

Data Science Workshop

Lecture 13: State of the Art

Marcin Luckner, PhD
mluckner@mini.pw.edu.pl

Version 1.1
November 19, 2020



**European
Funds**
Knowledge Education Development

**Warsaw University
of Technology**

European Union
European Social Fund



**MSc program in Data Science has been developed
as a part of task 10 of the project
„NERW PW. Science - Education - Development - Cooperation”
co-funded by European Union from European Social Fund.**

State of the Art



A Time to Harvest by Chaosium Inc.

State of the Art review

- Working on some issue, we should make a state of the art review.
- The review allows us to answer the following questions:
 - Is it worth to take up this issue?
 - What has been done in this area?
 - What left to achieve?

State of the Art objectives

- With a state of the art we can:
 - Determine if our solution is innovative in comparison to the existing solutions.
 - Determine a benchmark.
 - Determine how to verify and assess our solutions.

Sources



The screenshot displays a list of four references in a table-like format. Each entry includes a document icon, a title, author information, publication details, and a status indicator. The first three entries are about 'A flat earth society? Imagining academic freedom' by Boden, Rebecca; Epstein, Debbie, published in 'The Sociological Review' in August 2011. The fourth entry is 'Geology: The end of flat Earth' by Nature, 2008. Each entry has a green dot icon and the text 'Dostępny pełny tekst' (Full text available). Below each entry are links for 'Pełny tekst' (Full text) and 'Szczegóły' (Details). The word 'Artykuł' (Article) is also visible next to each entry.

Artykuł	☆ A flat earth society? Imagining academic freedom.(Report) Boden, Rebecca ; Epstein, Debbie The Sociological Review, August, 2011, Vol.59, p.476(20) [Czasopismo recenzowane] ● Dostępny pełny tekst Pełny tekst Szczegóły
Artykuł	☆ A Flat Earth Society? Imagining Academic Freedom Boden, Rebecca ; Epstein, Debbie The Sociological Review, August 2011, Vol.59(3), pp.476-495 [Czasopismo recenzowane] ● Dostępny pełny tekst Pełny tekst Szczegóły
Artykuł	☆ The Flat Earth Society: a rose by any other name? Nelson, Lawrence M Human Reproduction, 2014, Vol. 29(2), pp.190-192 [Czasopismo recenzowane] ● Dostępny pełny tekst Pełny tekst Szczegóły
Artykuł	☆ Geology: The end of flat Earth Nature, 2008, Vol.454(7205), p.671 [Czasopismo recenzowane] ● Dostępny pełny tekst Pełny tekst Szczegóły

- It is suggested to analyse scientific sources.
- Scientific works go through a revision process.
- Published works are a subject of a public academic discussion.
- That increases trust in published information.
- Works in Computer Science area are mostly published in an electronic version.

First steps of review

1. Determine the keywords that identify works in the examined area.
2. Search databases using the keywords.
3. Find **several dozen** of promising publications.
4. Carry out a primary selection of useful works.

Keywords

- Determine keywords that can be used to describe the examined area.
 - Find synonyms.
- When you find an article that covers the examined area, find the keywords inside the article and use them.

Publishers' databases

- Full-text databases.
- Usually limited to one publisher.
- The most extensive databases in the computer science area are Springer and IEEE.
 - Mostly, conference papers.
- ACM Digital Library
- IEEE/IEE Electronic Library
- Science Direct
- Springer Link
- Taylor and Francis Online
- Wiley Online Library

Sources

Intelligent Information and Database Systems
3,753 Result(s) for 'wi-fi positioning system' within Computer Science

Refine Your Search

Content Type

Chapter	3,815
Conference Paper	2,824
Article	895
Reference Work Entry	43

Discipline

Computer Science

Subdiscipline

Computer Communication Networks	1,552
Information Systems Applications (incl. Internet)	1,282
Artificial Intelligence	1,107
User Interfaces and Human Computer Interaction	893
Information Storage and Retrieval	548

Language

English	3,743
German	10

Sort By: Relevance | Newest First | Oldest First | Date Published | Page 1 of 188

Chapter and Conference Paper

Comparison of Indoor Positioning System Using Wi-Fi and UWB

Recently, as smart mobile devices become popular, location-based services are wide spreading. In the indoor positioning and guiding system, the accuracy and efficiency of the system are important. Various indoor

Jae-min Hong, Hyun Kim, Chong-Gun Kim in *Intelligent Information and Database Systems* (2018)

Chapter and Conference Paper

Open Source OwlPS 1.3: Towards a Reactive Wi-Fi Positioning System Sensitive to Dynamic Changes

Since 2004, our team has been developing an academic positioning system (PS for short) for hostile environments. The techniques involved are mainly Wi-Fi based. While at the very beginning the system architecture...

Philippe Camaleide, Mathieu Cyprien... in *Evaluating AAL Systems Through Competitive...* (2012)

Chapter and Conference Paper

Using Cameras to Improve Wi-Fi Based Indoor Positioning

Indoor positioning systems are increasingly being deployed to enable indoor navigation and other indoor location-based services. Systems based on Wi-Fi and video cameras rely on different technologies... show th...

Laura Padellai, Yael Moses... in *Web and Wireless Geographical Information...* (2016)

Showing 1-25 of 1,147 for **wi-fi positioning system**

Conferences (971) Journals (136) Magazines (24) Early Access Articles (0)
Books (3) Courses (3)

Select All on Page Sort By: Relevance

Show

All Results Open Access Only

Year

Single Year Range

2003 2020

From To

2000 2028

Abstract Full Text PDF HTML

A hybrid indoor **positioning system** based on **Wi-Fi** hotspot and **Wi-Fi** based nodes
Sudheer R. Delpinde, J. W. Baidal, Madhavi Ganes
2016 IEEE International Conference on Engineering and Technology (ICETECH)
Year: 2016 | Conference Paper | Publisher: IEEE
Cited by: Papers (1)

Abstract Full Text PDF HTML

Recurrent neural networks model for Wi-Fi-based indoor **positioning system**
Yuan Li, Li, Antonio Pascual-Chamorro
2017 International Conference on Smart Cities, Automation & Intelligent Computing Systems (ICON-SOINCS)
Year: 2017 | Conference Paper | Publisher: IEEE
Cited by: Papers (4)

Abstract Full Text PDF HTML

<https://ieeexplore.ieee.org/>

<https://link.springer.com/>

General databases

- Abstract databases.
- Collect entries from various sources.
- A good source of bibliographical data.
 - Respected databases in the computer science area are Web of Science and Scopus.
- The WUT Base of Knowledge
- DBPL Computer Science
- Google Scholar
- Scopus
- Web of Science

Results diversity

- An example of preliminary research for a given author.
 - Author: Marcin Luckner
 - Date: 31.01.2019
- The number of records
 - Google Scholar - 47
 - The WUT Base of Knowledge - 40
 - Scopus - 36
 - DBPL Computer Science - 35
 - Web of Science - 24

Database search guidelines

- Use more than one database.
- Use open sources and sources offered by the Main Library of WUT.
- Using full-text databases saves data access time.
- However, the abstract databases give us a more comprehensive view and sometimes link to the full texts.

Preliminary selection

- In case of a substantial number of articles covering the examined area, we can use the following criteria to narrow the set
 - A publication year
 - The newest publication should present the current State of the Art.
 - A citation number
 - Frequently cited articles may be important for the research area.
 - A publication place
 - Publications in respected journals (and in some conference materials) go through a multi-step review process.

Journal assessment

- The impact factor (IF) is an index that reflects the yearly average number of citations that articles published in the last two years in a given journal received.
- It is frequently used as a proxy for the relative importance of a journal within its field.
- Journals with higher impact factors are often deemed to be more important than those with lower ones.
- Journals with IF are grouped in Journal Citation Reports by Clarivate.

IF controversy

- IF is used to assess the authors.
- Some editors manipulate IP factor value
 - Self-citations and cross-citations.
 - The Polish journal *Przegląd Elektrotechniczny* was removed from JCR in 2013 with the self-citation index close to 77%.
- Predatory publishers use fake impact factors
 - Universal Impact Factor (UIF), Global Impact Factor (GIF), and Citefactor.

Conference assessment

- The CORE Conference Ranking provides assessments of significant conferences in the computing disciplines.
- Conferences are assigned to one of the following categories:
 - A* flagship conference, a leading venue in a discipline area
 - A excellent conference, and highly respected in a discipline area
 - B good conference, and well regarded in a discipline area
 - C other ranked conference venues that meet minimum standards

<https://www.core.edu.au/conference-portal>

Reference management

Reference Management Tools

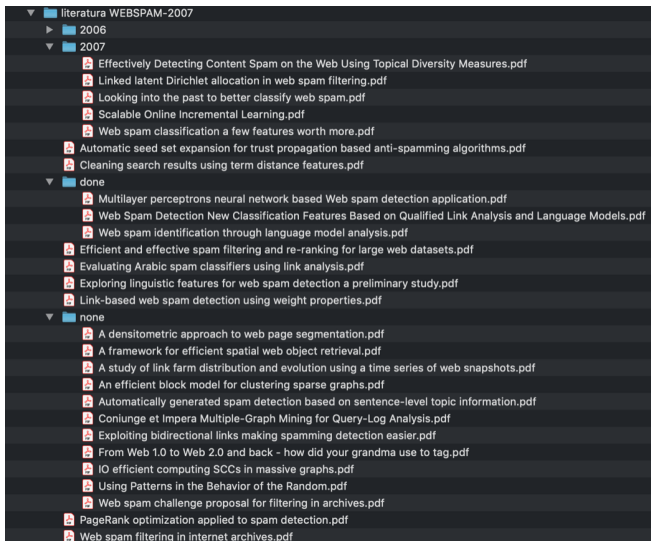


Using Reference Management Tools: EndNote and Zotero by Diarmuid Stokes

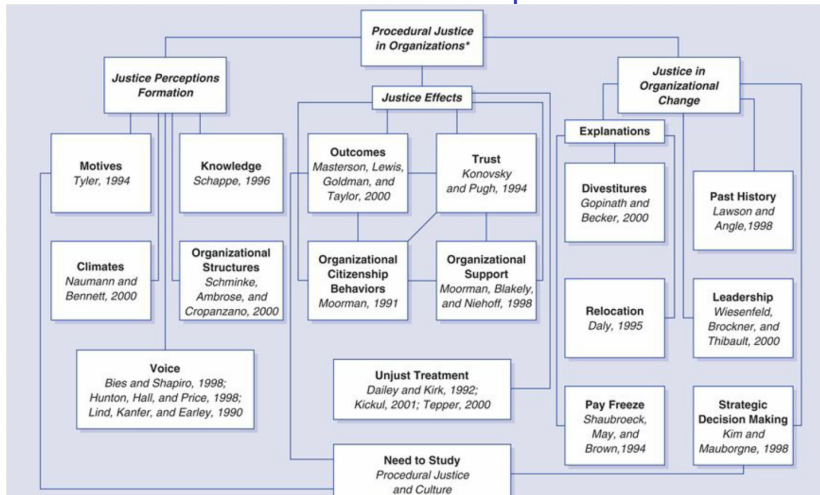
Reference management

- Working with several dozens of references, we have to keep them in order.
- An absolute minimum is to order bibliographical data and full texts in logical structure.
- Conceptually, we can be supported by hierarchical folder structure and literature map.
- A helpful tool is a reference manager.

Organised folder structure



Literature map



Research Design: Qualitative, Quantitative and Mixed Methods by John W. Creswell and J. David Creswell.

Organised bibliographical data

- Bibliographical data can be kept using bibtex format.
- That allows the author to create references quickly.
- A reference manager allows us to use this format comfortably.

```
@InProceedings{Bruckner2009,  
  Title           = {Nash Equilibria of Static Prediction  
    Games},  
  Author          = {Michael Br{"u"}ckner, and Tobias Scheffer},  
  Booktitle       = {NIPS},  
  Year            = {2009},  
  Pages           = {171-179},  
  Bibsource       = {DBLP, http://dblp.uni-trier.de},  
  Crossref        = {DBLP:conf/nips/2009},  
  EE              = {http://books.nips.cc/papers/files/nips22/  
    NIPS2009\_0534.pdf}  
}  
  
@inproceedings{Hao2016,  
  author = {Hao, Shuang and Kantchelian, Alex and Miller, Brad and  
    Paxson, Vern and Framster, Nick},  
  title = {PREDATOR: Proactive Recognition and Elimination of Domain  
    Abuse at Time-Of-Registration},  
  booktitle = {Proceedings of the 2016 ACM SIGSAC Conference on Computer  
    and Communications Security},  
  series = {CCS '16},  
  year = {2016},  
  isbn = {978-1-4503-4139-4},  
  location = {Vienna, Austria},  
  pages = {1568--1579},  
  numpages = {12},  
  url = (http://doi.acm.org/10.1145/2976749.2978317),  
  doi = {10.1145/2976749.2978317},  
  acmid = {2978317},  
  publisher = {ACM},  
  address = {New York, NY, USA},  
  keywords = {domain registration, early detection, reputation system},  
}
```

Reference manager

The screenshot displays the Mendeley Desktop application interface. At the top, there is a navigation bar with tabs for Feed, Library (selected), Suggest, Groups, Datasets, Careers, and Funding. A search bar labeled 'Library search' and a user profile for 'Marcin' are also present. Below the navigation bar, a toolbar includes buttons for '+ Add', 'Add to', 'Delete', 'Export to MS Word', and a dropdown for 'Added (newest)'. The main area is divided into two panes. The left pane, titled 'MY LIBRARY', shows a list of documents with columns for checkboxes, star ratings, document titles, authors, and dates. The right pane, titled 'Details', shows the details for the selected document, 'Multi-oriented text recognition in graphical documents using HMM' by Roy P., Roy S., and Pal U. The details include the journal name, DOI, ISBN, and a download button. The bottom of the interface shows a status bar with the text 'Processing of PDF started, will finish and succeed'.

Document Title	Author	Date
Automatic knowledge acquisition: Recognizing music notation with methods of centroids and classifications trees	Hornik R, Luckner M	10/10/15
Long term analysis of the localization model based on Wi-Fi network	Ghah R, Luckner M	10/10/15
Hybrid algorithm for floor detection using GSM signals in indoor localisation task	Luckner M, Ghah R	10/10/15
Conversion of decision tree into deterministic finite automaton for high accuracy online SYN flood detection	Luckner M	10/10/15
Global and local rejection option in multi-classification task	Luckner M	10/10/15
Labeled VoIP Data-set for Intrusion Detection Evaluation	Nasser M, Radu S, Fester O	30/09/15
A New Ensemble Model based on Linear Mapping, Nonlinear Mapping, and Probability Theory for Classification Problems	Charbonneau A, Jayen S	03/09/15
A review of dynamic vehicle routing problems	Pilav V, Gendreau M, Guenet C, et al.	28/07/15
Text Recognition for Information Retrieval in Images of Printed Circuit Boards	Li W, Neufkens S, Steier M, et al.	21/07/15
Multi-oriented text recognition in graphical documents using HMM	Roy P, Roy S, Pal U	21/07/15

<https://www.mendeley.com/>

Reference managers comparison

Reference management software	Cost	Does it integrate with Microsoft Word?	Can it capture a webpage to create a record?
RefWorks	US\$100/year or free at any subscribing academic institution	Yes (requires download of Write-N-Cite utility)	Yes
Zotero	Free/open source	Yes (also works with OpenOffice)	Yes
EndNote	US\$249.95	Yes (also works with OpenOffice)	No
Mendeley	Free	Yes (also works with OpenOffice)	Yes
Citationsy	Free	Yes (also works with OpenOffice)	Yes

https://www.scribendi.com/academy/articles/reference_management_software_solutions.en.html

Task

- In teams
 - Determine the keywords to prepare the State of the Art for your project.

References



J. W. Creswell and J. D. Creswell.

Research Design: Qualitative, Quantitative and Mixed Methods.

SAGE, 2014.



M. E. Falagas and V. G. Alexiou.

The top-ten in journal impact factor manipulation.

Archivum Immunologiae et Therapiae Experimentalis volume, 2008.



M. Jalalian.

The story of fake impact factor companies and how we detected them.

Electron Physician, 2015.



**European
Funds**
Knowledge Education Development

**Warsaw University
of Technology**

European Union
European Social Fund



**MSc program in Data Science has been developed
as a part of task 10 of the project
„NERW PW. Science - Education - Development - Cooperation”
co-funded by European Union from European Social Fund.**