Scientific Research in Computer Sciences

Marco Winckler, Hui-Yin Wu, and Teresa Colombi

Université Côte d'Azur (Polytech Nice) | I3S | SPARKS team | bureau 453

Marco.Winckler@univ-cotedazur.fr

http://www.i3s.unice.fr/~winckler/

Objectives

- Present methods and tools for a scientific approach of work
- Help to apply the method in the PER
- Develop capacities for :
  - finding scientific sources of information
  - analyzing scientifically existing solution
  - how to approach scientific literature
  - how to reason and argument using scientific sources of information
  - know how to write technical and scientific documents
  - how to communicate findings to a broad audience (using poster)

Planning

8h classes + hours of autonomous work

4 meetings

Date	Teacher	Room	Contents
Oct 16, 2024	Marco Winckler	C-104	Introduction
Sept 27, 2024	Hui-Yin Wu	C-104	Reading and Bibliography
Dec 11, 2024	Hui-Yin Wu	C-104	Scientific Writing
Jan 15, 2025	Teresa Colombi	A230	Communication Poster

LMS:
EIENSR9 - ECUE Scientific Research in Computer Science
https://lms.univ-cotedazur.fr/2024/course/view.php?id=13894
keyword **INTRO-research-2025**

Contents

- The scientific method :
  - What is it ? Why to use it ? How to make it work ? For whom ? When ?
- How to introduce scientific knowledge in the industrial work
- Sources of information
- How to approach the scientific literature
  - How of read scientific articles scientifically
  - How to make a review of the literature
  - Tools for dealing with references
- How to integrate scientific source in technical documents
- How to communicate results to a broader audience « vulgariser » using a poster

Evaluations: based on PER productions

Review of the literature

Must include:

- Bibliography (in section references)
- Citations to the references (in the main text)
- Citation of sources of images, tables, illustrations, etc. (ask permission for using someone's contents)
- Comparative analysis of references using a scientific method

Criteria:

- Coherence between the DoW and the review of the literature ;
- Writing skills for presenting results;
- Quality of Analysys (“et prise de recul”)
- Quality of Recommendations and conclusions.

Poster:

Must include :

- Identification of PER (author, supervisor, date, title, school, etc.)
- Synthetic presentation of the problem and scope of the study
- Figures of Merit (FoM)
- Comparative analysis and recommendations

Criteria

- Identification of the poster
- Legibility
- Quality of illustrations
- Quality of synthesis and analysis
- Use of language that corresponds to the target audience (**note vulgarization purposes**)

Why Conduct Research?

- To develop knowledge for professions.
- To develop effective policies.
- To solve practical problems.
- To make informed decisions.
- To increase the knowledge base of larger society.

Huge amounts of daily life and experience in our society are based on what we have learned using the logic and evidence involved in scientific research.

BRETAM model of the development of science technology

- Gaines, B.: Modeling and forecasting the information sciences. Inf Sci 57/58: (1999) 13-22

Modeling and Forecasting the Information Sciences

Brian R. Gaines
Knowledge Science Institute
University of Calgary
Alberta, Canada T2N 1N4

Abstract

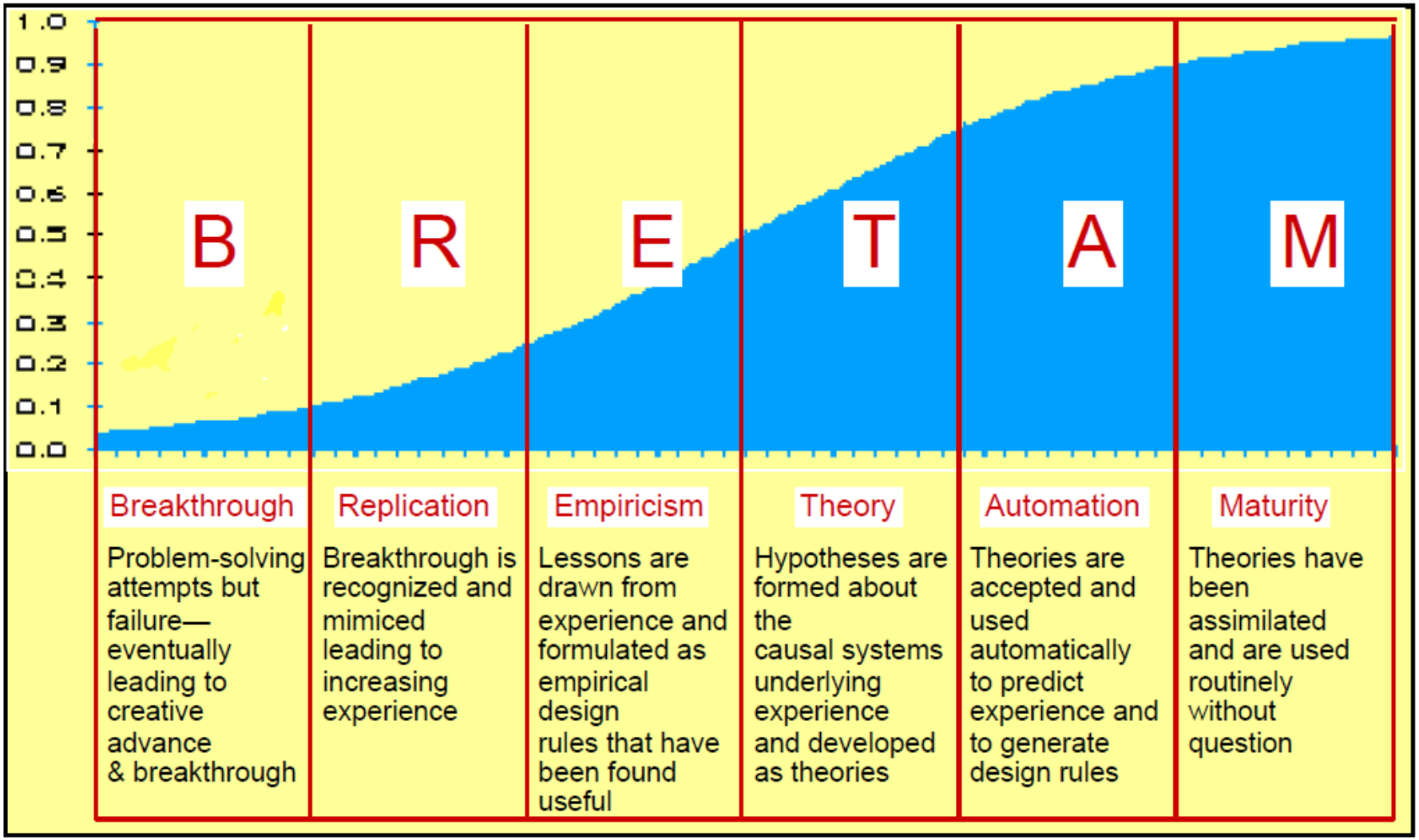
A model of the development of the information sciences is described and used to account for past events and predict future trends, particularly fifth and sixth generation priorities. The information sciences came into prominence as electronic device technology enabled the social need to cope with an increasingly complex world to be satisfied. Underlying all developments in computing is a tiered succession of learning curves which make up the infrastructure of the computing industry. The paper provides a framework for the information sciences based on this logical progression of developments. It links this empirically to key events in the development of computing. It links it theoretically to a model of economic, social, scientific and individual development as related learning processes with a simple phenomenological model. The fifth generation development program with its emphasis on human-computer interaction and artificial intelligence, and the sixth generation research program with its emphasis on knowledge science are natural developments in the foci of attention indicated by the model.

1 Introduction

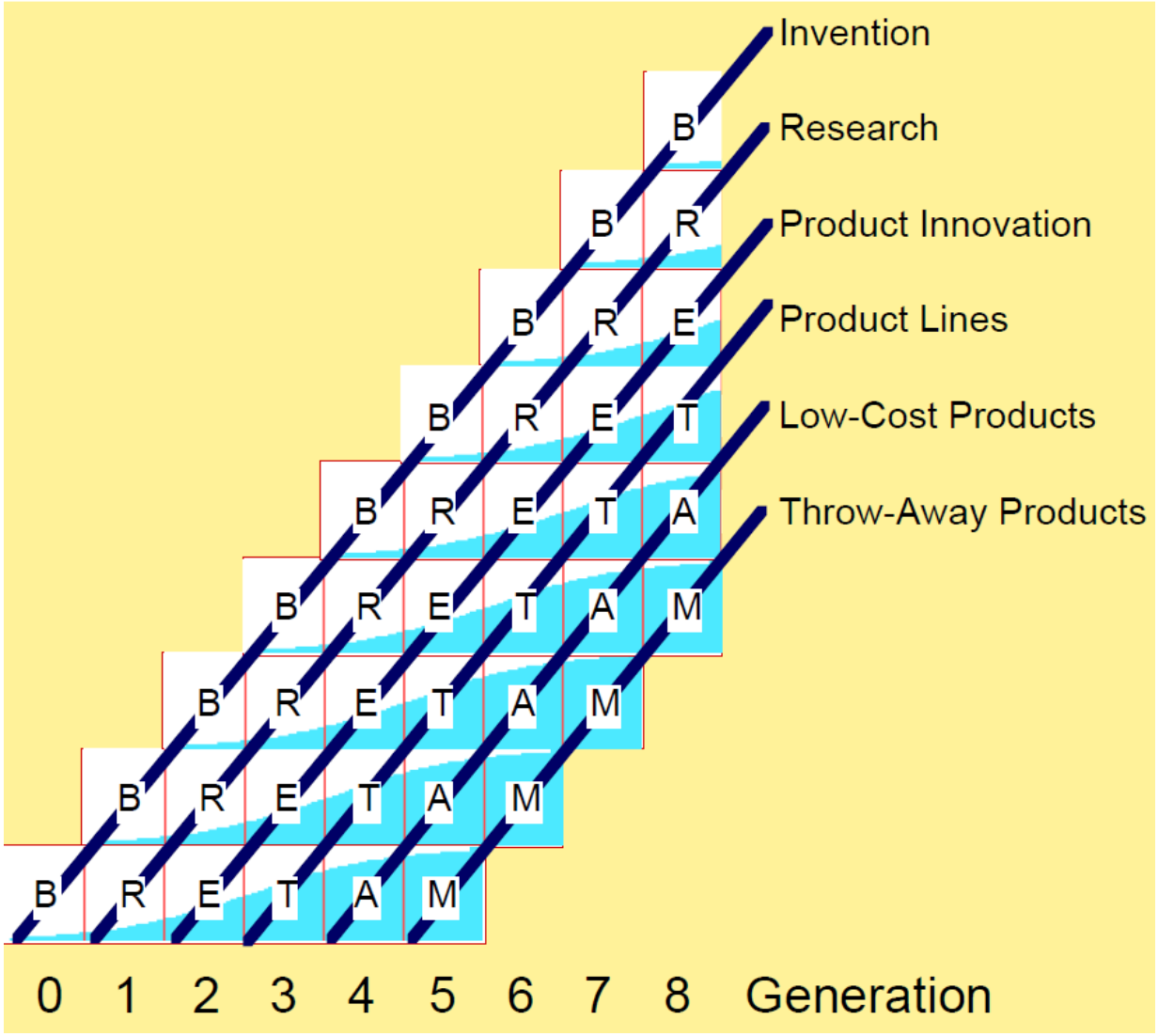
Forecasting advances in technology and their impact has a track record of making fools of the forecasters (Schnaars, 1989). However, the game of life is one of anticipating the future. We model the past that we may learn the lessons of experience, and we extrapolate our models into the future. That the future reserves the right not be anticipated is a meta-lesson that Hume taught many years ago. Nevertheless, our individual behaviors and our civilizations are founded on the assumption that anticipation is possible and, at some level of modeling, history does repeat itself. Our technological civilizations go one step further and change the universe to ratify our anticipations. The information sciences, in particular, structure a wholly artificial reality composed from the ultimate abstractions of the human mind. Paradoxically, they should be readily modeled because they are artifacts of our own mentation, but they may also be beyond modeling because to do so fully may involve ultimate understanding of ourselves.

This paper presents an integrative model of the information sciences that shows them as a tightly coupled, mutually supportive system. It first presents the underlying electronic device technology which provides both the basic support and physical constraints on information technology. It then presents a model of the learning curves of scientific and technological knowledge acquisition that underlies the development of the information sciences. The historic opportunities triggering successive advances in the information sciences are then analyzed and their learning curves superimposed to provide a model of the infrastructure of the information sciences. This model is extrapolated to provide forecasts of future directions, and fitted to various scientific and technological developments. The interactive synergies between levels are analyzed to show the basis of the positive feedback phenomena which continue to support the exponential growth of the information sciences and technologies.

Evolution of domain

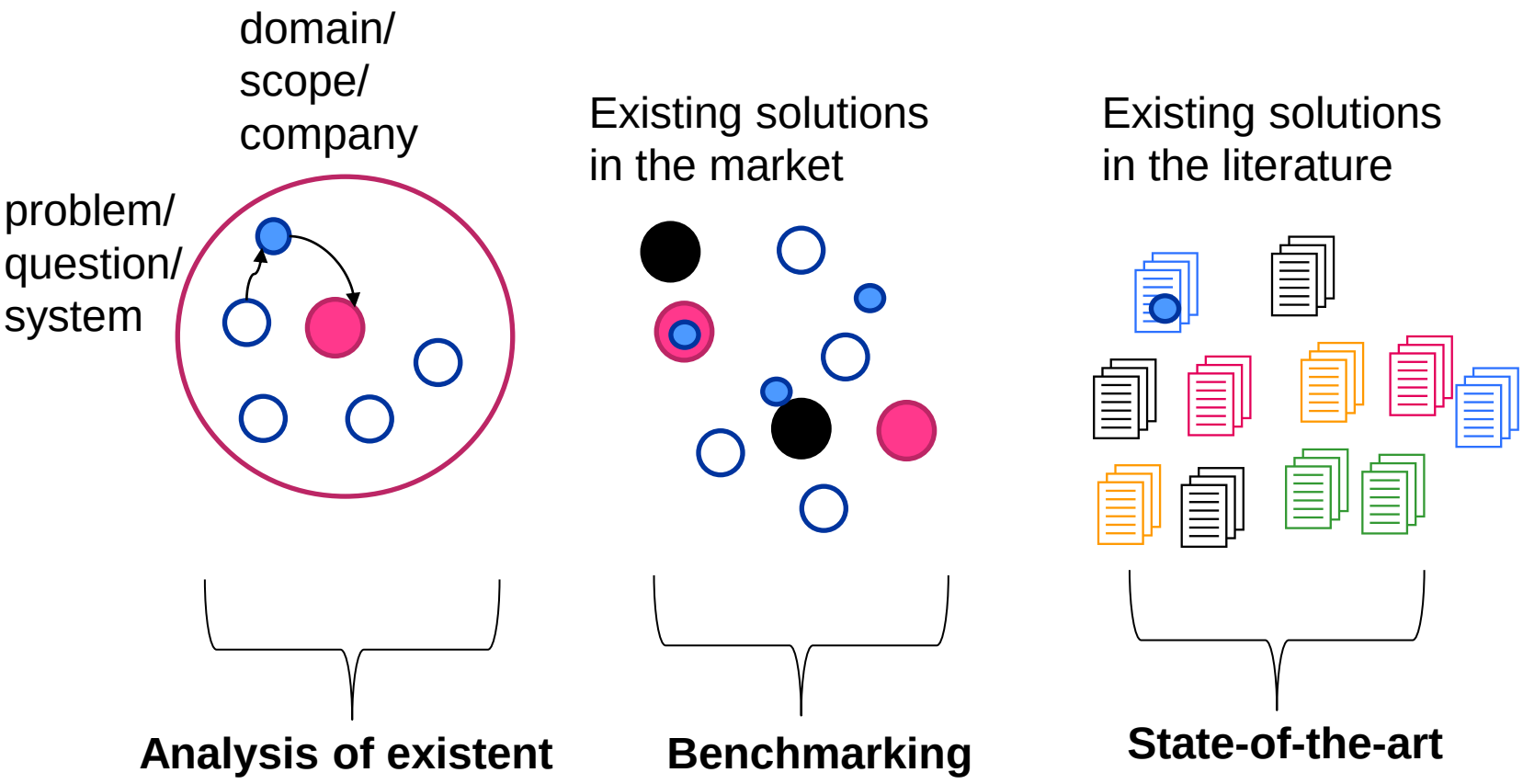


The process ladder



Keep updated

Context for the research



Research Approaches

- **Analyzing existing solutions**
  - process for auditing existing systems
- **Benchmarking**
  - process for comparing existing systems
- **State-of-the-art/review if the literature**
  - process allowing to discover and systematize knowledge

ANALYZING EXISTENT SOLUTIONS (AUDIT)

Questions for helping with the analysis (1/2)

- What is the environment of the organization ?
  - Rules and regulations
  - Strategic/management plan
  - Constraints
- Interactions between systems/actors?
  - What are the actors (people involved)
  - Geographical location
  - Current system architecture
  - Dependencies with other systems

Questions for helping with the analysis (2/2)

- Goals and needs ?
  - What are the indispensable needs without which the system cannot work
  - Points of failure/quality criteria
  - Non critical criteria
- Existing solutions ? (what does work)
  - Main functionalities delovered by the system
  - What are the current system performance
- Point of failures ? (what does not work)
  - What are the dysfunctionning parts of the system
  - The needs that are in the scope of the project

Examples of references commonly used

- Rules and regulations
  - ISO standards…
  - W3C standards…
  - GDPR, EDPB, CNIL https://www.cnil.fr/en/guidelines-and-recommendations
  - Accessibility guidelines ex. https://www.w3.org/TR/WCAG21/
  - Etc.
- Documents that explain the application domain
- Internal documentation
  - Strategic/management plan
  - System specification (before the study)
  - Maintenance and test specification (before the study)

BENCHMARKING

Steps in the benchmarking process

- Identify what you're going to benchmark
- Identify competitors and similar solutions
- Document your current processes.
  - Define criteria to compare solutions
  - Identify trends
- Collect and analyze data.
- Measure your performance against the data you've collected.
- Create a plan
  - Distribute tasks among people
  - Set deadlines
- Monitor results

Components in a benchmarking

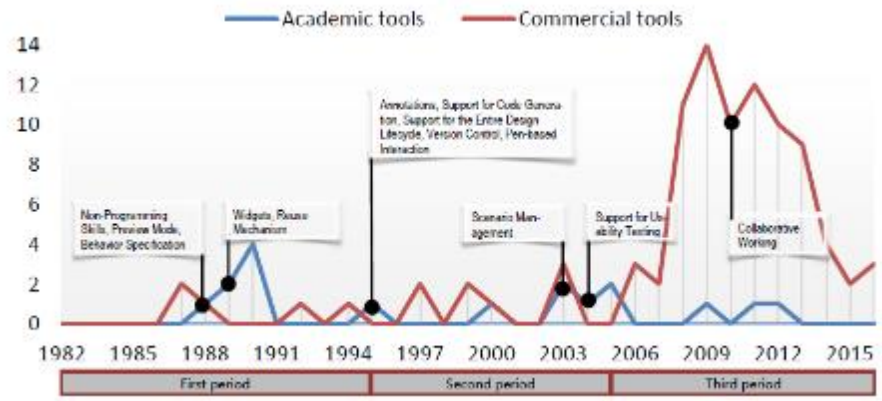
criteria

classification

synthesis

Non-Programming Skills
Pen-Based Interaction
Widgets
Behavior Specification
Collaborative Working
Reuse Mechanism
Scenario Management
Preview Mode
Support for Usability Testing
Support for Code Generation
Version Control
Annotations
Support for the Entire Design Lifecycle

Tool	Reference	Year	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]
iPhoneMockup	http://iphonemockup.lkmc.ch/	?	■		■		■	■							
iRise	http://www.irise.com/	?	■		■	■	■			■			■	■	
JBart	http://www.artwarcsoft.com/#?page=jbart	?			■	■	■					■			■
Mockup Designer	http://fatiherildi.github.io/mockup-designer/	?	■		■		■								
Omnigraffle	https://www.omnigroup.com/omnigraffle	?	■		■	■									
ProcessOn	https://www.processon.com/	?	■		■		■	■					■	■	
Protostrap	http://protostrap.ch/	?	■		■	■		■		■		■			
Serena Prototype Composer	http://www.serena.com/index.php/en/products/other-products/prototype-composer/	?	■		■	■				■					
SoftAndGUI	http://www.softandgui.co.uk/	?	■		■	■		■		■				■	
UXPin	https://www.uxpin.com/	?	■		■	■	■	■	■	■	■		■	■	
Adobe Illustrator	http://www.adobe.com/fr/products/illustrator.html	1987	■	■	■			■							
Microsoft PowerPoint	http://products.office.com/fr-fr/powerpoint	1987	■		■			■	■	■				■	
Adobe Photoshop	http://www.adobe.com/fr/products/photoshop.html	1988	■	■	■			■							



References to use with bechnmarking

- Documents that explain the application domain
- Articles describing the criteria to be used in the classification
- Reference to tools/articles used in the classification

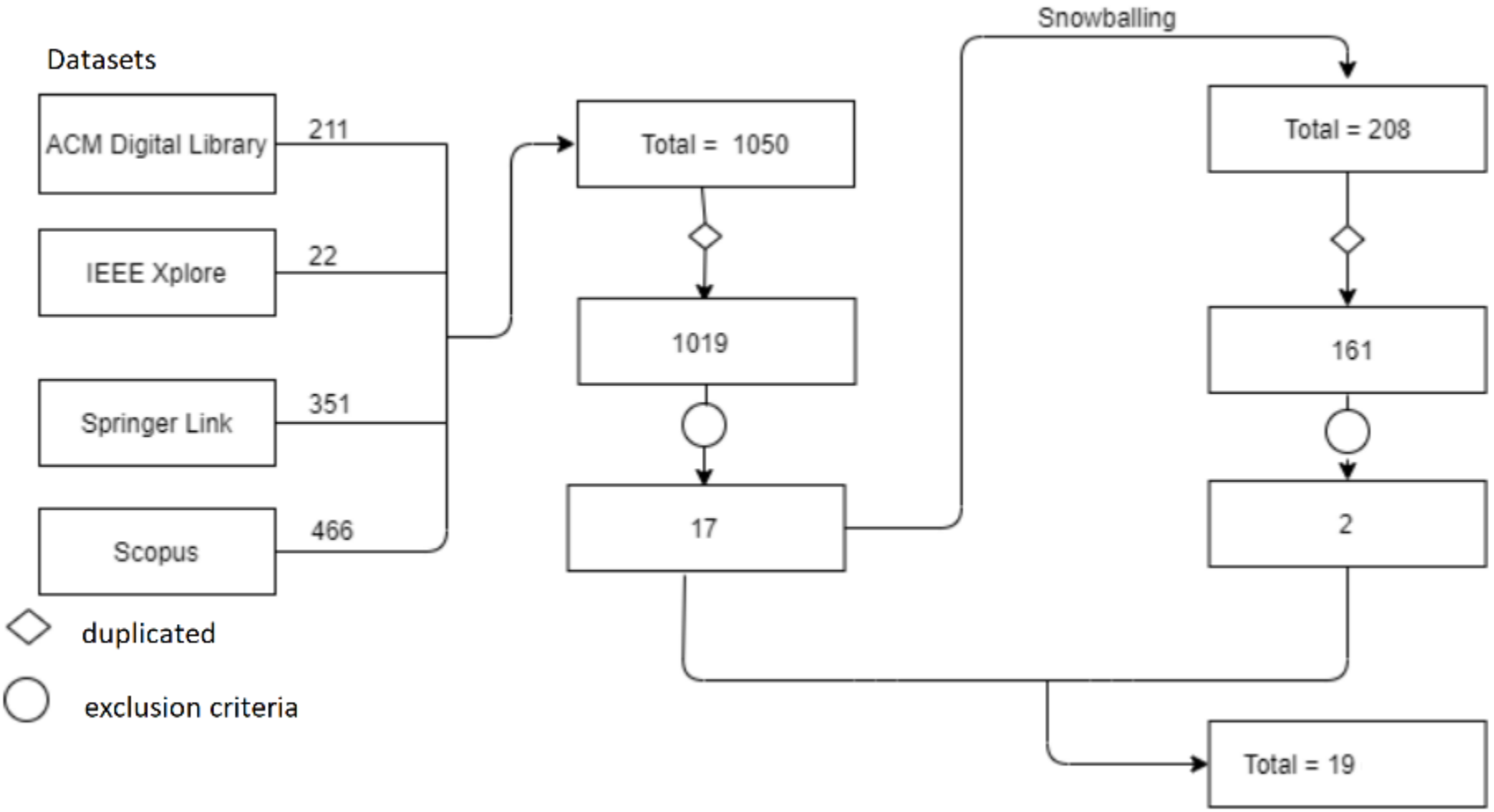
LITERATURE REVIEW

Systematic Literature Review

- Formulate your research questions
- Define the scope (ex. web sites to search)
- Define keywords (in relation with keywords)
- Determine criteria for inclusion/exclusion
  - Ex for inclusion:
    - keywords in the title, abstract or keyword sections
    - Published in the period of …
    - Contribution can be replicated…
  - Ex for exclusion
    - Non peer reviewed papers
    - Private or papers that require payment
    - Papers written in other language than English
    - …
- Snowballing

Overall process for looking for scientific information

Selection process



Selected list

ID	Author	Ref	Base
ID-0	Meshram, Rahul and Kaza, Kesav	IEEE	[13]
ID-1	Nguyen, Minh-Duc and Cho, Yoon-Sik	IEEE	[15]
ID-2	Chong, Sunshine and Abeliuk, AndrÃ©s	IEEE	[5]
ID-7	Liu, Taiheng and He, Zhaoshui and Wang, Peitao	IEEE	[12]
ID-12	Zhou, Diao and Hao, Shengnan and Zhang, Haiyang and Dai, Chenxu and An, Yongli and Ji, Zhanlin and Ganchev, Ivan	IEEE	[29]
ID-13	Yang, Zhenyu and Zhang, Mingge	IEEE	[26]
ID-19	Saranya and Sowmya, A S and Mohammed Shebin, K K and Mohan, Anuraj	IEEE	[19]
ID-20	Shibamoto, Eriko and Kittirojrattana, Chalisa and Koopipat, Chawan and Hansuebsai, Aran and Takano, Kosuke	IEEE	[21]
ID-1052	Xu, Xuhai and Hassan Awadallah, Ahmed and T. Dumais, Susan and Omar, Farheen and Popp, Bogdan and Rounthwaite, Robert and Jahanbakhsh, Farnaz	ACM	[1]
ID-1058	Misztal-Radecka, Joanna and Indurkhya, Bipin	ACM	[14]
ID-1100	Chang, Jianxin and Gao, Chen and Zheng, Yu and Hui, Yiqun and Niu, Yanan and Song, Yang and Jin, Depeng and Li, Yong	ACM	[4]
ID-1540	Jianjun, M.	IEEE	[8]
ID-1566	Guo, S. and Li, C.	IEEE	[7]
ID-1568	Gao, Q. and Ma, P.	Scopus	[18]
ID-1582	Nguyen, T. and Ngo Van, L. and Than, K.	Scopus	[16]
ID-1700	Walek, B.	Scopus	[23]
ID-2445	Widiyaningtyas, Triyanna and Hidayah, Indriana and Adji, Teguh B.	Springer	[24]
ID-15117	Y. Koren . ; R. Bell . ; C. Volinsky. ;	IEEE	[11]
15123	L. Zheng . ; V. Noroozi . ; P. S. Yu. ;	ACM	[28]

Classification

	Q1	Q1.1	Q1.2	Q1.3
ID	Descrição	Técnicas	Modelos	Dynamicity
0	Específico	CB	Processo de decisão de Markov parcialmente observável (POMDP)	Comportamento, Tempo
1	Genérico	CB e CF	Modelo híbrido generativo baseado em LDA	Comportamento
2	Genérico	CF	Correlação de Pearson e fatoração de matriz (MF)	Comportamento
7	Genérico	CF	Modelo de fator latente (LFM)	Comportamento
12	Genérico	CF	Matriz Filtragem com gradiente descendente; Tecnologia de aprendizado profundo (DL) ;	Comportamento
13	Genérico	CB e CF	ON-LSTM em conjunto	Comportamento
19	Específico	CB	Rede heterogênea de itens de usuário (HUI) e similaridade de Cosseno;	Tempo; comportamento
20	Específico	CF	Não identificado	Comportamento
1052	Específico	CB e CF	Uso de heurística post-hoc para identificar as explicações	Contexto; Tempo; Comportamento;
1058	Específico	CF	Algoritimo Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN).	Comportamento
1100	Específico	CB	Não identificado	Comportamento
1540	Genérico	CF	Algoritmo de filtragem colaborativa e coeficiente de Pearson;	Tempo, Comportamento
1566	Específico	CB e CF	Algoritmo K-means e Funk-SVD e algoritmo de recomendação híbrido KMFSCF;	Comportamento
1568	Específico	CB	Uso CA-GNN	Contexto; Comportamento;
1582	Especícico	CB	Não identificado	Comportamento; tempo
1700	Genérico	CB	Não identificado	Comportamento
2445	Específico	CB e CF	UPCSim algoritmo,	Comportamento
15117	Específico	CB e CF	Matriz para controlar a mudança temporal	Tempo, Comportamento
15123	Genérico	CB e CF	DeepCoNN	Comportamento;

Sources of scientific information

Types of scientific articles (1/4)

- Book and Book chapters
  - Presentation and discussions on a topic making part of a big picture (the whole book)
  - Each chapter might be written by an independent author
- Journal publications
  - Focused on extensive discussion of original results

  or
  - Survey of state-of-the-art: comprehensive review of the literature
- News Articles / Magazines
  - Provide a brief overview of scientific findings for a general audience
  - Useful when you are searching for the latest scientific research

Peer Review Process

WILEY

Common criteria selection

- Fit for the target audience/scope of journal/conference
- Originality of the work
- Quality of the results
- Review of previous work
  - State-of-the-art, related work, previous work, benchmark, etc.
  - Proper citation of relevant work
- Quality of the text
  - Language and structure
- Format of the paper

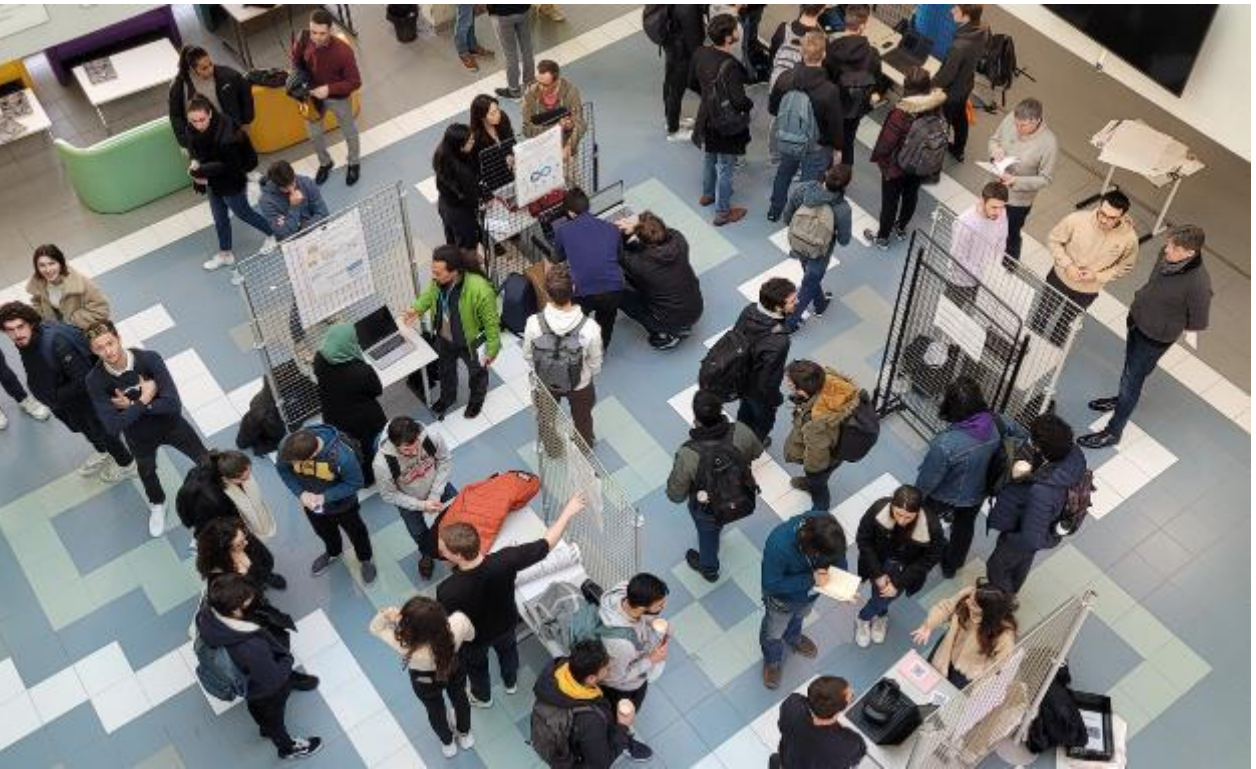
Types of scientific articles (2/4)

- Articles in conferences
  - Original research presented at conferences
  - Useful when you are looking for recent research in the scientific literature
  - (main) Types of contributions:
    - Keynotes (invited papers from outstanding senior researchers)
    - Full papers (long in extension, more prestigious, high selection)
    - Short Papers/Late breaking results (short contributions)
    - Poster/demonstrations (maybe part of the main / adjunct proceedings)
    - …

Plenary session

Types of scientific articles (3/4)

- Poster sessions
  - More informal way of approaching the public
  - The poster is a catcher for attacking people passing by
  - Common format for professional forums as well



Types of scientific articles (4/4)

- Workshop papers
  - Brief descriptions of ongoing work
  - Some workshops are like small conferences whilst other workshops are focused on joint work and only require a position paper
  - Position paper: a brief statement (1-2 pages) describing an opinion

Sources of scientific information

- Attending conferences, symposiums, workshops, … direct access to the book proceedings
- Digital libraries
  - ACM Digital Library: http://dl.acm.org/
  - Springer Link: https://link.springer.com/
  - IEEExplorer: https://ieeexplore.ieee.org/
  - HAL Archives Ouvertes (France): https://hal.archives-ouvertes.fr
- Collaborative Platforms
  - Interaction Design Fondation: https://www.interaction-design.org/
  - Research Gate https://www.researchgate.net/
  - Mendely: https://www.mendeley.com/

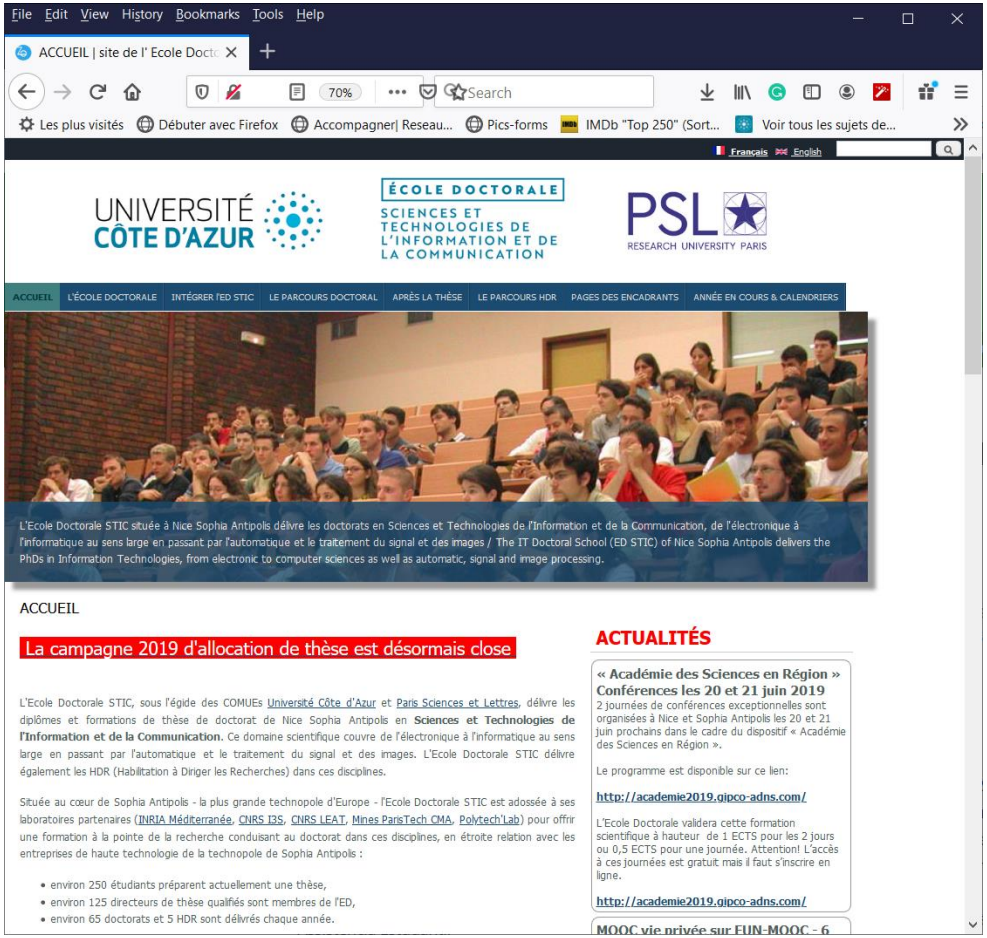
PhD Education

- BAC+8 Diploma
- World-wide recognized as the highest level of education
- ~3 years training on scientific research
- Participation of conferences, travels in perspective
- Some teaching activates might be required
  - Depends on the scholarship
- Final assessment based on a manuscript, the PhD thesis
  - Assessed by a jury of expert in the domain

PhD

- At a Doctoral School/Ecole Doctorale
  - Ex. STIC
- A few classes during the PhD
  - usually for professional matters
- Supervision by an skilled researcher
  - HDR/Full Professor
  - Co-supervision is possible
- Students receive an scholarship/salary during the studies

Ecole Doctoral STIC



http://edstic.i3s.unice.fr/

Source of funding for a PhD

- Ministère de l'Education (voir école doctorale)
- Bourse Région (ex. PACA)
- Research projects (ex. ANR, EU, etc.)
- Convention CIFRE
  - Cooperation 1 university and 1 enterprise
  - Kind of similar PhD in "alternance"
  - Ex. Conventions CIFRE
    - http://www.anrt.asso.fr/fr/cifre-7843

How to find a PhD thesis ?

- Offer at the “Ecole Doctoral”
  - Mandatory for scholarship Region and from Education Nationale
- Monitor mailing lists
  - Ex. annonces@afihm.org, seworld@sigsoft.org, etc.
- Contact supervisors
  - Look at the Web pages for offers of PhD thesis
  - Write an email sending the CV
  - Pay attention to the records of the supervisors such as experience with projects, students' supervision, publications, etc…

Job careers for PhDs at Academia

- Post-doc positions
- Permanent positions with teaching duties
  - *Teaching correspond to part time of the duties*
  - Associate professor (*Maître de Conferênce*)
  - Full Professor (*Professeur des Universités*)
- Full-time research in scientific institutions
  - Ex. CNRS, INRIA, INRA, etc.
  - No official teaching duties but teaching might be possible

Job careers for PhDs at Industry

- Business and Senior developer
- Research & Development (DR) at the industry
- Project leader
- Scientific expert

- Start-ups!
  - Always looking for innovative minds ☺

Why a PhDs thesis relevant in the Industry

- Good entry point for innovation !
- Possibility of participate in EU and National projects (ex. ANR) for building cooperation with the industry
- European and national projects
  - Consortium of many universities and enterprises
  - Goal is to foster technological transfer to the industry

Research-Oriented Internship

- Local labs:
  - I3S: https://www.i3s.unice.fr/fr/
  - INRIA: https://www.inria.fr/fr/centre-inria-universite-cote-azur
  - LEAT: https://leat.univ-cotedazur.fr/
- Funding
  - ~650€ / month
  - **Open call** for DS4H/Polytech students. Deadline November 11. Start by finding a supervisor.

Exercise

- Identify your topic of research (object of study)
- Identify which approach of research (analysis audit, benchmarking, literature review) best fits your PER.
- Define the scope of your work (domain, research question, public, etc.)
- Define exclusion and inclusion criteria
- Select keywords
- Search keywords in the following datasets
  - DBLP.org ACM DL IEEE Explorer Google scholars
- Find at least five relevant papers
- Create a bibliography dataset
  - Add info citation paper and category (ex; conf./workshop/journal, etc.)
- Start reading the papers and bring them for the next class