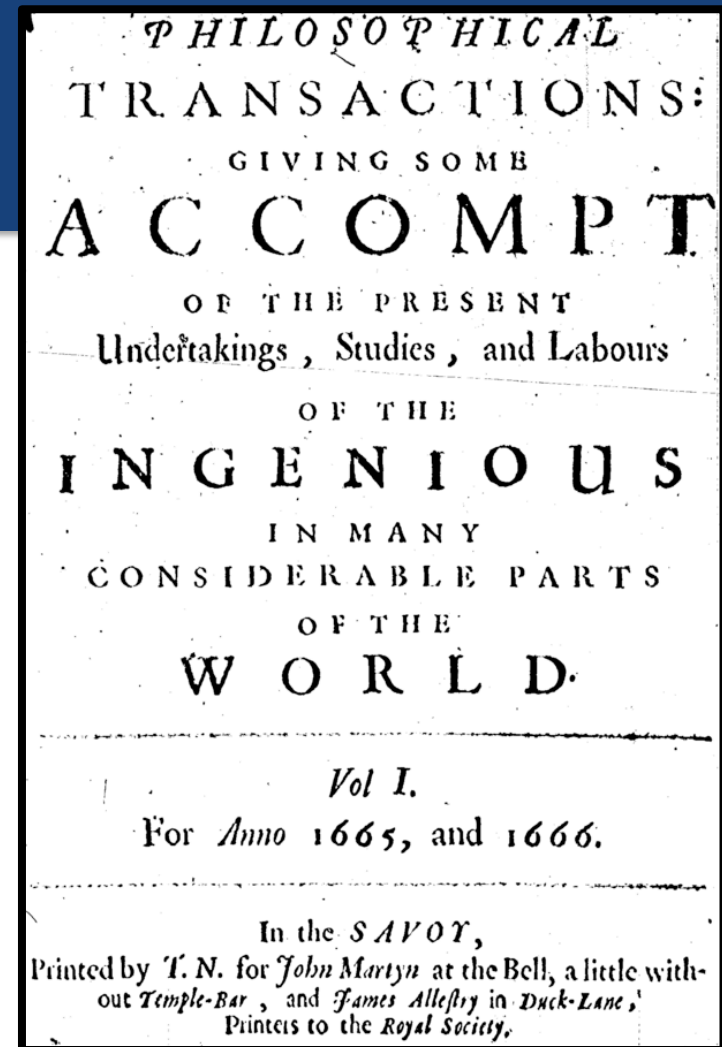
<http://www.scoop.it/t/toolsandmethodologyforresearch>

Historical introduction (I)

- (in the west)
- Since the greeks
 - circulation of knowledge works
 - books, horses, libraries, manual copies (after print, 1456, mechanical copies), only few specific “scientific works”
- Turn of the 16th-17th centuries
 - notions of author, anteriority of discovery
 - e.g. Galileo sends Kepler his encrypted discovery of Jupiter’s satellites
 - organisation of scientific communication
 - 1635: *Academia Parisiensis*
 - Marin Mersenne (1588-1648, monk, philosopher, mathematician): communication with scholars, collect and diffusion of discoveries using postal services
 - 1662: Royal Society of London
 - 1666: Académie des Sciences

Historical introduction (2)

- Academy of Sciences
 - scientific communication are read during meetings
 - articles/minutes are then published by academies
- Scholarly societies
 - idem
- Professional publishers for scientific and medical material
 - because institutions were not that good at publishing
- Periodical journals (19th)
- Exponential growth of scientific material
 - need for means of finding scientific information: databases, abstracting, etc.
- Here: focus on computer science



Title page of Philosophical Transactions
of the Royal Society, Vol. I

by Royal Society is Public Domain

Outline

- Different types of scientific documents
- Principles of publication
- Economics of publishing
- Bibliometrics

Outline

- **Different types of scientific documents**
- Principles of publication
- Economics of publishing
- Bibliometrics

Journal articles

- Oldest and most considered publications in the world of research
 - *Nature*, *Science* (not for computer science!)
 - most journal are focused on a (sub-)discipline
- Important articles which describe mature, solid research and results
 - often the best publications of a researcher
- Various publication rates
 - 1 to 12 issues per year, with *numbers*
 - an issue comprise 4 to 10 articles
 - general or special issues
 - all the issues of the year compose a *volume*
- Computer science publishers
 - Elsevier, Springer, ACM, IEEE, *etc.*



Conference articles

- Article are presented at a conference, and published in the proceedings
- Important in computer science
 - (not in every discipline!)
- Focus on sub-disciplines
 - e.g. ICDM, VLDB, CHI
- Various levels of prestige
 - top level international conferences article as good as journal articles in computer science
 - top researchers in the program committee / attending
 - international and national conferences
 - full (long) or short papers
- Mostly annual

DIS 2012 • In the Wild

June 11–15, 2012 • Newcastle, UK

DIAM : Towards a Model for Describing Appropriation Processes Through the Evolution of Digital Artifacts

Amaury Belin, Yannick Prié
Université de Lyon, CNRS Université Lyon 1,
LIRIS, UMR5205, F-69622
{amaury.belin,yannick.prie}@liris.cnrs.fr

ABSTRACT

Appropriation of technology is a process by which users complete the work of designers by making interactive systems functional within the frame of their situated activities. While existing theories and studies about appropriation are oriented toward the psychological or organizational dimension of this process, we propose a model to describe it through evolutions of digital artifacts and information structures. We also present a case study demonstrating how this model helps to identify particular user operations, and related digital transformations, as a part of the appropriation process. These findings open perspectives to bridge scattered theoretical approaches of appropriation around a low-level, artifact-oriented, and objective way of describing appropriation. Our model could also improve the way appropriation is taken into account in design, by bringing more focus on technical aspects of interactive systems.

Author Keywords
Appropriation of technology, design-for-appropriation, information structures.

ACM Classification Keywords
I5.2 Information interfaces and presentation (e.g., HCI): Theory and methods.

General Terms
Design, Theory, Human Factors.

INTRODUCTION

Appropriation is the global process by which people continuously integrate artifacts into their practices. Appropriation of interactive systems involves a variety of sub-processes, such as customization of tools or repurposing and reconfiguration of practices [4,16]. The appropriation process is clearly revealed when users develop workarounds to adapt a tool to their practices, or when they develop usages that were not anticipated by

designers. This is all the more true for activities carried in open environments, where rapid evolutions of their practices lead users to spend as much time appropriating or re-appropriating artifacts (*constructive dimension of the activity*) as being directly "oriented toward the production of results" (*productive dimension of the activity*) [13].

It is through appropriation processes that a user completes the work of designers by making interactive systems functional within the scope of their situated activity [1]. Appropriability is an important characteristic of an interactive system, and should be thoroughly considered in design. Several studies, guidelines, models and methodologies have been developed in this regard. This body of work builds on various scientific fields, covering multiple dimensions of human activity, from sociology [1] and ethnomethodology [4] to psychology [16] and organization science [2]. Such a variety of origins reveals the complexity and richness of the notions of appropriation, and underlines the situated nature of this process; users considering various dimensions (technical, organizational, social, etc.) in order to integrate artifacts in their own particular practices [4]. These works often focus on only one theoretical dimension of appropriation, and are mostly based on case studies. Therefore, findings about the design of interactive systems are either limited by the specificities of the observed situations (difficulty to provide general and powerful statements about appropriability) or by remaining too abstract or distant from technical concerns (use in design is not straightforward). To sum up, an important gap appears between over-descriptive and theoretically scattered works on the one hand, and pragmatic designers' needs of concepts clearly related to the system they have to design on the other hand [1,3,4,8,15].

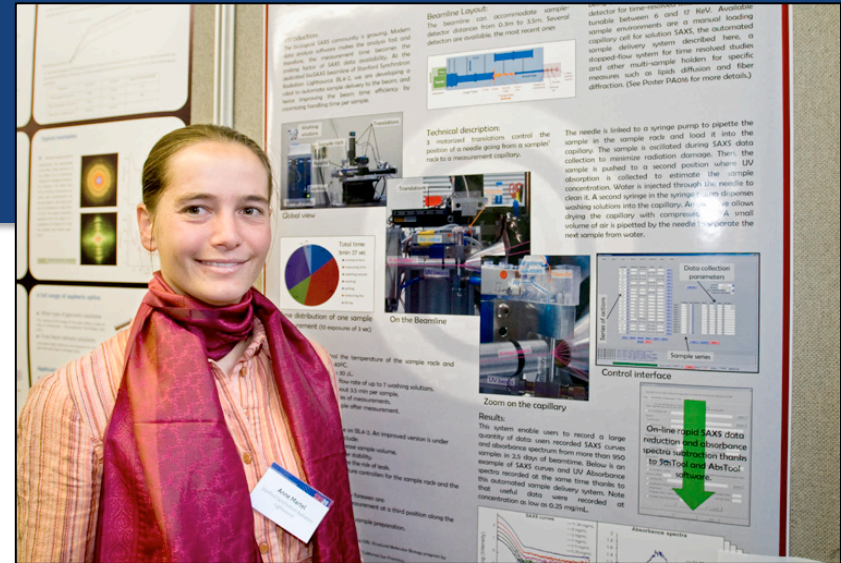
In this paper we propose a step towards bridging this gap by bringing light to the ways digital artifacts, and related information structures, evolve as they are appropriated. For that purpose, we propose a theoretical model articulated around the notion of *digital instruments*, defined as a stabilized functional unit composed of customized artifacts and utilization schemes developed by a user. We show how the appropriation process could be seen as a dynamic co-construction of these two components. We propose the notion of *evolutions of digital structures* to describe the typical co-evolutions of information structures and tool customization that are observable on the artifact side, and

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
DIS 2012, June 11–15, 2012, Newcastle, UK.
Copyright 2012 ACM 978-1-4503-1210-3/12/06...\$10.00.

645

Posters

- Posters are presented in dedicated sessions of conferences
 - several stand-up presentations
- Research result that were not sufficient for publication in the main program
 - not finished
 - only preliminary ideas
 - can be associated to a short paper or abstract in the proceedings
 - (the occasion to attend important conferences without a paper)



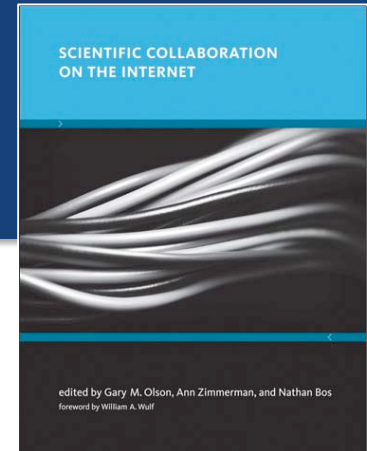
Anne Martel at the poster session by SAS-2009 Oxford is licensed under [CC BY 2.0](https://creativecommons.org/licenses/by/2.0/)

Workshop articles

- Workshops are small conferences focused on dedicated topics
 - aimed at discussing hot subjects in a more informal atmosphere
 - position papers, on-going work
 - key researchers participate to workshops
- Various kinds of workshops
 - 10 to 100 participants
 - with or without proceedings
 - recurring or one-shot
 - open or invitation-only
 - independent or associated to a conference (shared accommodation)
- Workshops can lead to special issues of journals

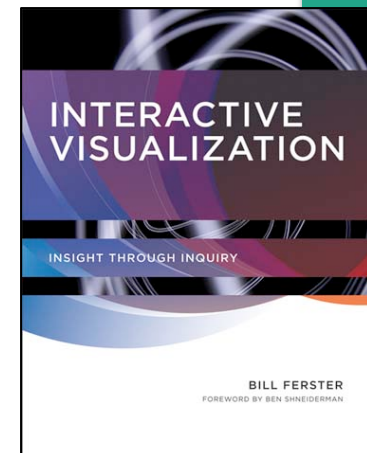
Books and book chapters

- Book: the most ancient mode of disseminating knowledge
 - e.g. dialogues of Plato
- Academic books
 - classical: one or several authors
 - “chapter-based”: one or several editor
 - one or several authors for each chapter
- Various quality
 - various publishers, various book series
 - books or book chapter are generally written upon request
 - not the same evaluation processes as journals
 - books must be sold
 - editorial policy, marketing effect, etc.



edited by Gary M. Olson, Ann Zimmerman, and Nathan Bos
foreword by William A. Wulf

© MIT Press



BILL FERSTER
FOREWORD BY BEN SHNEIDERMAN

© MIT Press

Research reports

- Articles under submission
- Preprints
- Technical report from which articles can be extracted

- A means to declare anteriority: a report has a number and a date, is published by an institution

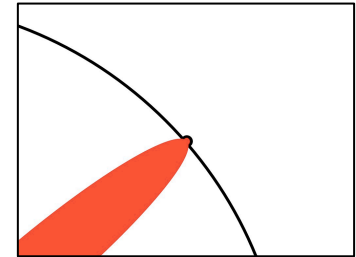
Data, additional material

- Generally associated with articles
 - full results
 - code
 - experimental protocol
 - ...

PhD thesis (and HDR)

<http://matt.might.net/articles/phd-school-in-pictures/>

- PhD Thesis
 - various national systems / various forms
 - describes PhD work and achievements
 - main interest: bibliographical study on a particular topic
- HDR (Habilitation à Diriger des Recherches)
 - French particularity
 - various forms



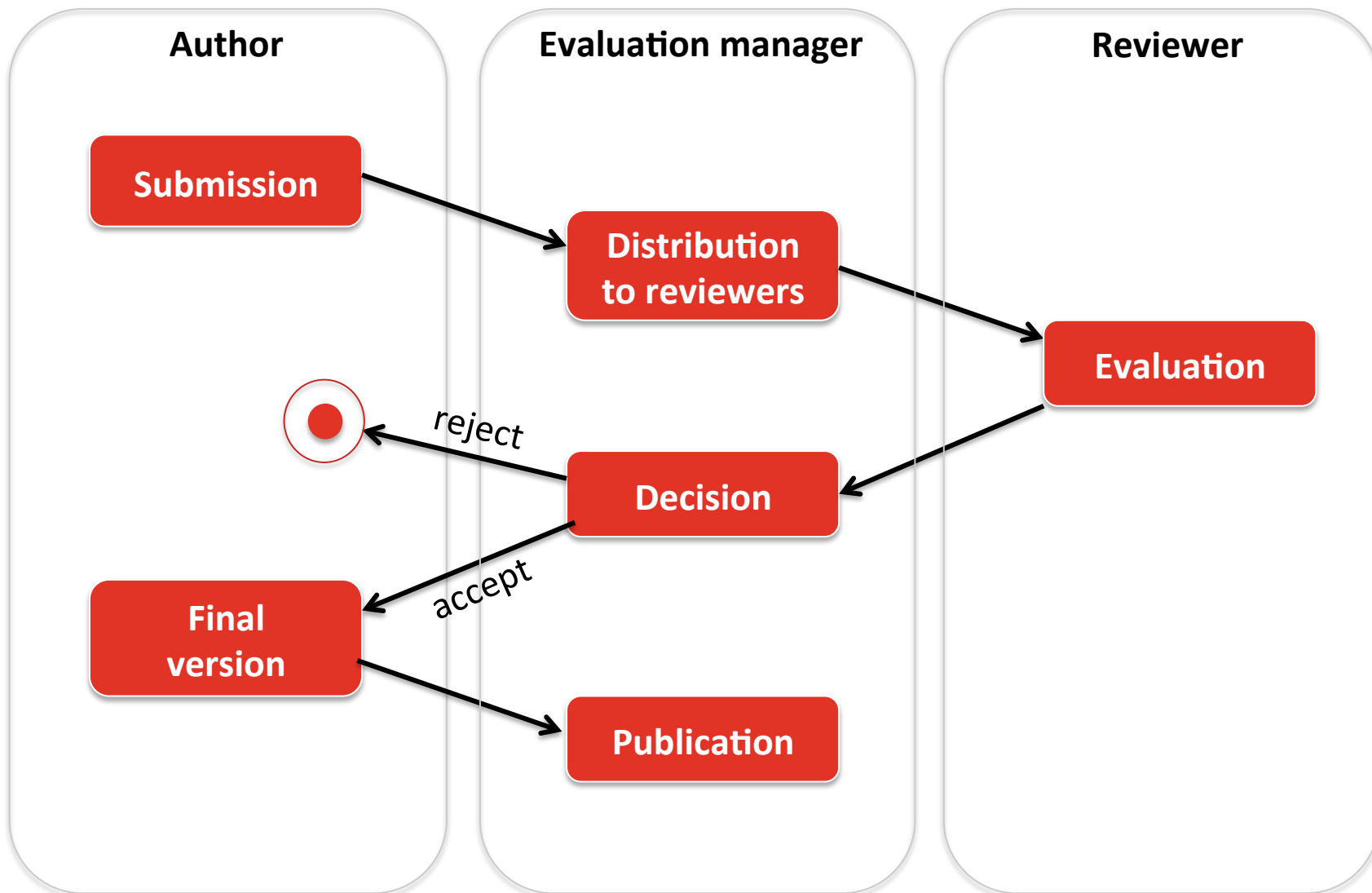
Scientific popularisation material

- Books
- Articles in journals targeted towards the general public
 - Scientific journal
 - Institutional journals (eg. CNRS)
 - Classical journals
- Videos
- Websites
- ...

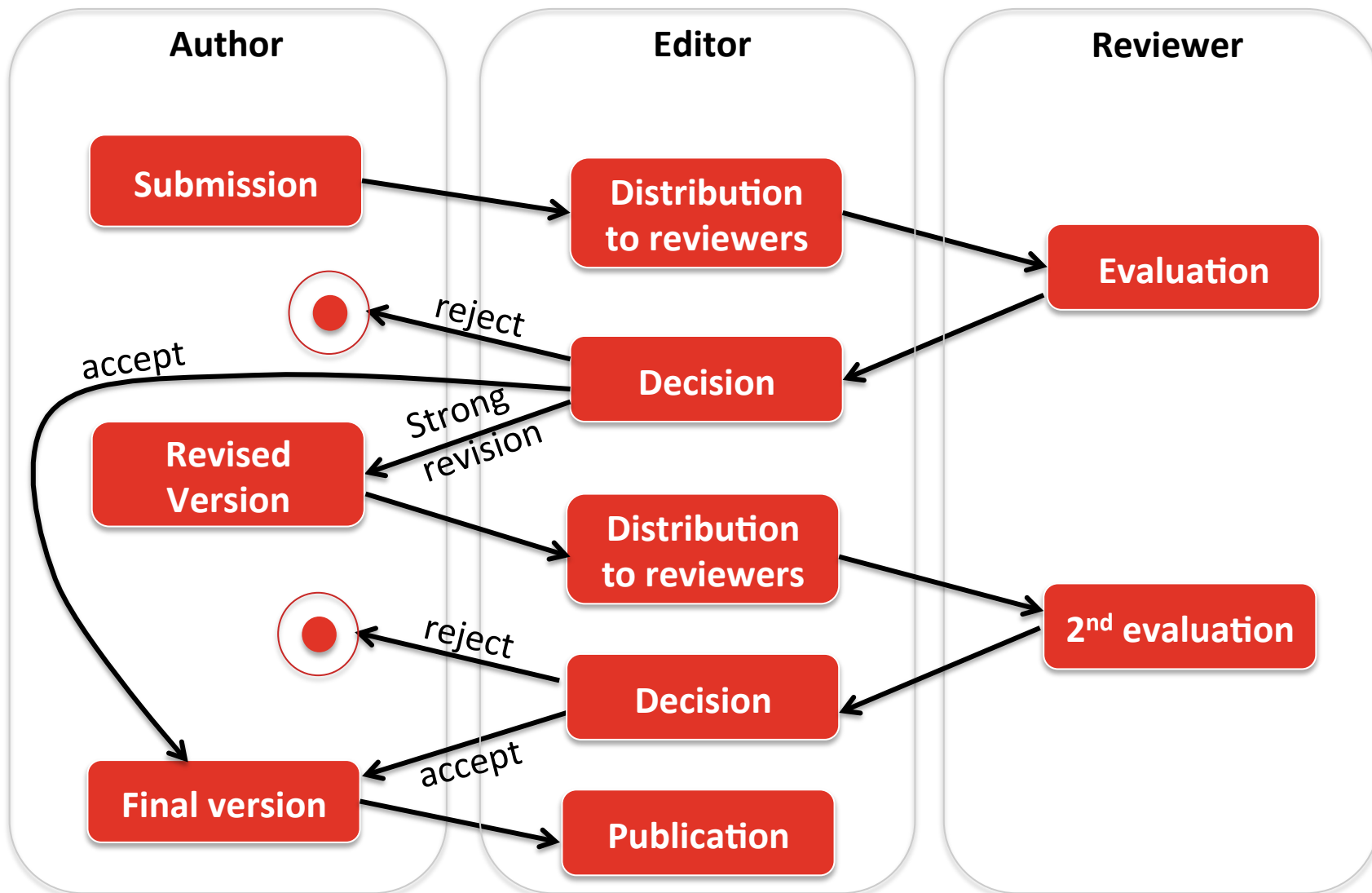
Outline

- Different types of scientific documents
- **Principles of publication**
- Economics of publishing
- Bibliometrics

General workflow



Journal workflow (I)



Journal workflow (2)

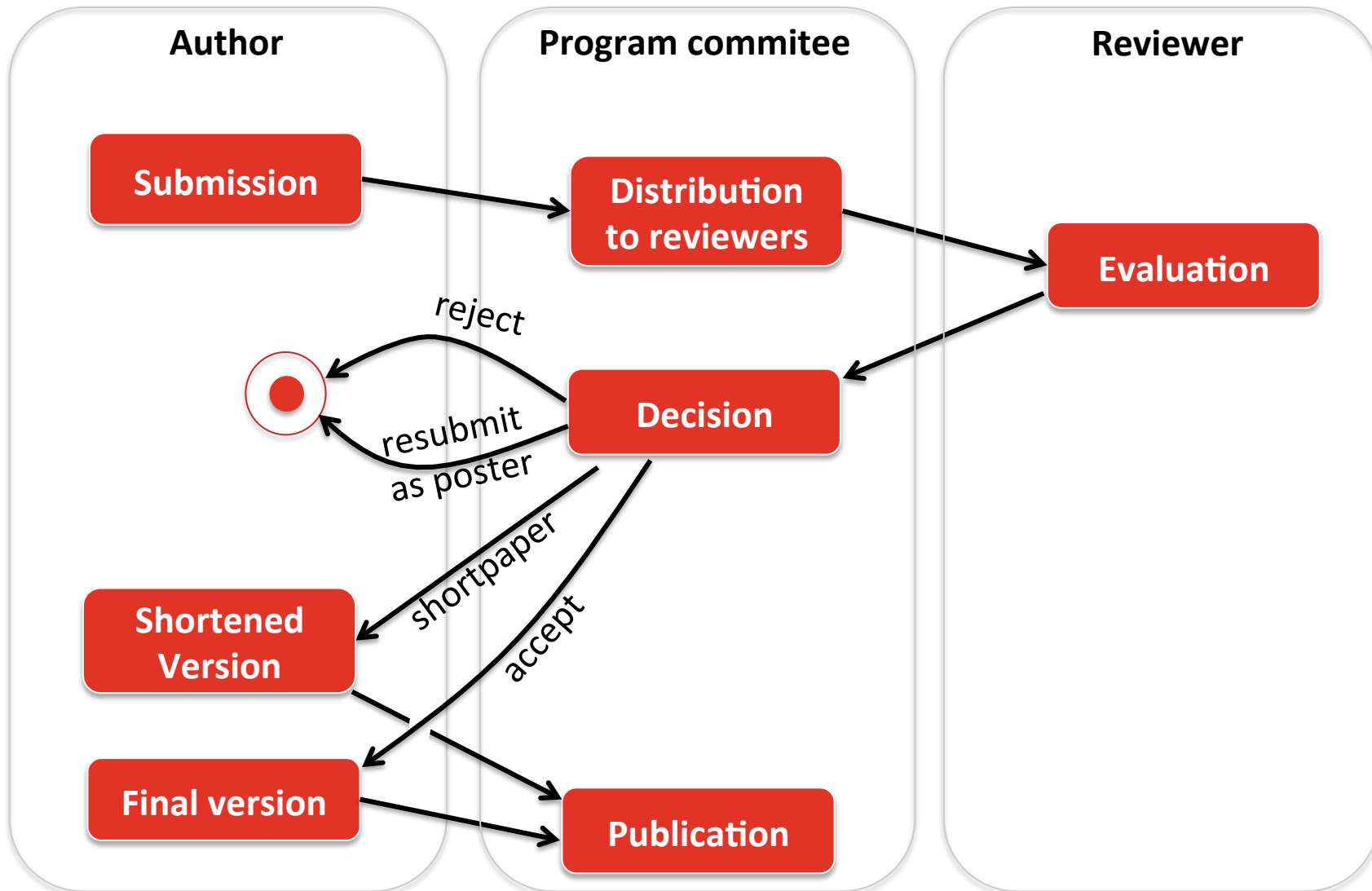
- Authors can get a chance to improve a potential valuable paper
 - possible because evaluation takes times (up to several years)

HCI Editorial Record. First manuscript received December 31, 2008. Revisions received March 8, 2009, and July 1, 2010. Final manuscript received August 29, 2010. Accepted by John Carroll. — *Editor*

Published nov. 2012!

- Major revisions are accompanied with a response to reviewers
 - stating how their highly valuable remarks have been carefully taken into account
- Generally 2 or 3 reviewers, more if they cannot reach an agreement

Conference workflow (I)



Conference workflow (2)

- Full evaluation process takes 4-7 months
- Up to 4 reviewers for highly disputed papers
- Variants
 - poster can be accepted automatically
 - rebuttal: a few days to respond to reviewers before final decision
 - meta-reviewers: members of CP, choose reviewers, write a meta-review
- Abstract-only conferences:
 - acceptance is based on a 1-2 pages abstract, paper is written if accepted
 - in many disciplines (hard science or social science)
 - but NOT in computer science

Getting one's work from one publication to another

- Getting more chance to have one's work read
 - workshop paper → special issue journal paper
 - conference paper → journal paper
 - national conference → international conference
- This is why you could read several time the same paper
- Depend on the sub-discipline's stance on republishing
 - may need serious extension
 - e.g. at least 40%-50% of new material

Reviewing = evaluating a paper

- Giving one's opinion on the value of an article
 - originality (regarding the state of the art)
 - technical quality (soundness, precision)
 - presentation quality (language, clarity, figures)
 - appropriateness to the journal/conference
 - confidence of the reviewer
 - general evaluation, recommendation
- Giving comments on how to improve it
 - Very important!

Example for PLOS journals

PLOS: Public Library of Science

- What are the main claims of the paper and how important are they?
- Are these claims novel?
- Are the claims properly placed in the context of the previous literature?
- Do the results support the claims?
- If a protocol is provided, for example for a randomized controlled trial, are there any important deviations from it?
- Would any other experiments or additional information improve the paper?
- Is this paper outstanding in its discipline?
- Who would find this paper of interest? Why?
- If the paper is considered unsuitable for publication in its present form, does the study itself show sufficient enough potential that the authors should be encouraged to resubmit a revised version?

Outline

- Different types of scientific documents
- Publishing principles
- **Economics of publishing**
- Bibliometrics

Economics of publishing

- **Various jobs**
 - Book and journal publishing
 - Editing, printing, selling (journal > subscription)
 - Conference organisation
 - Editing, (printing), organising
 - Bibliography and ranking
 - Collecting notices, calculating indicators
- **Dominant model**
 - 90% of the editing job is done benevolently by researchers who are state-funded
 - authors give up their copyrights
 - articles are hidden behind pay walls
 - subscriptions are paid by libraries which are state-funded

Publishers

- Big players
 - Springer Verlag, Elsevier, Kluwer, etc.
- Private players associated to universities
 - MIT Press, Oxford University Press, etc.
- National players
 - Lavoisier (Hermès)

Scholarly organisations

- More or less thematic
 - Have members that pay a fee
 - Edit journals
 - Sponsor workgroups (eg. Special Interest Group)
 - Sponsor conferences
 - Give awards
- Big international players in computer science
 - ACM: Association for Computing Machinery
 - good label for conference
 - IEEE: Institute of Electrical and Electronics Engineers
 - careful with IEEE sponsored conferences

Citations indexes

- A necessity with the increase in the number of articles published, even in a sub-discipline
 - ISI – Web of Knowledge (since 1960)
 - Source of the impact factor indicator
 - Owned by Thomson/Reuters
 - Scopus
 - Owned by Elsevier
 - Publishers indexes / digital libraries
 - IEEE Xplore, ACM DL, etc.
 - Recent players
 - Google Scholar, Microsoft Academic Search, CiteSeerX

Spreading one's work anyway

- Publish on the web a quite final version of the work
 - version n-1, preprint
- Get authorization from the publisher
- Use open access (see later)

Outline

- Different types of scientific documents
- Publishing principles
- Economics of publishing
- **Bibliometrics: evaluating the impact of research work**

Articles

- Classification
 - A+, A, B, C
 - Based on the rank of the associated journals / conferences
- Count of citations
 - Measure of interest of the publication
 - “success” of its content (either positive or negative)
 - Base of scientometrics
 - bibliometrics applied to science

Journals: impact factor

- Frequency of citations or the journal's articles
 - average number of citations a paper in a journal gets

$$\text{IF (t)} = \frac{\text{number of citations to articles of the journal (t-1,t-2)}}{\text{number of published articles (t-1,t-2)}}$$

- Journals and citations from *ISI Web of Knowledge*

Conferences: classificationS

- Rating
 - A+, A, B, C + “not in the classification”
- No generally accepted rules
- Classification based on
 - prestige
 - “The premier conference in...”
 - selection rate
 - 5% to 50%
 - durability
 - “First conf.” vs “24th conf.”
 - discipline of the classifier
 - Bias toward

Researchers: h-index

- Goes further than the number of publications: also uses the number of citations
 - *“a scientist has index h if h of his/her N papers have at least h citations each, and the other $(N - h)$ papers have no more than h citations each” (wk)*
 - “h-index = 10” means that there are 10 articles that have been cited more than 10 times
 - can be limited to a recent period (e.g. 5 years)

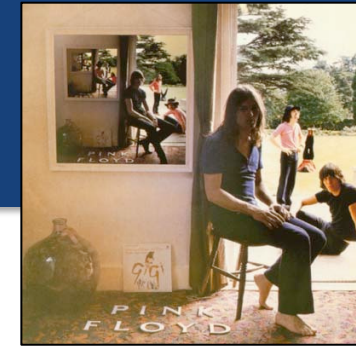
Bias (I)

Indicators are just... indicators

- Indicators are easy to design and calculate
 - it depends on the aims
 - e.g. h-index not adapted to short careers
- Citation number does not directly measure quality
- Impact factor is related to journal, not to article
- H-index is not suited to short careers
- Differences between disciplines
 - ways of citing
 - number of authors
 - journals alone or journals + conferences

Bias (2)

Careful with that h-index, Eugen



[ummagumma - pink floyd 1969](#)
by [Ian Burt](#) is licensed [CC BY 2.0](#)

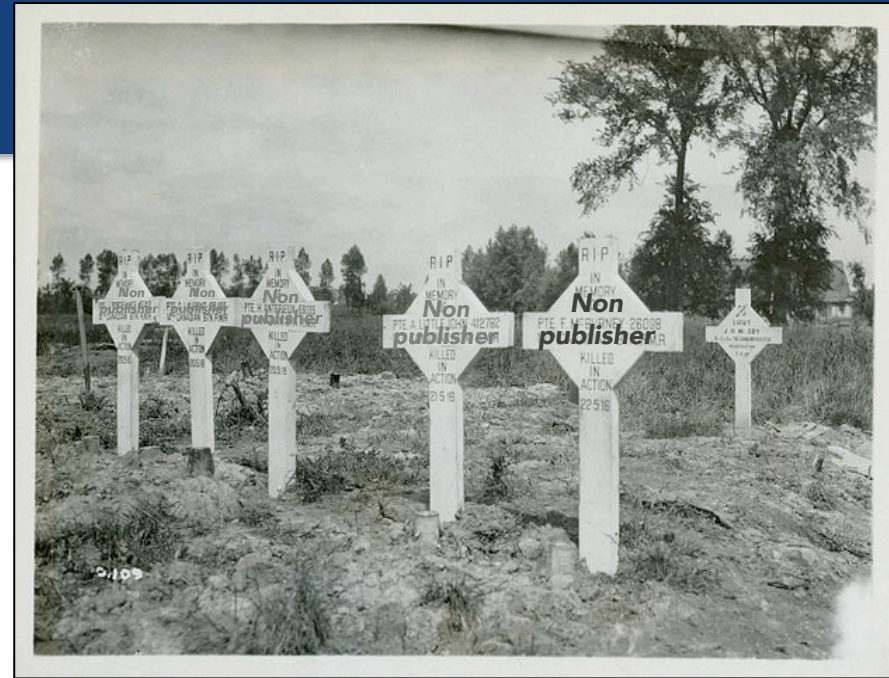
- Indicators can be manipulated
 - h-index: auto-citations
 - impact-factor: e.g. an editorial that cites the best recent papers of the review itself
- Indicators depend upon the organisation that makes the calculation
 - various h-indexes, depending on what articles are counted
 - auto-citations or not
 - only peer-reviewer article vs any pdf on the web
 - but if a student's work cites an article, it is indeed a measure of its influence! (cf. pagerank)
 - IF depends on ISI
 - not all journals are taken into account

Outline

- Different types of scientific documents
- Publishing principles
- Economics of publishing
- Bibliometrics
- **Conclusion**

Publish or perish

- Researcher are evaluated using quantitative indicators
 - even automatically!



Modified from [Canadian Corps - Canadian war graves](#) by [Library and Archives Canada](#) is Public Domain

- This induces dedicated behaviours / strategies
- Mixed with economical considerations in the publishing work

Good strategies

- Target appropriated conferences / journals
- Try to have your papers read
 - disseminate (pdf on the web)
 - do good research
- Help indexing robots
 - good name of organisation

Bad strategies

- Do auto-plagiarism
- Cheat indicators
- Go over conflict of interest
 - e.g. review your friend's papers
- Knowingly publish in “false conferences” or “false journals”
 - <http://www.qualityofconferences.com/>
- Declare false results
 - *“There is increasing concern,” declared epidemiologist John Ioannidis in a highly cited 2005 paper in PLoS Medicine, “that in modern research, false findings may be the majority or even the vast majority of published research claims.”*
http://www.sciencenews.org/view/feature/id/57091/title/Odds_Are,_Its_Wrong
 - See <http://retractionwatch.wordpress.com/>

Why is it bad?

- Because the world of research functions with peer reviewing evaluation
- If the system is cheated, the huge amount of time spend in reviewing is lost
- Trust is a vital necessity



Resistance

- Slow science movement
- San Francisco Declaration on Research Assessment
 - *Putting science into the assessment of research*
- Open science



Annex: PLOS evaluation sheet

- What are the main claims of the paper and how important are they?
- Are these claims novel?
 - If not, please specify papers that weaken the claims to the originality of this one.
- Are the claims properly placed in the context of the previous literature?
- Do the results support the claims?
 - If not, what other evidence is required?
- If a protocol is provided, for example for a randomized controlled trial, are there any important deviations from it?
 - If so, have the authors explained adequately why the deviations occurred?
- Would any other experiments or additional information improve the paper?
 - How much better would the paper be if this extra work was done, and how difficult would such work be to do, or to provide?
- Is this paper outstanding in its discipline?
 - If yes, what makes it outstanding? If not, why not?
- Who would find this paper of interest? Why?
- If the paper is considered unsuitable for publication in its present form,
 - does the study itself show sufficient enough potential that the authors should be encouraged to resubmit a revised version?