Andreas Holzinger

VO 706.996 W DiplomandInnen Seminar

VO 706.997 W DissertantInnen Seminar

Welcome to the Holzinger Group HCI-KDD

Part 4: How do we contribute to the international scientific community

a.holzinger@hci-kdd.org

Web: http://hci-kdd.org/scientific-working-for-students

Please do not forget:

- The “best” is the enemy of the “good” – whenever you try to be “perfect” – there is the danger that you finalize nothing*) …”

*) zero, nada, null

François-Marie Arouet (1694 – 1778)
known as “Voltaire”

Motto of the Holzinger Group

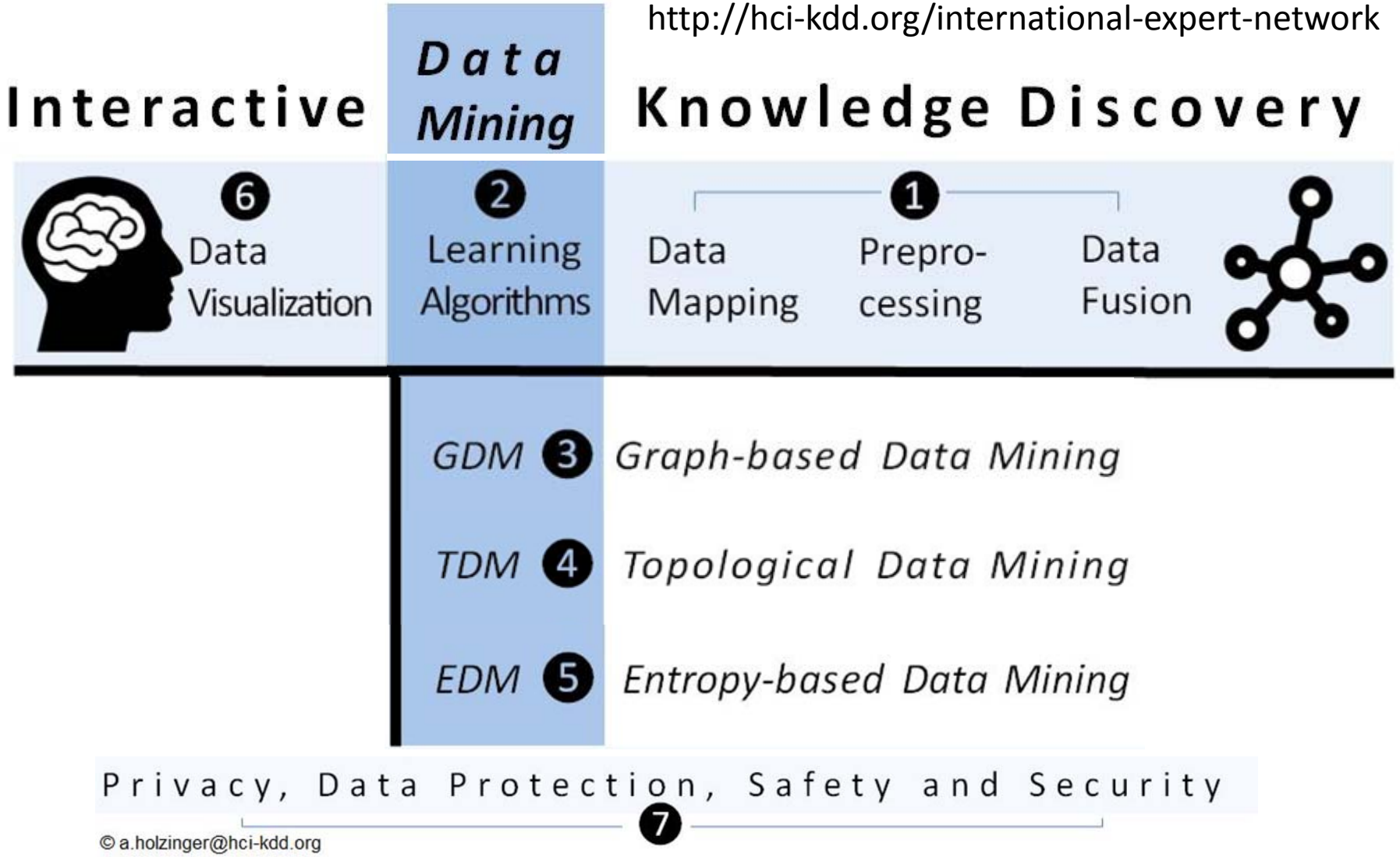
Science is to test crazy ideas – Engineering is put these ideas into Business!

Learning Goals

- At the end of this seminar you should
- … be familiar with the formal requirements
- … be aware of the requirements for a PhD
- … know the HCI-KDD approach
- … have an overview on our research topics
- … understand what research is
- … getting started with your work
- … understand how to write a paper

ML needs a concerted effort fostering integrated research

http://hci-kdd.org/international-expert-network



Holzinger, A. 2014. Trends in Interactive Knowledge Discovery for Personalized Medicine: **Cognitive Science meets Machine Learning.** IEEE Intelligent Informatics Bulletin, 15, (1), 6-14.

Summary “Mantel-PhD”´(cumulative PhD)

- 180 + 120 = 300 ECTS min. for **PhD admission**
- Duration: **36 months** regular (max. 48 months)
- Classes over the 36 months: **14 course hours***)
  e.g. this course LV 706.997 has one hour, LV 706.315 has two hours, 706.046 has three hours
- **PhD-Plan** for the 36 months with deliverables
- Deliverables in form of **papers** to international peer-reviewed conferences/journals
- **Recommended output:** 2 papers per year (e.g. one conference and one journal)
- **Minimum: 1 paper per year**

*) in German: Semesterwochenstunden (semester course hours), there are no ECTS necessary!

Friday 14 October lecture on variational inference.

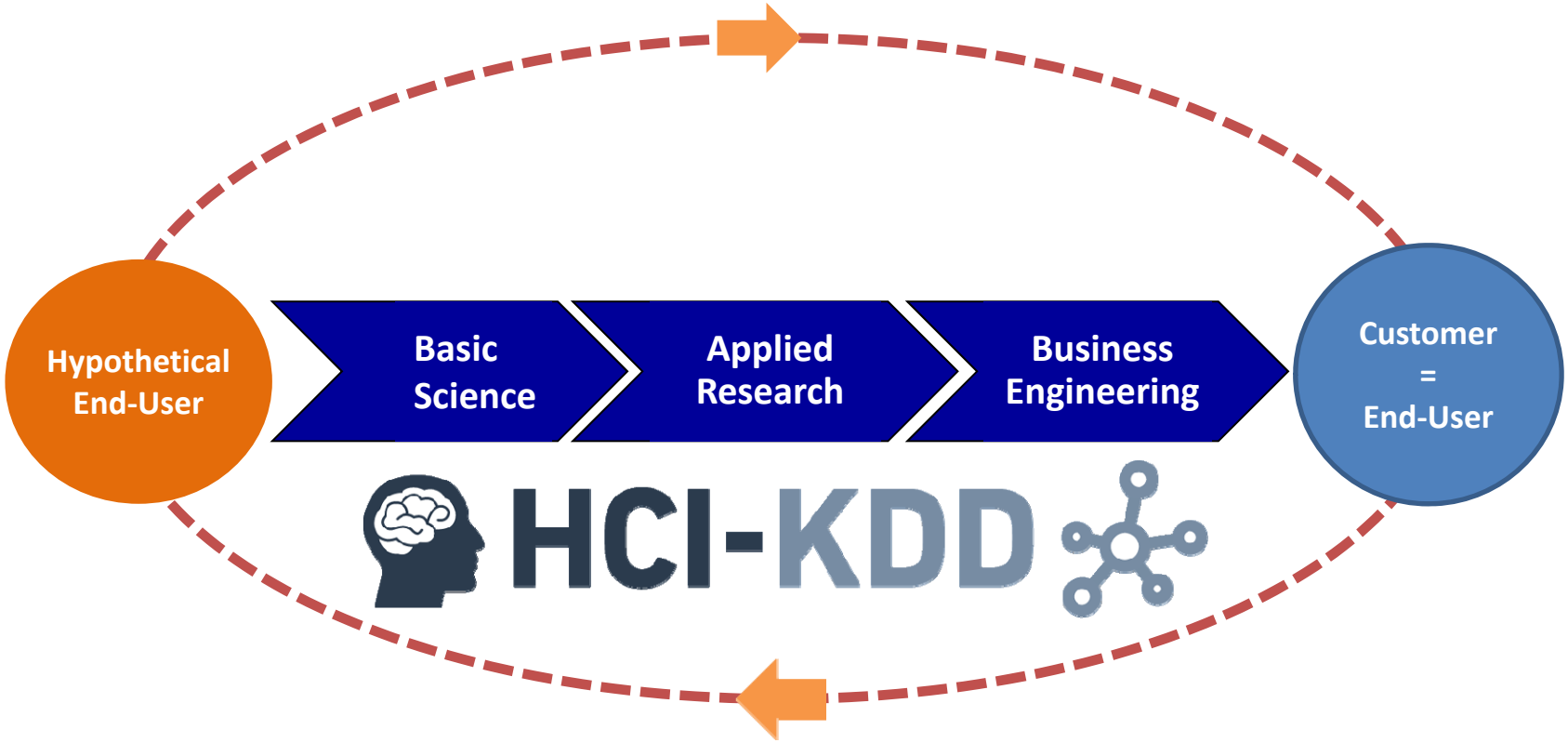
Final Project

In the second half of the course, you will complete a project. The ideal outcome of this project would be a paper that could be submitted to a top-tier machine learning conference such as NIPS, ICML, UAI, AISTATS, or KDD. There are different ways to approach this project, which are discussed in a more comprehensive document that is available from the course website under the Files tab. There are four separate components of the project:

… to contribute to the international scientific community

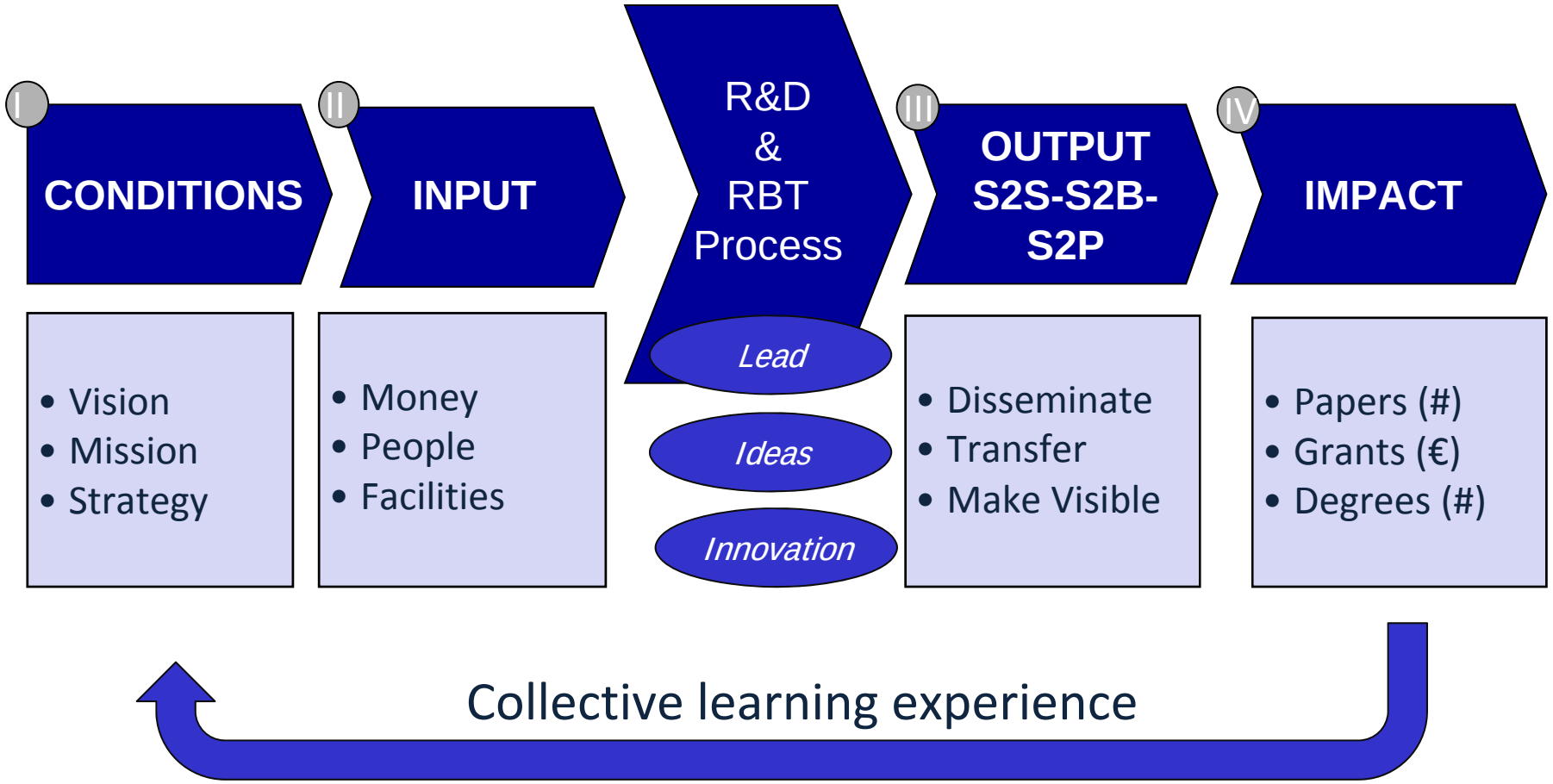
How to contribute to the international scientific community?

Science is testing crazy ideas – Engineering is putting these ideas into Business



Holzinger, A. 2011. Successful Management of Research and Development, Norderstedt: BoD.

Workflows for a Research Group



Holzinger, A. 2011. *Successful Management of Research and Development,* Norderstedt, BoD.

Papers are communications to the scientific community

In Journals

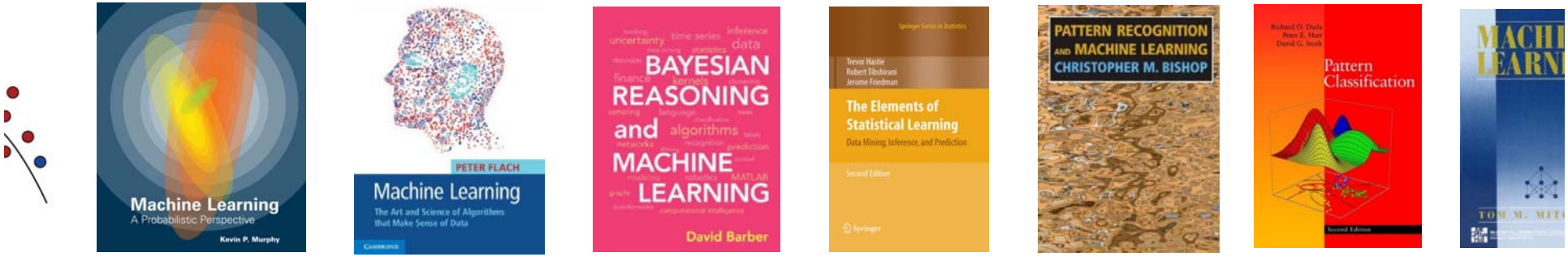
In Proceedings

In Magazines

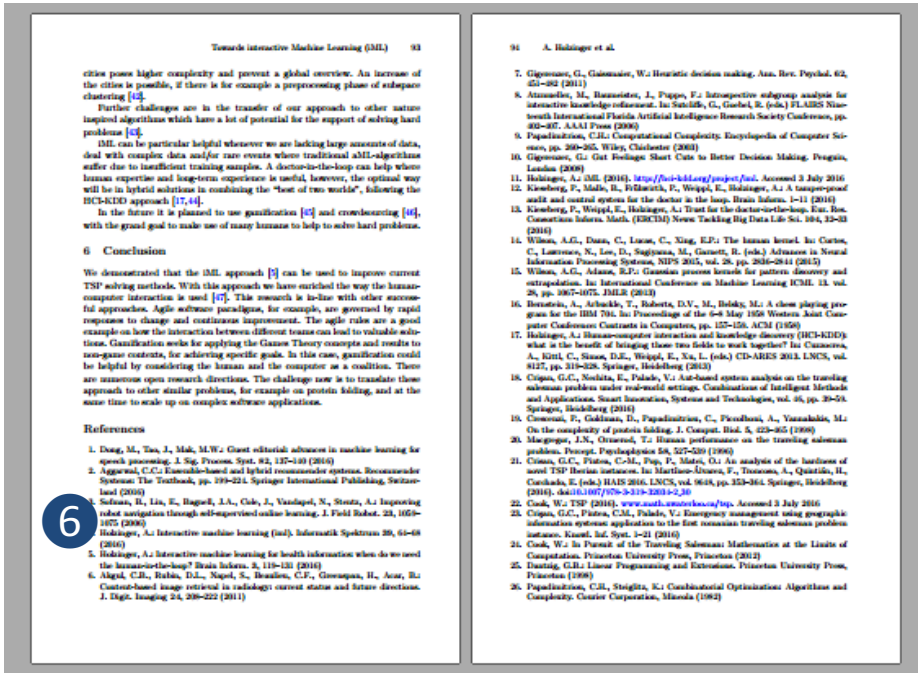
Book Chapters

Note: Should be included in the DBLP: http://dblp.uni-trier.de/db/

Not to be confused with Student Textbooks or Monographs:



Scientific communication via Papers

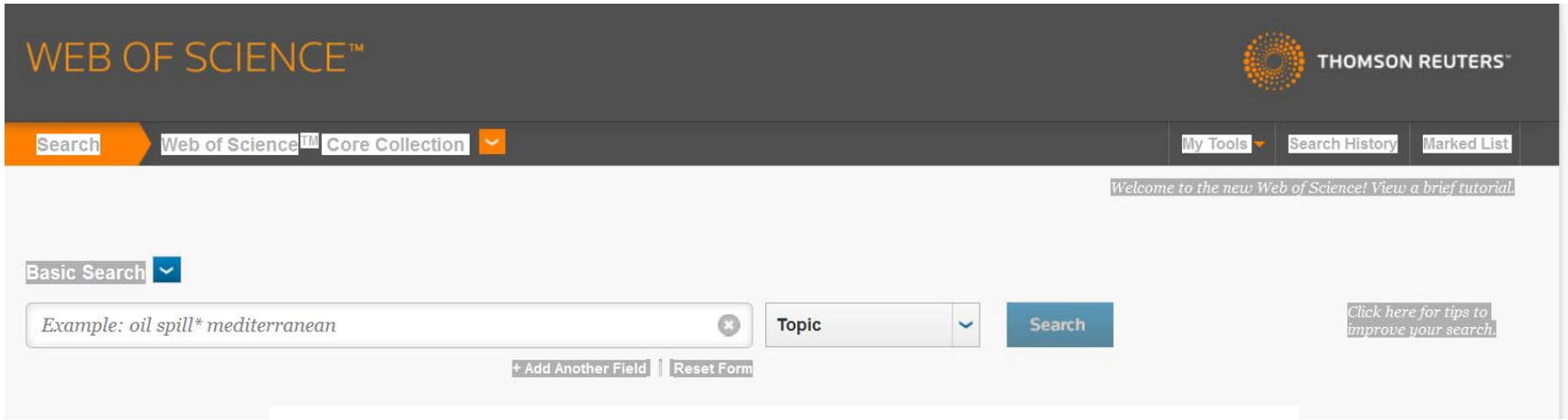


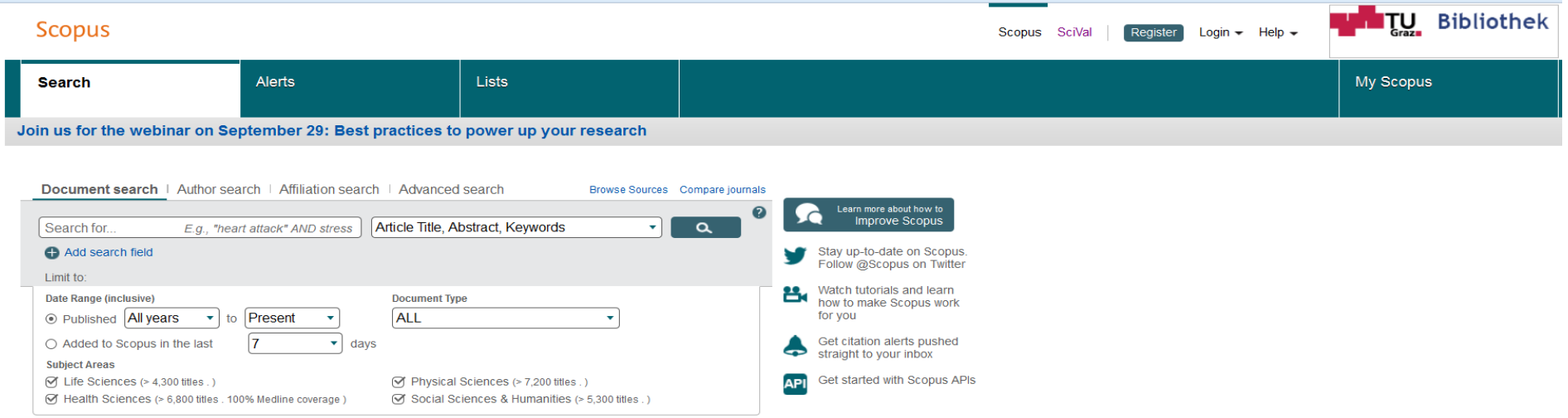
- 1 = Paper title
- 2 = Authors with Affiliations
- 3 = Abstract
- 4 = Keywords
- 5 = Content (formally divided into: 1) Introduction and Motivation for Research > 2) Background and Related Work, 3) Experimental method, setting, results, 4) Discussion, 5) Future Work, 6) Conclusion)
- 6 = References

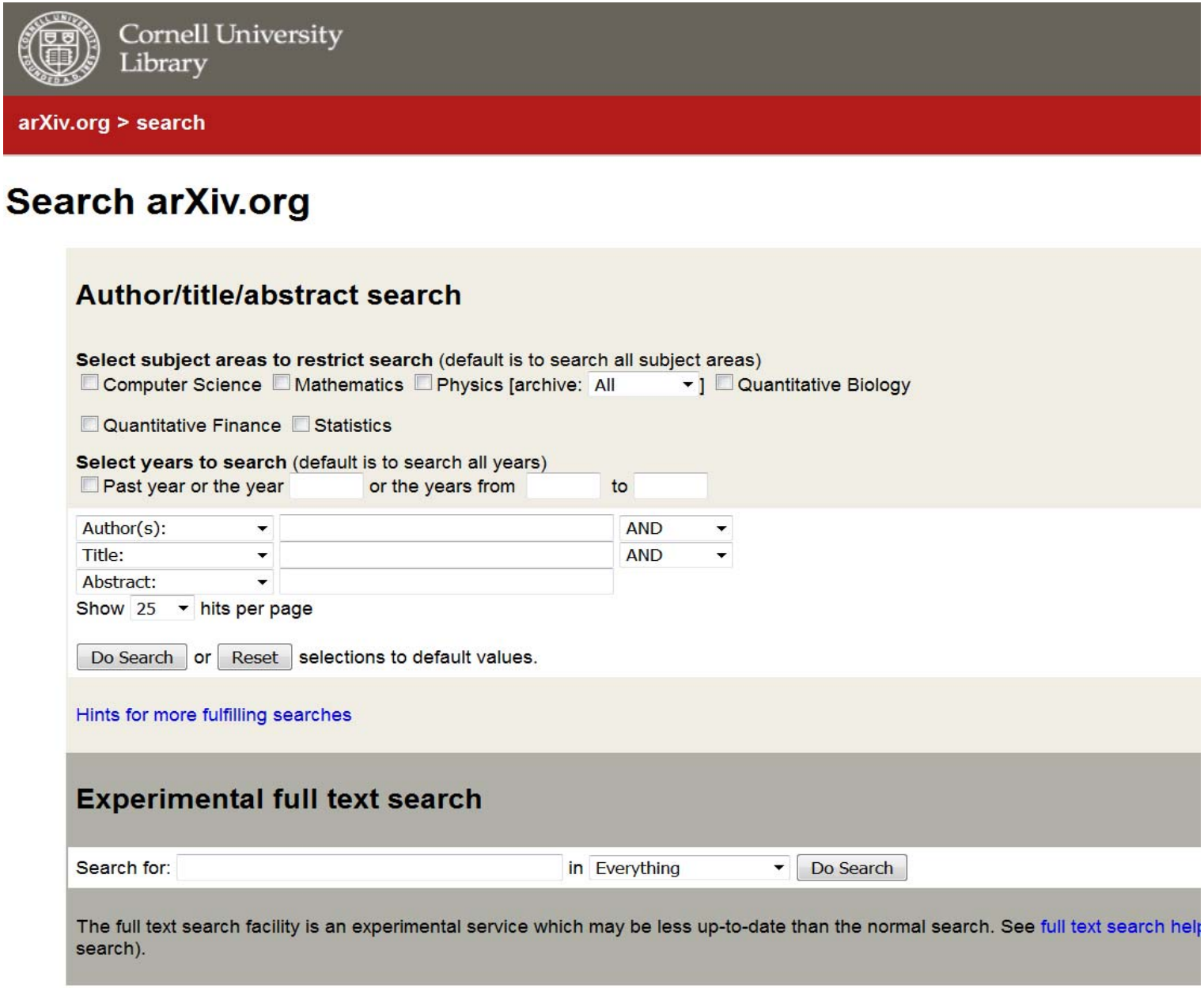
A scientific paper …

- is a **message** to the international research community
- is written in the scientific language today - **English** (Latin in medieval times and Greek in ancient times)
- reports something **of value** for other researchers
- should be useful, and what other researchers can use, will be **referenced** (brings citations)
- is subject to **peer review**
- has **specific form** (style and format)
- appears within a scientific **journal** or in **conference** proceedings (e.g. Springer LNCS)

Where do I find papers?






Cornell University Library
arXiv.org > search
Search arXiv.org
Author/title/abstract search
Select subject areas to restrict search (default is to search all subject areas)
Computer Science Mathematics Physics [archive: All] Quantitative Biology
Quantitative Finance Statistics
Select years to search (default is to search all years)
Past year or the year or the years from to
Author(s): AND
Title: AND
Abstract:
Show 25 hits per page
Do Search or Reset selections to default values.
Hints for more fulfilling searches
Experimental full text search
Search for: in Everything Do Search
The full text search facility is an experimental service which may be less up-to-date than the normal search. See full text search help
search).

How to read efficiently = how to write efficiently

- Stop Reading! Start Writing!
- Read along when writing:
- SQ3R-Method:
  - Survey (read title, abstract, conclusion, subheadings)
  - Question (what are the major insights of this paper?)
  - Read: with regard to the question above
  - Recite: summarize with your own words
  - Review: Try to reflect the major insights of the paper
- Do not waste time! Be economic! It is simply impossible to read everything and all!

- **Classic Newtonian** approach:
  - Ask question > develop theory > form a hypothesis to proof/disproof theory > conduct experiments > compare data with hypothesis > accept/reject theory
- **Computer Science** approach:
  - Find open problems to solve > form hypothesis how to solve the problem > experiment > evaluate > present new solution to the problem
- **Machine Learning** approach:
  - Setting up experiments to answer questions including:
    How does model m perform on data from domain D?
    Which of these models have the best performance?
    Much is feature engineering and precision and recall are your best friends!

Work evolution

- 1. Set goal (e.g. to bring paper into conference x or journal y) – write a preliminary (!) title and abstract
- 2. Study published work related to your topic
- 3. A good start is on the “future outlook” sections of published papers – outline intended work on one single page (birds eye view)
- 4. Start Writing! Discuss the related work and the theoretical background – leave gaps
- 5. Now bring in your ideas, experiments and results
- 6. Write Introduction, Conclusion, revise abstract, revise the title accordingly
- 7. Submit your paper
- 8. Carefully read the reviews, revise accordingly

Always start with an abstract

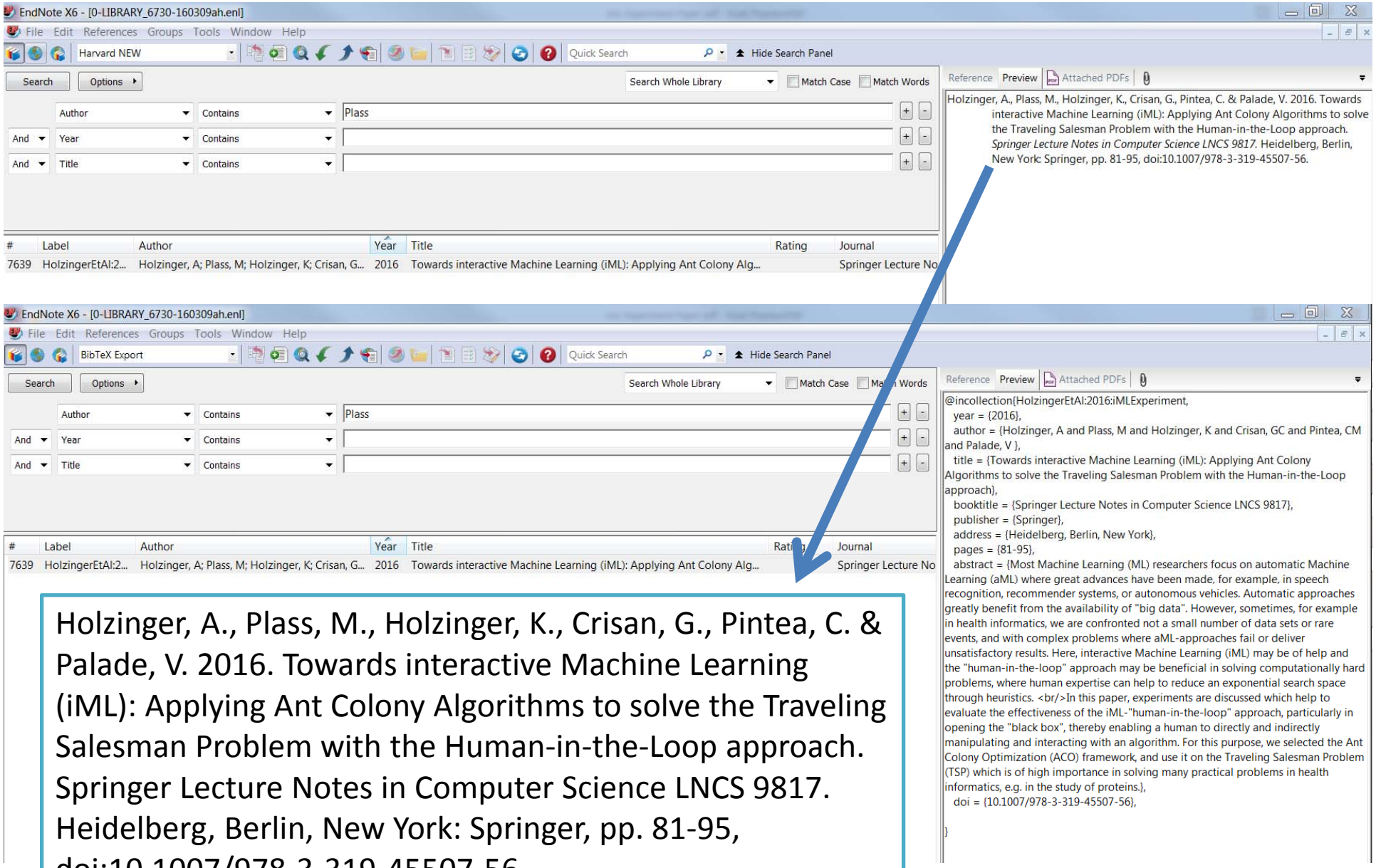
- What is the problem? Is it challenging?
- How can the problem be solved – alternative methods, background, related work?
- How well is the problem solved (evaluation)?
- How useful is the result to the intended readers?
- Example:
  - We propose a method … this is important because … we solve this problem via … finally we demonstrate that our method outperforms the state-of-the-art …

THOMSON REUTERS
ENDNOTE

LaTeX

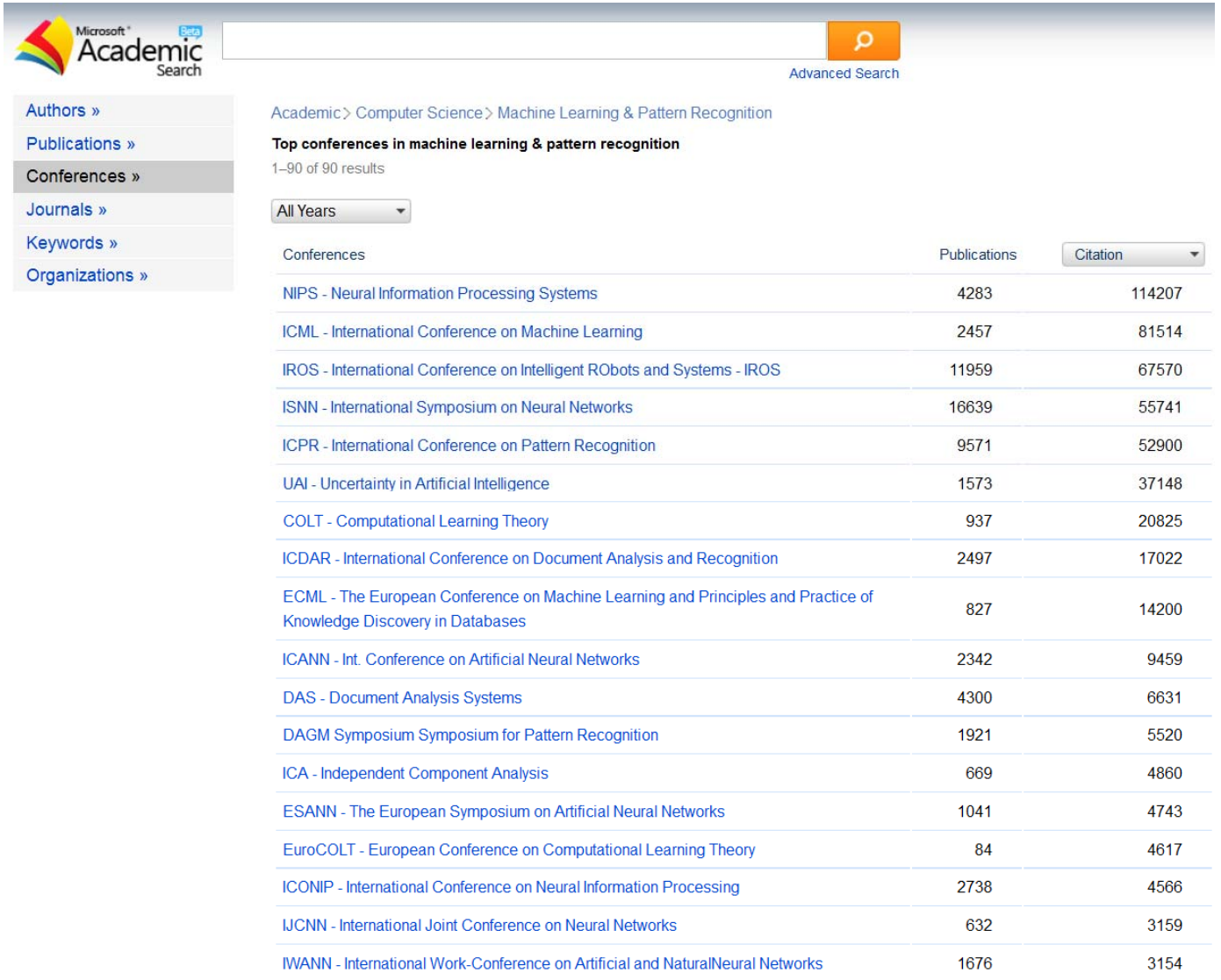
BibTeX

EndNote Library (format incl. Bibtex)



Holzinger, A., Plass, M., Holzinger, K., Crisan, G., Pintea, C. & Palade, V. 2016. Towards interactive Machine Learning (iML): Applying Ant Colony Algorithms to solve the Traveling Salesman Problem with the Human-in-the-Loop approach. Springer Lecture Notes in Computer Science LNCS 9817. Heidelberg, Berlin, New York: Springer, pp. 81-95, doi:10.1007/978-3-319-45507-56.

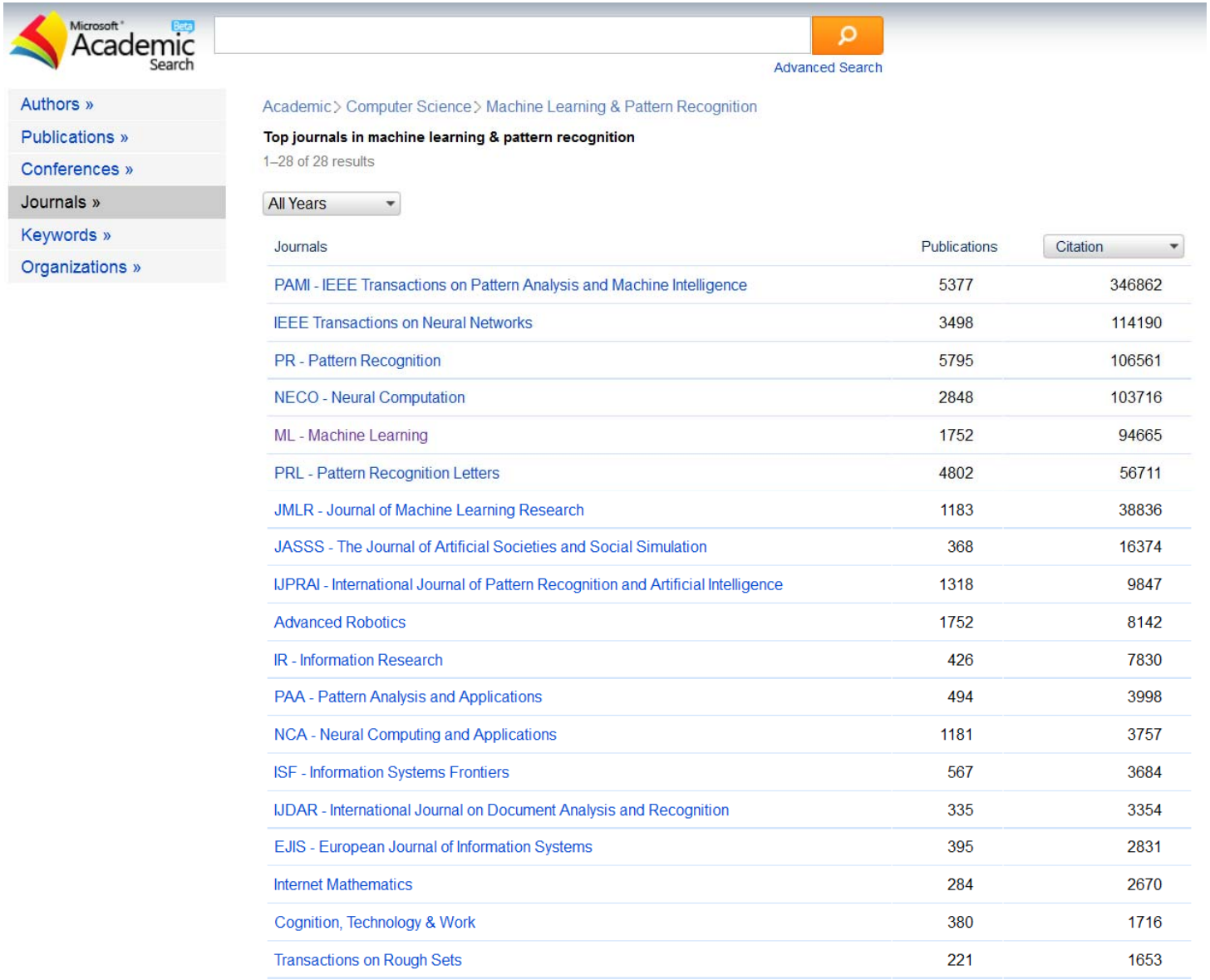
Publishing: Conferences



Conferences	Publications	Citation
NIPS - Neural Information Processing Systems	4283	114207
ICML - International Conference on Machine Learning	2457	81514
IROS - International Conference on Intelligent RObots and Systems - IROS	11959	67570
ISNN - International Symposium on Neural Networks	16639	55741
ICPR - International Conference on Pattern Recognition	9571	52900
UAI - Uncertainty in Artificial Intelligence	1573	37148
COLT - Computational Learning Theory	937	20825
ICDAR - International Conference on Document Analysis and Recognition	2497	17022
ECML - The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases	827	14200
ICANN - Int. Conference on Artificial Neural Networks	2342	9459
DAS - Document Analysis Systems	4300	6631
DAGM Symposium Symposium for Pattern Recognition	1921	5520
ICA - Independent Component Analysis	669	4860
ESANN - The European Symposium on Artificial Neural Networks	1041	4743
EuroCOLT - European Conference on Computational Learning Theory	84	4617
ICONIP - International Conference on Neural Information Processing	2738	4566
IJCNN - International Joint Conference on Neural Networks	632	3159
IWANN - International Work-Conference on Artificial and NaturalNeural Networks	1676	3154

http://academic.research.microsoft.com/RankList?entitytype=3&topDomainID=2&subDomainID=6

Publishing: Journals



Journals	Publications	Citation
PAMI - IEEE Transactions on Pattern Analysis and Machine Intelligence	5377	346862
IEEE Transactions on Neural Networks	3498	114190
PR - Pattern Recognition	5795	106561
NECO - Neural Computation	2848	103716
ML - Machine Learning	1752	94665
PRL - Pattern Recognition Letters	4802	56711
JMLR - Journal of Machine Learning Research	1183	38836
JASSS - The Journal of Artificial Societies and Social Simulation	368	16374
IJPRAI - International Journal of Pattern Recognition and Artificial Intelligence	1318	9847
Advanced Robotics	1752	8142
IR - Information Research	426	7830
PAA - Pattern Analysis and Applications	494	3998
NCA - Neural Computing and Applications	1181	3757
ISF - Information Systems Frontiers	567	3684
IJDAR - International Journal on Document Analysis and Recognition	335	3354
EJIS - European Journal of Information Systems	395	2831
Internet Mathematics	284	2670
Cognition, Technology & Work	380	1716
Transactions on Rough Sets	221	1653

http://academic.research.microsoft.com/RankList?entitytype=4&topdomainid=2&subdomainid=6&last=0&orderby=1

Peer Review

http://hci-kdd.org/wordpress/wp-content/uploads/2014/10/REVIEW_TEMPLATE_XXXX.docx

http://hci-kdd.org/wordpress/wp-content/uploads/2014/10/REVIEW_TEMPLATE_XXXX.pdf

HCI-KDD

***begin of Double-Blind Review

1) Title of the Paper:

2) Please describe briefly with your own words what this paper is about:

This paper reports on x ...

3) Originality: Does the paper contain significant content to justify publication? Which are the novel aspects? Is the paper original, i.e. did you check for plagiarism (at least with a quick Google search)?

Novel aspects include the topic x ...

4) Literature: Does the authors demonstrate an adequate understanding of the relevant related work? Are any relevant references missing? Please provide recommendations.

The following important related papers are missing ...

5) Methodology: Is the paper's argument built on an appropriate base of theory and concepts? Are the methods used appropriately described?

The methods x are ...

6) Results: Are the results presented clearly and appropriately? Do the conclusions adequately tie together the other elements of the paper?

The results are ...

7) Quality of Communication: Is the paper well written? Has attention been paid to the clarity of expression and readability, sentence structure, acronym explanation, etc.

The paper is ...

8) Your recommendation (mark or underline):
A) Accept – B) Minor Revision – C) Major Revision – D) Reject

9) On a scale between 0 (rubbish) **to 100** (excellent) **– how would you rate this paper:**

100	90	80	70	60	50	40	30	20	10	0

If you recommended: a) b) or c) – please outline how the authors can improve their paper: What should the authors do? What should they expand/remove etc.? What should they improve?

(use additional space as you need it)

***end of double-blind review

Review [illegible] 5/2015 Thank you very much for your time and effort – your help is most appreciated

Thank you!

Some literature recommendations