



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





■ 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"



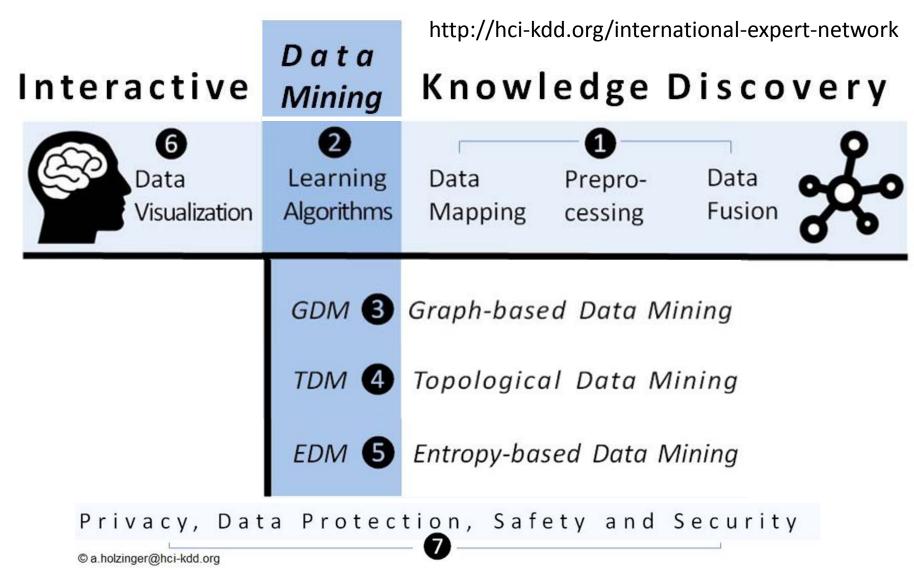




Science is to test crazy ideas – Engineering is put these ideas into Business!



- 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



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



- 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!



rriday 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



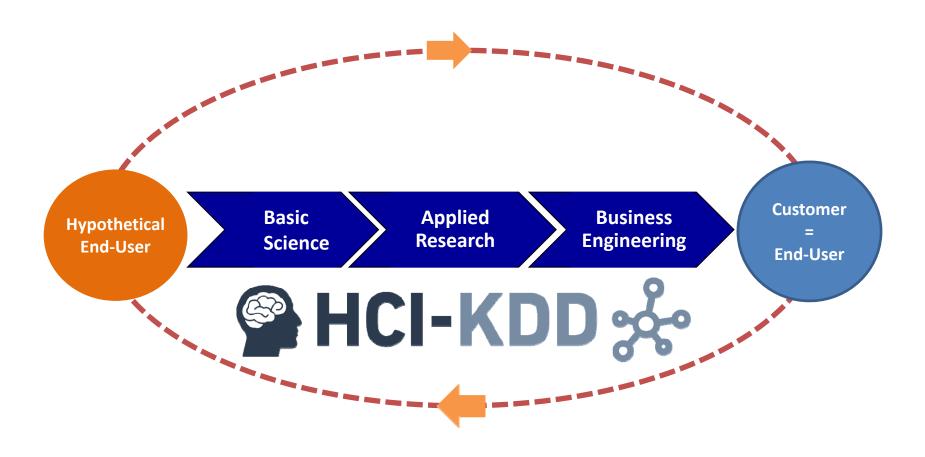


How to contribute to the international scientific community?





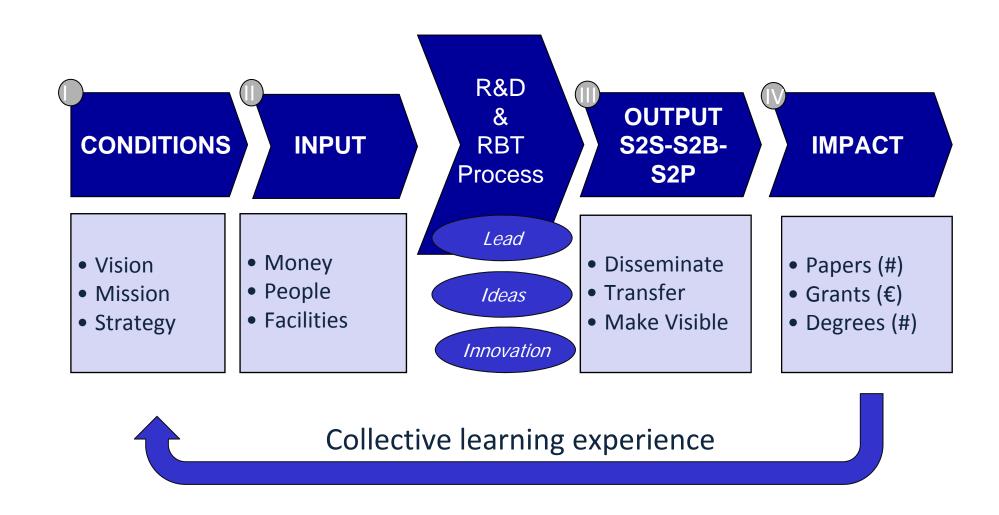
Science is testing crazy ideas – Engineering is putting these ideas into Business



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







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



Papers are communications to the scientific community

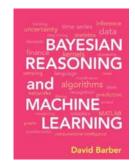




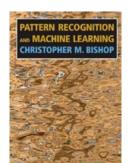
Not to be confused with Student Textbooks or Monographs:

















Scientific communication via Papers





Andreas Holzinger ¹⁹⁶), Markus Plass¹, Katharina Holzinger¹, Gioria Cerasela Crigan², Camella-M. Pintea³, and Vasile Palade⁴

Holzinger Group HCI-KDD, Institute for Medical Information Statistics & Documentation, Medical University Graz, Graz, Austria (a. holzinger, a. plans, k. holzinger) thei-told. org Vasile Alexandri University of Backs, Backs, Romania

ceraselacrizantub.re Technical University of Chij-Napora, Chij-Napora, Romania dr.camelia.pintentinee.org
Faculty of Engineering, Environment and Computing, Coventry University
Coventry, UK

Abstract. Most Machine Learning (ML) researchers focus on automatic Machine Learning (ML) where great advances have been made, for example, in query-broughther, commencing sprinting as autonomous whether. Automatic supramedure greatly benefit from the antichality of the contract of the

olving many practical problems in health informatios, e.g. in the study

Keywords: interactive Machine Learning - Human-in-the-loop Traveling Salesman Problem - Ant Colony Optimization

1 Introduction and Motivation for Research

Automatic Machine Learning (aML) is increasingly making big theoretical a well as practical advances in many application domains, for example, in speech recognition [1], recommender systems [3], or autonomous vehicles [3]. The aML-approaches sometimes fail or deliver unsatisfactory results, when

The abili-opproaches sometimes hall or deliver insatisfactory counts, when the programme of the property of th

classic supervised learning approach, the first question is to define the human's role in this loop (see Fig. 1), [3].



Fig. 1. The IML human-in-the-loop approach. The main issue is that lumnars are not only anolyod in preparessing, by selecting data or features, but actually during the locating phase they are discretely interacting with the algorithm, thus shifting away the black-box approach to a glaze-box; there might also be more than one lumnan agent interacting with the expertation approach to a glaze-box; there might also be more than one lumnan agent interacting with the computational approach(s), always for convoluncing or gamilloxium and the computational approach(s), always for convoluncing or gamilloxium and the computational approach(s), always for convoluncing or gamilloxium and the computational approach(s), always the convoluncing or gamilloxium and the computational approach are considered as the convolunce of the convoluncing or gamilloxium and the computational approach are considered as the convolunce of the convo

There is evidence that humans sometimes still outperform MI-algorithms, e.g. in the latticities, often almost interataneous interpretation of complex patterns and the latticities, and the latticities, and the second properties of the latticities and the lattice and the lattice and lattice and

no evidence of how such interaction may really optimize these algorithms as it is a subject that is still being studied by cognitive scientists for quite a while and even though "natural" intelligent agents are present in large numbers throughout

even though "natural" intelligant agents are present in suge incasers torongoment the world [1]. From the theory of human problem solving it is known that for example, medical decines can often make diagnoses with great reliability - but without being able to explain their rules explicitly. Here MM could help to equip algo-rithms with early "admitstrative" knowledge and learn thereof. The importance of rules are considered to the contractive of the contractive of

Towards interactive Machine Learning (iML) 93

cities poses higher complexity and prevent a global overview. An increase of the cities is possible, if there is for example a preprocessing phase of subspace clustering [12]. Further challenges are in the transfer of our approach to other nature

iML can be particular helpful whenever we are lacking large amounts of data, and it is a first particular adopts tensioner to the next page angular and in-side of the statistical and the statistical and

with the grand goal to make use of many humans to help to solve hard proble

We demonstrated that the iML approach [5] can be used to improve current TSP solving methods. With this approach we have enriched the way the human-computer interaction is used [47]. This research is in-line with other successtompore interaction is since [11]. He seekees in Jane who can be accessed in approaches, Agile software paradigms, for consupie, are governed by rapid responses to change and continuous improvement. The agile rules are a good example on how the interaction between different teams can local to valuable solutions. Gazulfication seeks for applying the Games Theory concepts and results to tools. Commutation seem for appring the Canise Theory concepts and results to non-game contexts, for arbitrary apertife goods. In this case, gaunifestion could be helpful by considering the human and the computer as a condition. There are numerous open research directions. The challenge now is to translate these approach to other similar problems, for example on protein folding, and at the same time to scale up on complete software applications.

- speech processing. J. Sig. Process. Syst. 82, 137-140 (2016)
 Aggarwal, C.C.: Execuble-based and hybrid recommender systems.
 Systems: The Textbook, pp. 199-221. Springer International Publi

- Intellinger, A. Internative incension stationage for install instructions were to be seen the learner-inche-loop? Brain Inform. 3, 119–131 (2016)
 Alegal, C.B., Rutha, D.L., Napel, S., Bourlies, C.F., Greenquan, H., Acar, R.: Content-based image retrieval in radiology: current status and future directions.
 J. Bigh. Imaging 24, 208–222 (2011)

- Gigeneauer, G., Gaismaier, W.; Henristic decision making. Ann. Rev. Psychol. 62, 63:1–82 (2011)
 Attmeeller, M.; Banneieter, J., Pappe, F.; Introspective subgroup analysis for interactive knowledge ordeneaut. In Scholler, G., Gesbel, E. (eds.) P. ALISS Nincolneck Interactional Facials Artificial Recolleges reference facilities of the School Schoo

- dillusion. Adult resea passing and Completelly, Eurydopedia of Computer Science, pp. 204–205. Wiley, Chalester (2001)

 16. Gigenesse, G. Cut Fersings Short Clerk to Herter Decision Making, Posspin, Landan (2008).

 16. Reingesser, G. Cut Fersings Short Clerk to Herter Decision Making, Posspin, Landan (2008). https://doi.org/10.1011/journal.poss.com/10.1011/journal.poss.
- (2016) 14. Wilson, A.G., Dann, C., Lucas, C., Xing, E.P.: The human hernel. In: Cortes, C., Lauvence, N., Lee, D., Sagiyama, M., Garnett, R. (eds.) Advances in Neural Information Processing Systems, NIPS 2015, vol. 28, pp. 2816–2841 (2015) 15. Wilson, A.G., Adams, R.P.: Gaussian process bears for pattern discovery and
- extrapolation. In: International Conference on Machine Learning ICML 13, vol. 28, pp. 1067–1075. JMLR (2013)

- extrapolation. In International Conference on Manthes Learning (CMI. 13 vol. 18 processes of the Conference of the Confe

- Cook, W.: TSP (2016). www.math.unnordon.or/tsp. Accessed 3 July 2016
 Cripan, G.C., Pinton, C.M., Palade, V.: Emergency management using geographic information systems application to the first romanian traveling salesman problem.

- Princeton (1998)

 26. Papadimitrion, C.H., Steiglitz, K.: Combinatorial Optimiz Complexity. Courier Corporation, Miscola (1982)

- 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)

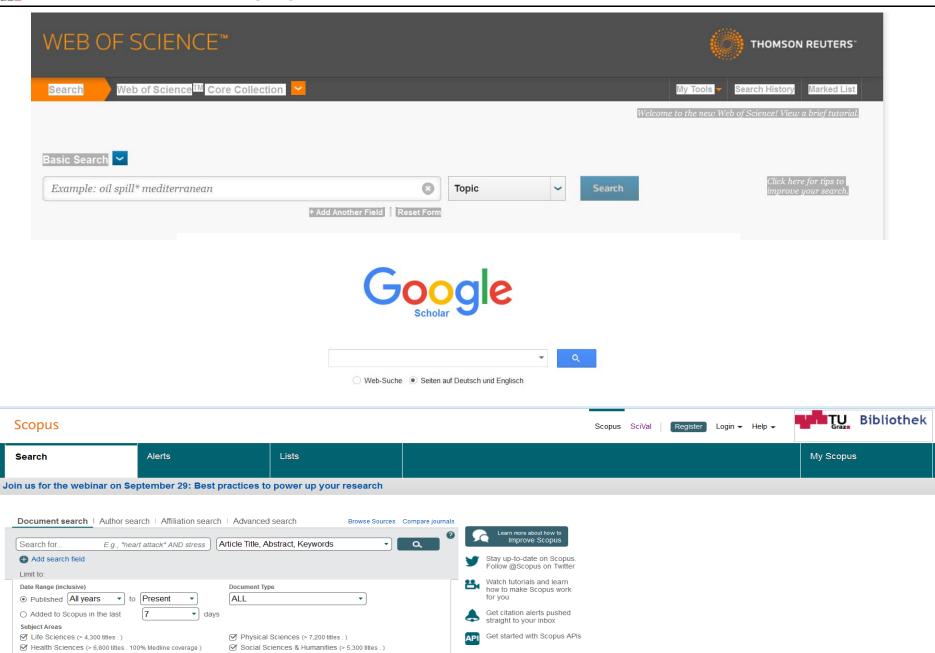
6 = References



- 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?



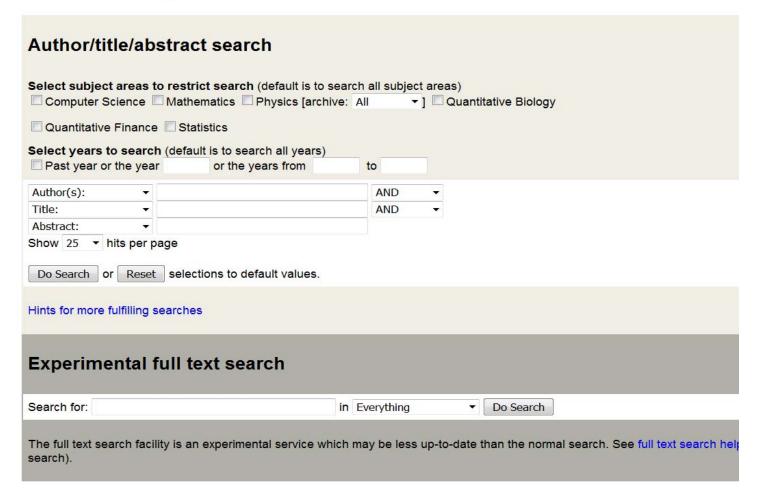








Search arXiv.org







- 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!



- 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



- 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 ...





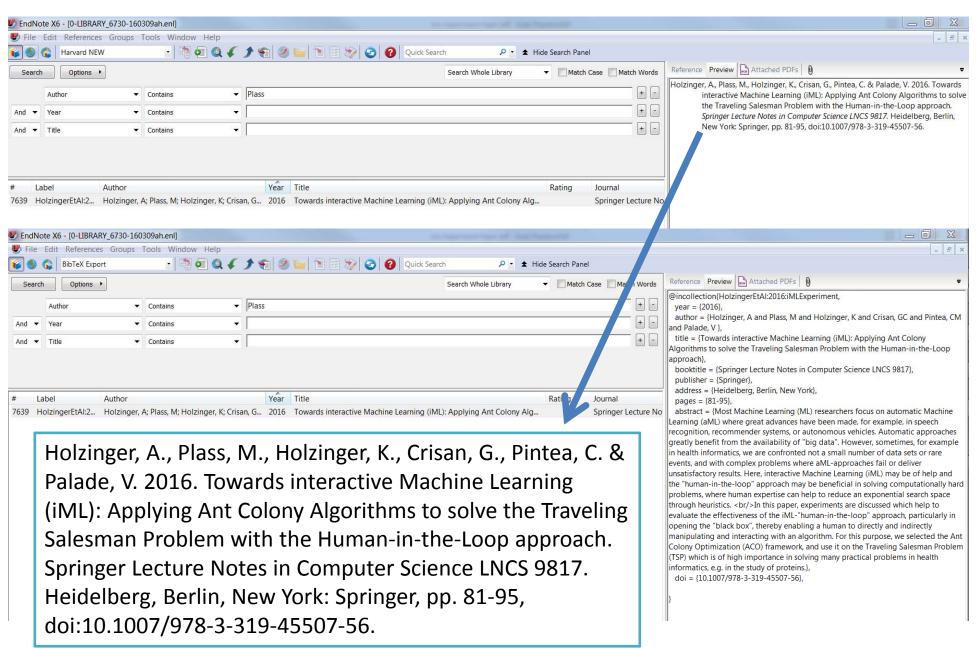




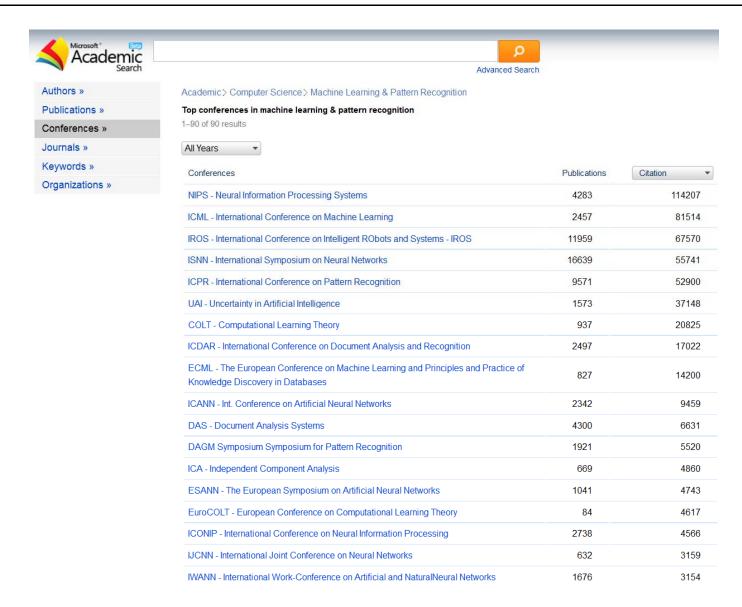


EndNote Library (format incl. Bibtex)



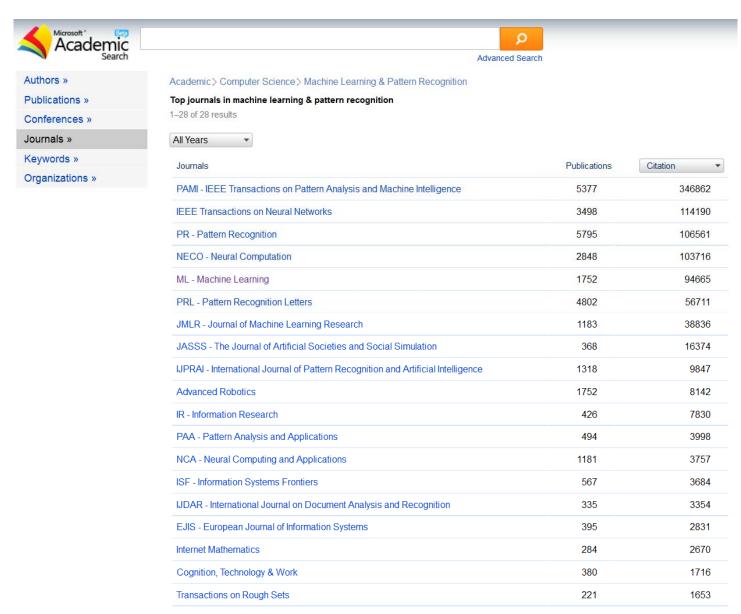






http://academic.research.microsoft.com/RankList?entitytype=3&topDomainID=2&subDomainID=6





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



HCI-KDD %

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

```
***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·O·(rubbish)·to·100·(excellent)·--how·would·you·rate·this·paper:·¶
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 Farmulay 5/2015 Thank you very much for your time and effort - your halp is most appraciated?



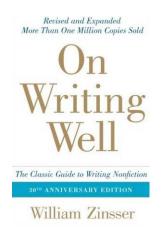


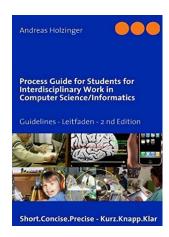
Thank you!

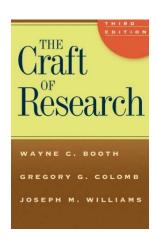
Some literature recommendations

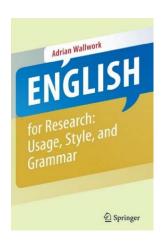


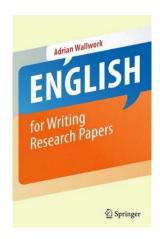


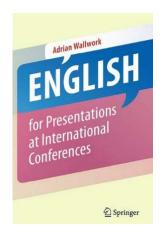


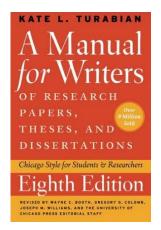












26