Do Not Let Them Trash Your Paper!

The Alchemy for Higher Chances to be Accepted

Vadim Ermolayev
Zaporizhzhia National Univ., Ukraine

Kherson State University, Ukraine
14 June 2019
Why this Guy?

• Can we trust him?
• Rival Trench
  – Trashed a lot and knows why …
• Chaired Program Committees, e.g. ICTERI
• Reviewed for:
  – Signature CS Conferences, e.g.: TheWebConf, ISWC, ER
  – Major Intl Journals, e.g.: IEEE IC, ACM TIIS, KAIS
  – European Commission: in FP6, FP7, H2020
• Authored/Edited:
  – Technical reports (international projects)
  – Workshop and Conference papers
  – Journal papers
  – Books
• Details: Scan QR Code – check his profile
Academic Writing

Course topics

- An Introduction into Academic Writing in CS/DS
- Planning and Structuring Your Manuscript for the Purpose
- Working with Literature. A Review of the Related Work
- Writing a Technical Report
- Writing a Thesis
- Writing and Publishing Papers. How to Learn to Write a Good Paper
- Talking and Presenting
- Overview and Summary

Prerequisites

None

https://cms.ucu.edu.ua/course/view.php?id=1430
Target Audience

• Tried, got accepted, willing to improve
• Tried but got rejected – interested why and how to succeed
  – Willing to fully understand the comments by reviewers
• Planning to try – looking for good practices
• Hesitating to accept PC membership
...
• All those who wish to improve their research papers
  – Academic performance
You are Actively Involved

• Feel free to **stop me at any time** for:
  - Posing a **question**
  - **Debating** a statement

• When you see anything **red** in a slide:
  - Feel invited to **react** – e.g. **answer**
Agenda

SCOPE: Conference Research Papers

- **Motives** (wish to)
- **Commitments** (have to)
- **Pre-requisites:**
  - Doing research (**Scientific Method**)
  - Choosing the **Venue** (estimating the chances)
- **Ingredients:** aspects to be addressed
- An **after-taste** check
  - Did your paper finally happen?
- **Final Remarks**
Disclaimer

• All citations, textual or pictorial, used in this slide-set are provided for illustrative purposes ONLY

• The use of these citations does not mean expressing any support or otherwise critique of the authors / owners and their respective views
  – Just a bit of attitude, perhaps …
Why 😁 would I Write this Paper?
A Pathway to a Solution

• Written texts, like professional papers, are for ... ?
  – Readers
  – Looking for a solution of their problem

• Your manuscript
  – A pathway for readers in a solution space
    – From their initial state (have the problem)
    – To the target state (have the solution)

• The better your manuscript is written
  – The shorter and more comfortable the path is
Good Reasons to Write

I **might wish** to:

- **Convey my results** to professional readers
- My peers **cross-evaluate** my results against theirs
- Put my brain dump in a **proper structure and order**
- **Transfer** my solution to the professional **community**, including **industry**
- **Impact** the technology
- Hoh, **be cited**

I **have** to:

- Do it **anyway** … as the **partial fulfilment** …
- I’m **appraised** based on what I write and publish

• **More reasons that brought you here?**
Poor Reasons to Write

• I will **then know what to do** in my project …
  - Exactly **on the contrary** …
  - Project **proposals**

• I dream to be **famous** …
  - You’d better **write fiction** then …

• My **super(visor/ordinate)** tasked me …
  - If she wants to get published, please **let her write** …
  - Teamwork is a **slightly different** thing

• Non-exhaustive list …
Pre-Writing

(before deciding if you wish to ...)
• “Let us first deliver and then ... we might wish to write a couple of papers ...”
  - Prof. Hans Akkermans, BI, VUA at an EU project meeting (last century)
  - One of the “giants” in CS

• Why is \textit{h-index} important?

• Why we might “\textit{wish to write}”?  
  - Getting a result is important, not \textit{writing} ... 

• Yet more important: 
  - How to \textit{do research} in a proper way?
How to: the Scientific Method

RESEARCH? - Getting knowledgeable about what has NOT been known

A pattern*: 

- **Very similar** in different sciences / branches 
  - Philosophy … CS … Medicine
- **Iterative**
  - Posing a Question (1)
  - Hypothesis (2)
  - Prediction (3)
  - Test (4)
  - Revision / Refinement (5)
  - Evaluation / Benchmarking (6)

Research Methodology (CS)

- A derivative of the Scientific Method
  - Subclass-of
- Different sorts – for various kinds of research
  - Theoretic
  - Experimental
  - Mixes / Synthetic

- **Pattern:**
Scientific Method ⇒ Structure

- **Iterative**
  - Question (1)
  - Hypothesis (2)
  - Prediction (3)
  - Test (4)
  - Revision / Refinement (5)
  - Evaluation / Benchmarking (6)

- **Introduction**
- **Motivation**
- **Problem Setting**
- **Related Work**
- **Approach to Solution** (background/transition/foreground/methodology)
- **Solution** (and Discussion)
- **Evaluation** (Objective, Plan, Data, Tool, Results, Discussion)
- **Conclusions, Recommendations, and Future Work**
Is your Result Publishable?

• = Demanded by the Readership
  - Helps them solve THEIR problem (Relevance)
  - Is, indeed, NEW (Novelty, Related Work)
  - Does BETTER than predecessors (Efficacy, Evaluation)
  - Available for USE and REPRODUCTION (Public Access)
  - Potentially, has IMPACT

• How to know?
  - Write a technical report
  - Present to your peers
  - Make it public (e.g. ResearchGate)
    - NDA (cross-check with your Legal)
  - Monitor the interest
Cross-Evaluation of Automated Term Extraction Tools

September 2017
Report number: TS-RTDC-TR-2017-1 · Affiliation: Zaporizhzhia National University, Ukraine
Projects: SemData: Semantic Data Management · OntoElec: a Methodology for Domain Ontology Refinement
Lab: Intelligent Systems

Cross-Evaluation of Automated Term Extraction Tools by Measuring Terminological Saturation

Authors: Victoria Kosa, David Chaves-Fraga, Dmitriy Naumenko, Eugene Yuschenko, Carlos Badenes-Olmedo, Vadim Ermolayev, Aliaksandr Birukou

This paper reports on cross-evaluating the two software tools for automated term extraction (ATE) from English texts: NaCTeM TerMine and UPm Term Extractor. The objective was to find the most fitting software for extracting the bags of terms to be...
Where to Publish?

• No universal answer – always a trade-off:

1. Choose a **RELEVANT** conference (scope)
2. Check **previous proceedings**
3. Compare your **contribution** (value)
4. Assess your **chance**
E.g. Impact vs Competition

- International Semantic Web Conference (ISWC)

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Publication years</td>
<td>1998-2018</td>
</tr>
<tr>
<td>Citation years</td>
<td>21 (1998-2019)</td>
</tr>
<tr>
<td>Papers</td>
<td>150</td>
</tr>
<tr>
<td>Citations</td>
<td>3733</td>
</tr>
<tr>
<td>Cites/year</td>
<td>177.76</td>
</tr>
<tr>
<td>Cites/paper</td>
<td>24.89</td>
</tr>
<tr>
<td>Cites/author</td>
<td>1257.78</td>
</tr>
<tr>
<td>Papers/author</td>
<td>56.56</td>
</tr>
<tr>
<td>Authors/paper</td>
<td>3.27</td>
</tr>
<tr>
<td>h-index</td>
<td>28</td>
</tr>
<tr>
<td>g-index*</td>
<td>56</td>
</tr>
<tr>
<td>h, norm</td>
<td>18</td>
</tr>
<tr>
<td>h, Annual</td>
<td>0.86</td>
</tr>
<tr>
<td>*Count</td>
<td>6</td>
</tr>
</tbody>
</table>

- H-index: 28
- G-index*: 56
- CORE: A-Conference
- Acceptance: < 10 %
- Charge quite some money as a conference fee

*g-index is the largest number such that the top g articles received together at least g^2 citations

Publish of Perish

June 17, 2019
E.g. Impact vs Competition

- ICT in Education, Research, and Industrial Applications (ICTERI)

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Help</th>
</tr>
</thead>
<tbody>
<tr>
<td>Publication years:</td>
<td>2012-2018</td>
</tr>
<tr>
<td>Citation years:</td>
<td>7 (2012-2019)</td>
</tr>
<tr>
<td>Papers:</td>
<td>150</td>
</tr>
<tr>
<td>Citations:</td>
<td>470</td>
</tr>
<tr>
<td>Cites/year:</td>
<td>67.14</td>
</tr>
<tr>
<td>Cites/paper:</td>
<td>3.13</td>
</tr>
<tr>
<td>Cites/author:</td>
<td>231.50</td>
</tr>
<tr>
<td>Papers/author:</td>
<td>83.85</td>
</tr>
<tr>
<td>Authors/paper:</td>
<td>2.35</td>
</tr>
<tr>
<td>h-index:</td>
<td>11</td>
</tr>
<tr>
<td>g-index:</td>
<td>14</td>
</tr>
<tr>
<td>hI,norm:</td>
<td>7</td>
</tr>
<tr>
<td>hI,annual:</td>
<td>1.00</td>
</tr>
<tr>
<td>*count:</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cites</th>
<th>Per year</th>
<th>Rank</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>h</td>
<td>27</td>
<td>6.75</td>
<td>1 N Kussul, A Shelestov, R Basarab, S S...</td>
</tr>
<tr>
<td>h</td>
<td>25</td>
<td>4.17</td>
<td>2 NV Morze, O Kuzminska, G Protsenko</td>
</tr>
<tr>
<td>h</td>
<td>19</td>
<td>3.17</td>
<td>3 V Kukharenko</td>
</tr>
<tr>
<td>h</td>
<td>16</td>
<td>4.00</td>
<td>4 H Kravtsov</td>
</tr>
<tr>
<td>h</td>
<td>15</td>
<td>3.75</td>
<td>5 IP Atamanyuk, Y P Kondratenko</td>
</tr>
<tr>
<td>h</td>
<td>14</td>
<td>2.00</td>
<td>6 E Lavrischeva, A Ostrovski, I Radetski</td>
</tr>
<tr>
<td>h</td>
<td>14</td>
<td>2.00</td>
<td>7 A Kolesnyk, O Slabospitskaya</td>
</tr>
<tr>
<td>h</td>
<td>14</td>
<td>4.67</td>
<td>8 OG Glazunova, T Voloshyna</td>
</tr>
<tr>
<td>h</td>
<td>12</td>
<td>3.00</td>
<td>11 IP Atamanyuk, Y P Kondratenko</td>
</tr>
<tr>
<td>h</td>
<td>11</td>
<td>1.83</td>
<td>9 VS Kharchenko, O Odarushchenko</td>
</tr>
<tr>
<td>h</td>
<td>11</td>
<td>3.67</td>
<td>10 A Sachenko, V Kochan, V Kharchenko,</td>
</tr>
<tr>
<td>h</td>
<td>10</td>
<td>1.67</td>
<td>12 D Bodnенко</td>
</tr>
</tbody>
</table>

- H-index: 11
- G-index: 14
- CORE: C-Conference
- Acceptance: < 50 %, this year – 30 % on average
- Free participation

June 17, 2019
Alchemy to Note: Pre-Writing

**Do Research**
- Scientific Method
- Relevant Methodology
- Get Result(s)

**Select Venue**
- Find Relevant
- Check Impact
- Check Competition
- Assess Chance

**Assess if Publishable**
- Present
- Technical Report
- Monitor Interest

June 17, 2019
Ingredients

(while writing ... )
To Pay Attention to

- Relevance and Readership
- Structure and Logic (Methodology and Story)
- Background, Approach, and Foreground
- Problem setting, Focus, and Contribution
- Related Work
- Terminology
- Rigor, Evidence, and Evaluation
- Illustrations
- Reproducibility
- Look and Fill (Template and Page Limit)
What is Relevance?

• **RQ**: Why Britons are ~20% taller than Japanese, on average? …

• Which of the **RA** would be **relevant** for ICTERI?

**Why?**

- We looked at what those Nations **eat** – Curry is 20% more energetic than Vasabi

- We checked the **excavation records** from X century on for those Regions – smaller people had ~20% more chances to survive in Japan

- We collected **digital medical records** for ~2 decades; we developed **software** to analyze these – Britons are genetically ~20% taller
Relevance, is it Important?

• Is Dietology your **readership**?
  - Would they **use** (and cite) your **result**?

• As a Computer Scientist, do you **contribute** to Archeology?

• To check **Venue relevance**:
  - Look into the **Scope** (topics of interest)
  - Check the previous **Proceedings**
  - Check the **Program Committee** (your community?)

• **Multi-disciplinary** research …
  - Check your **focus** and **contribution** vs the **Scope**
Story and Methodology

- Present the parts of your text in a **logical sequence** – map your **research workflow** into the paper

- **Why important?**
  - A reader expects you follow a **methodology** recognized in your field

- **Pattern:**
  - Situation-Problem-Process-Solution-Evaluation/Proof
  - Explain the **situation** and **motive** to improve it
    - Great if **shared** by your readership
  - Explain the **problem** that led to the situation
  - Look around – at the **related work**
  - Describe how your **approach** to the solution
  - Present your **solution** to the problem
  - Present your **evaluation**: was it correct, useful, efficient, effective, accepted by the victims, …
  - **Conclude**, **recommend**, and outline your **future work**
A different objective – **guide through** and **keep excited**

- A pattern to **convey your idea** through the narrative

**General to Specific**
- Start with a **general** statement / situation
- Elaborate in **more detail** / specifics
- Finally, put in the **context** of a **broader** statement

**Specific to General**
- Begin with a **specific focus**
  - E.g. event, a piece of art, an individual (a patient in a case report in Medicine), an organization (a company in a business case study)
- Progressively become **more general**

**Better Not to mix** in one paper
- There are cases when **you may wish** to, but **be wise**

Partitioning

• What would you feel if you see?:
  - Introduction
  - Main Part
  - Conclusion

Reviewer: This paper is not properly structured

• Parts become clear after defining your logical structure:
  - Introduction, perhaps including Motivation
  - Related Work
  - Problem Setting
  - Approach to Solution
    - Background-Transition-Foreground
  - Solution (and Discussion)
  - Evaluation (and Discussion)
  - Conclusions, Recommendations, and Future Work
• A.k.a. B, A, F
  - **Background**: what you had before doing this research
    - *Initial* state
  - **Approach**: how you did it
    - Research *path*
    - Through the *solutions* space
    - One of *many*
  - **Foreground**: what has been achieved
    - *Goal* state
• For example:
  - **B**: Informed that Britons are ~20% taller than Japanese
    - Had clinical data
    - Had programming skill
  - **A**: Developed software
    - Analyzed clinical data
  - **F**: Learned why — genetically
Approach

- Often underestimated, misunderstood, missing
- Not a Methodology
- Explains **how** the Methodology was
  - Chosen
  - Applied
- And **why** that way
- Reasons to choose:
  - The **mainstream** in the field
    - To compare with the Related Work
  - **Known do not help** solve the problem
    - Have to develop a new / improve the existing
Problem Setting

• A **formal** representation of a **situation** to be resolved
  - Got a rock (clinical data)
  - Have the motive (curious)
  - Have some background (coding)
  - Chose the approach
    - E.g.: remove all the unnecessary

• **Anything else?**

• Research **hypothesis**
  - Formally (e.g. a **theorem**)
    - E.g. Rodin: if you remove all the unnecessary, then you have a masterpiece

[Image]


https://www.nga.gov/collection/art-object-page.1005.html
Choose as appropriate for a conference paper (CS):

- We demystify the predominance of Europeans over Asians
- We explain why Britons are taller than Japanese
- We analyze clinical records to find out why Britons are 20% taller than Japanese
- ... Your variant (write down to further compare)
Focus

- **Choose as appropriate for a conference paper (CS):**

  - We demystify the predominance of Europeans over Asians
  - We explain why Britons are taller than Japanese
  - We analyze clinical records to find out why Britons are 20% taller than Japanese
  - … *Your variant*

https://www.sciencedaily.com/releases/2019/05/190521101505.htm
• **Choose as appropriate for a conference paper (CS):**
  
  - We demystify the predominance of Europeans over Asians
  
  - We explain why Britons are taller than Japanese
  
  - We analyze clinical records to find out why Britons are 20% taller than Japanese
  
  - … *Your variant*

http://4usky.com/download/164832620.html
Focus

• Choose as appropriate for a conference paper (CS):

  - We demystify the predominance of Europeans over Asians
  - We explain why Britons are taller than Japanese
  - We analyze clinical records to find out why Britons are 20% taller than Japanese
  - ... Your variant

UK: https://www.letsbebrief.co.uk/the-guardian-the-uk-election-results-map/
• Choose as appropriate for a conference paper (CS):

- We demystify the predominance of Europeans over Asians
- We explain why Britons are taller than Japanese
- We analyze clinical records to find out why Britons are 20% taller than Japanese
- Using data analytics and following the XYZ methodology, we analyze the digital DNA records of patients. This data has been collected in the last 2 decades in Britain and Japan. We aim to explain why Britons are ~20% taller than Japanese.

https://www.amazon.in/Big-Data-Analytics-Bioinformatics-Healthcare/dp/1466666110
Focus: Aspects

- Disciplinary
- Temporal
- Spatial
- Methodological
- Situational
- Pragmatic

- To be chosen as **precisely** as possible
- To be cross-checked by **relevance**

Using data analytics and following the XYZ methodology, we analyze the digital DNA records of patients. This data has been collected in the last 2 decades in Britain and Japan. We aim to explain why Britons are \( \sim 20\% \) taller than Japanese.
Contribution

• Given:
  - The problem
  - The focus

• C: How do you solve the problem in focus?
  - Compared to the Related Work?

• C: Oh, what is NOT solved?
  - Equally important

http://www.freeworldmaps.net/europe/ukraine/political.html
https://www.backmarket.com/refurbished-smartphones.html
What is Related Work?

• **How would you denote?**
  - The **stuff** in the RW section
  - **Similar work** people did
  - The work by some others, **not helpful**, just related
  - The **results** to which to be **compared** (cross-evaluation)
  - Something by the “**giants**” who will review my work
  - The mentions of the **groups** working on the same / similar problem
  - The mentions of your **own prior work** or the work by your **cooperatives**
  - The mentions of the **members** of your M.Sci / Ph.D committee
  - **Anything missing?**
Related Work?

• How would you denote?
  – Similar work people did
  – The results to which to be compared (cross-evaluation)
  – The mentions of the groups working on the same / similar problem - their results

• Anything missing?
  – A roadmap to your idea and contribution
Reasons to Review the RW

- Demonstrate that you know what happens in your field
- Showcase the recent achievements in your field
- Put your own work in the context within your field
- Outline the highlights and lowlights of the results by the others
- Tame your footprint towards extending the State-of-the-Art – the contributions
- List those who do similar things – to raise their interest
  - E.g. in cross-evaluation
- L.B.N.L, pay respect to your community
  - No matter if they are GIANTS or PEERS
  - To be respected as a member of your specialist community ...
  - in return
- **EXAMPLES** … good and bad reviews
We are sorry to confess that we **failed to find any published reports dealing with similar research problems and settings.** More specifically, the research on measuring the impact of the order of adding documents on terminological saturation, and also a bit broader – on investigating how different orders influence the quality of automated term extraction, has not been published yet, to the best of our knowledge.


- **Is it good or bad? Why?**
  - **Good:** if true, it might be a revolution in science
  - **Bad:** more likely, they missed something …
The problem of informatisation of education, application of information and communication technologies (ICT) in the educational process is the subject of research of many Ukrainian scholars. For instance, V. Bykov, R. Gurevich, M. Kademii, L. Nakonechna, L. Petrova, S. Sysoeva study the development of the conceptual apparatus on informatisation of education; … Dzyuba, M. Zhaladak, S. Kizim, S. Loboda; I. Vorotnikov, A. Kocharyan, N. Morse, S. Petrenko – peculiarities of formation of information competence of future teachers by means of ICT; …[2].


- **Is it good or bad? Why?**
  - Also **non-existent** – with all respect, names are not results
…The best known lossy compression techniques are JPEG [10], Zstandard by Facebook [11] and algorithms based on Singular Value Decomposition (SVD). … One of the well known lossless compression techniques is Discrete Wavelet Transformation (DWT). In this work, we are proposing an overview of a lossy algorithm combining SVD and DWT [1]. …

One of the CourseWork Reports submitted in the Linear Algebra course at UCU CS-DS M.Sci Program

- Is it **good** or **bad**? **Why**?
  - Are those all relevant results?
In the recent surveys on text similarity measurement approaches, e.g. [31], [32], methods (or measures) are grouped based on analysing: (i) characters and their sequences; (ii) tokens; (iii) terms; (iv) text corpora; or (v) synsets. In [32] hybrid measures that allow fuzzy matching between tokens are also mentioned. Brief characteristics of the groups are given immediately below. The individual methods belonging to the groups are detailed in Table 1. …


• Is it good or bad?
  - Good – mention alternative paths
E.g.4: Detail and Compare (2/4)

**Table 1: The overview of text similarity / distance measures**

<table>
<thead>
<tr>
<th>Name, source</th>
<th>Description</th>
<th>Specifics</th>
<th>Relevance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Longest Common Substring [33]</strong></td>
<td>common character sequence based measure</td>
<td>returns the integer length of the longest common substring; could be normalized by the total length</td>
<td>moderate</td>
</tr>
<tr>
<td><strong>Levenshtein distance [17]</strong></td>
<td>edit distance based measure</td>
<td>returns an integer number of required edits</td>
<td>marginal</td>
</tr>
<tr>
<td><strong>Hamming distance [25]</strong></td>
<td>edit distance based measure</td>
<td>strings have to be of equal length</td>
<td>marginal</td>
</tr>
<tr>
<td><strong>Monger-Elkan distance [18]</strong></td>
<td>edit distance based measure</td>
<td>returns an integer number of required edits</td>
<td>marginal</td>
</tr>
<tr>
<td><strong>Jaro distance [19]</strong></td>
<td>counts the minimal number of one character transforms in one</td>
<td>returns a normalized real value from [0, 1]</td>
<td>good</td>
</tr>
</tbody>
</table>

- **Is it good or bad?**
  - **Good** – most relevant in more detail
In difference and complementary to the abovementioned relevant work, we contribute several novel things. Firstly, we propose a way to rationally choose the thresholds that are used to regard string similarity as term similarity (Section 3). Secondly, we develop an algorithm for similar terms grouping that uses string similarity measures and term similarity thresholds (Section 4). Based on its use, we propose the refinement of the baseline THD algorithm [2] for measuring terminological difference between two subsequent text datasets (Section 4). …

• Is it good or bad? Why?
  – Good – to tame a land and motivate to read further
In Philosophy and Computer Science, time has been and is being taken as a fundamental aspect in representing and reasoning about change and dynamics. In this paper however we are interested not in the gems of the philosophical thought, but in the theories that are in use – in applications. Many theories comprising a temporal aspect have been developed, covering the contexts of change and actions, events, situations, communication, execution and control, distributed real-time systems, planning and processes, data and information mining, cognitive sciences and linguistics, etc. The most prominent of these theories are [Lamport (1978)][8651]*; [Allen (1983)][7894]; [Pinto (1994)][220] based on [Kowalski and Sergot (1986)][1708]; [Prior (1967)][1496]; [McDermott (1982)][1130]; [Sandewall (1995)][404]; [Halpern and Shoham (1991)][389]; [Bacchus and Kabanza (1998)][230] based on [Alur et al. (1996)][399]; [Williams (1986)][198]; [Koubarakis (1992)][55]; [Iwasaki et al. (1995)][52]. …

* References are ordered by the number of citations (given in round brackets) as of Aug. 24, 2014. Source: Google Scholar.

• Is it good or bad? Why? Good - methodology

Tips to Consider

• All most **up to date** and **influential** to be mentioned – explain the **choice(s)**

• Good to give the evidence of **completeness**

• Analyze, **detect gaps** regarding the **problem**
  – Outline **strengths**
  – Mention **weaknesses / shortcomings**

• Pick the **short list** of the most relevant
  – Compare in more detail – table or diagram

• Outline **your contribution** compared to the RW

June 17, 2019
Use of Proper Terms

• **E.g.:** imitational modeling / simulation modelling
  - To be: **Simulation**
• **E.g.:** own value (in mathematical sense)
  - To be: **Eigenvalue**
• **E.g.:** plurality (in mathematical sense)
  - To be: **Manifold**
• **E.g.:** country in smartphone
  - To be: **eGovernance**
• Wrong terminology – the most blatant fallacy
  - Often interpreted as:
    - Lack of basic education, skills (e.g. English)
    - Lack of professional knowledge
• For having proper terms:
  - Never trust proof-reading companies
  - Read mainstream publications in your field

• Proper terms indicate **rigor** in your work
Rigor

- In CS (science)
  - No room for plurality / ambiguity in interpretation
- Clear formulations
  - Focus, Problem, Method, Solution
- Use of proper notation
  - Formal statements and proofs
  - Diagrams
  - Pseudocode
- A signature of being professional
Evidence

• Does not work in Science
• You have to say **HOW** and **WHY**!
• You have to provide facts that **prove** your declaration
• Examples:
  - A **formal proof** of a statement
  - A **ground truth** that supports the claim
  - An **experimental result** that proves the hypothesis

Does it offer any **evidence**?

It says:
**New Course for Ukraine**
**Peace and Security**
Julia Tymoshenko
([http://yahovor.arbat.name/?p=13037](http://yahovor.arbat.name/?p=13037))
Lemma 1 (The total frequency of nested occurrences). The total value of the frequency of nested occurrences, in $D$, of a term candidate string $s$ in longer term candidate strings $ls$ is the sum of the total frequency values of nested occurrences in all partial collections $D_i$ of $D$:

$$tnest(s) = \sum_{ls \in T^s} f(ls) = \sum_{i=1}^{n} \left( \sum_{ls \in T^s_{i}} f_{i}(ls) \right) = \sum_{i=1}^{n} (tnest_{i}(s)). \tag{4}$$

Proof. It implies from Definition 2 (of partial c-value), that $tnest_{i}(s)$ is the total number of occurrences of the term candidate string $s$ in all longer term candidate strings $ls$ extracted from the partial collection $D_i$. The number of these longer term candidate strings equals to $P(T^s_{i})$. Due to the disjointness of the partial collections $D_i$ (Condition 2 of Definition 1), $f(ls) = \sum_{i=1}^{n} f_{i}(ls)$. Therefore, and due to the Condition 1 of Definition 1, the total number of occurrences of $s$ in all $ls$ extracted from $D$ is:

$$tnest(s) = \sum_{ls \in T^s} f(ls) = \sum_{ls \in \bigcup_{i=1}^{n} (T^s_{i})} (\sum_{i=1}^{n} f_{i}(ls)) = \sum_{i=1}^{n} \left( \sum_{ls \in T^s_{i}} f_{i}(ls) \right) = \sum_{i=1}^{n} (tnest_{i}(s)).$$

Evidence: Example

Experimental evidence*

- In **all** compared Bags of Terms (conv. – vs – merged)
  - **Terminological difference** less than **individual term significance threshold**
  - Also confirmed by superimposed **saturation curves**

- *(h1) holds valid*
- **Theorem 1** proven

---

Evaluation

• What is Evaluation?
• Finding if there is a value …
• Checking if
  – What you found
  – Corresponds to the claim
    (what was promised)
• Different types:
  – Formal – check the proof, find contradiction
  – Experimental – run / repeat stuff, compare result
• Reproducibility …

Would you pay € 9.75? How to find out?
Let Them Do What You Did

- A way to **convince**
- A kind of **evidence**
- An invite for **collaboration**
- A way to **attract** early adopters
- A way to **pay respect**

E.g.:

- All **code** and **data** related to our study are available for research purposes.
  - Data: [http://dx.doi.org/10.17632/knb8fgyr8n.1](http://dx.doi.org/10.17632/knb8fgyr8n.1), available under a Creative Commons Attribution 4.0 International license.
Evaluation findings to be discussed, in sufficient detail:
- If you request evidence
- Then you have to provide evidence, in reply (also reviewers)
- Both to be sufficient for reproduction

E.g.: Rioja Vega Reserva, experimental:
- I risked € 9.75 (experimental setting and environment)
- I invited my 5 best friends (statistically representative)
- We had a wonderful party (properly measured)
- We ordered THREE more (measurable outcome)
- Great value for the price (objective assessment)
Types of Illustrations

• What is an illustration?
  - A figure (chart, picture, etc.);
  - But also: a table; an example; a fact; a link
  - A piece that supports your story/statement

• Which types of illustrations are appropriate?
  - All those above, that help you
    - Save space (a look is worth a thousand words)
    - Make your explanation simple and clear

• Make sure that
  - Use permissions are granted
  - Sources are mentioned:
    - Like the above: https://en.wikipedia.org/wiki/A_picture_is_worth_a_thousand_words
Template and Page Limit

• How would you choose your wine?
• Same for your reviewer
  - How it looks, which implies
  - How it tastes
• A paper to look nobly and brandy
  - Right template to be used
    - Label is white - content is right
  - Page limit to be strictly obeyed
    - Never pay full price for half a bottle
    - Never > 0.75 l
• Not easy
  - have to give up things …

June 17, 2019
Alchemy to Note: While Writing

Structural and Logical Approach
- Background
- Transition
- Foreground
- Explain
- Explain the Link
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
- Gap
- Formalization
- Motive
Beaujolais Nouveau or Rioja Reserva?

(after writing ... did the paper happen?)
Your research methodology is the **SCENARIO**

https://www.slideshare.net/garr/sample-slides-by-garr-reynolds/
E.g. Exposition & Climax

Motivation: What is this about?

- What is this proceedings volume about?
  - Without reading ...
- A word cloud?
  - Of-the-shelf services available
    - e.g. http://www.wordclouds.com/
    - Too long text - stop words pop up
- A word cloud on abstracts?
  - Also not perfect:
    - Displays single words only
    - Terminology is often multi-word
  - So, trust keywords?
    - Too coarse:
    - No way to see which word is used more or less often

Motivation: Problems with ATE

- Does not scale well
  - E.g. you want to know what is Knowledge Management about
  - Need to extract terms from gigabytes of text
  - Springer KM
    - 15 journals; > 9 000 papers
    - 7 554 papers in full texts (1997 - 2016)
- Need an approach to extract:
  - A representative sub-collection
  - Of minimal size
- Solution: OntoElect – terminological saturation

Motivation: ATE

- Automated Term Extraction
- Feed your full text (e.g. proceedings) into the extractor
- Get the ranked list of multi-word terms
  - Without stop words
  - Cut off insignificant terms

General statement

↓

Elaboration on the statement

↓

More detailed elaboration

Broader statement

E.g. Resolution

“Simplicity is about subtracting the obvious, and adding the meaningful.”

— John Maeda

What to **SUBTRACT**?

https://www.presentationzen.com/presentationzen/2006/11/presentations_a_1.html
It was Easy ... that Far

• If you knew the **method**
• If you had some **experience**
• If you had some **resource**
• If you had a bit of **talent**
• And ... if you **did the work**

**How would you achieve THIS?**

• So you **arrived** at ...
• And the **goal** was ...

June 17, 2019
Rodin and ... the Gang

• Albert Einstein
  - A German-born theoretical physicist
  - The theory of relativity
    - One of THE two pillars of modern physics

• A 3 step procedure: ...
  - 1. Take a Universe
  - 2. Think
  - 3. Write …

How to learn THIS?

E = mc²
Yeah, ... the Tips ...

• Find a **Super Star** in your field
• Take your **time**
• **Observe** how s/he does
• **Learn**
• **Try**

• **... Suffer ...**
• Do not **mimic**
• Be **yourself**
• Become a **SUPER NOVA**

• Perhaps, the **only** feasible **way** is ...

Apprenticeship and Responsibility

Sorry, that ABSTRACT
“It’s taken me all my life to learn what not to play.”

— Dizzy-Gillespie

https://www.slideshare.net/garr/sample-slides-by-garr-reynolds/
Check the After-Taste …

• After some time passed …. are you still happy about your paper?
• Do colleagues read it?
• Do they cite it?
• May well surprise you …
Alchemy to Note: After Writing

June 17, 2019
Final Remarks
Putting All Together

You have to work hard

**Before** writing

- Do Research
  - Scientific Method
  - Relevant Methodology
  - Get Results

- Select Venue
  - Find Relevant
  - Check Impact
  - Check Competition
  - Assess Chance

- Assess if Publishable
  - Present
  - Technical Report
  - Monitor Interest

**While** writing

- Structure and Logic
  - Methods as a Pattern
  - Intentionality
  - General vs Specific
  - Partition

- Text
  - Read Scope
  - Proceed
  - Proof

- Illustration
  - Figure
  - Table
  - Example
  - Fact
  - Code
  - Help save space

- Analysis
  - Motive
  - Gap
  - Comparison

- Terminology & Rigor
  - Proper terms
  - Mainstream
  - Accuracy
  - Ambiguity
  - Proper formal notation

- Reproducibility
  - Discussion

**Problem & Focus**

- Background
- Transition
- Foreground

- Explain
- The
- How

**Approach**

- Template
  - Template to strictly follow
- Page limit
  - Look implies
  - Feel

- RW & Contribution
  - Results by
  - Others
  - Motive

- Keep it Simple
  - Keep important
  - Sacrifice
  - Obvious
  - Suffer
  - Be responsible

**After** writing

- Check the Story
  - Exposition
  - Conflict & Climax
  - Resolution

- Check the After-Taste
  - Are you happy?
  - Is it read and cited?
  - Be surprised

Keep on suffering … Start early …

**Anything else to add?**
Start Early: M.Sci Students

1st Masters Symposium on Advances in Data Mining, Machine Learning, and Computer Vision (MS-AMLV 2019)
November 15-16, 2019, Lviv - Ukraine

Proceedings: CEUR-WS
Venue: Ukrainian Catholic University
Submission: https://easychair.org/conferences/?conf=msamlv2019

- Monday, 16.09.2019 – paper abstract submission deadline
- Monday, 30.09.2019 – paper submission deadline
- Tuesday, 05.11.2019 – acceptance notification
- Monday, 14.11.2019 – camera-ready papers submission deadline
- Monday, 14.11.2019 – registration deadline
- Friday, 15 – Saturday, 16.11.2019 – symposium days

- CfP printouts available for pick-up
Yeah, if You Still WannAdvise

“Master your instrument. Master the music. And then forget all that bullshit and just play.”

—Charlie-Parker

https://www.slideshare.net/garr/sample-slides-by-garr-reynolds/
• Many thanks to all those, whose
  - Images
  - Slide samples and fragments
  - Papers
  - Ideas
  - Recommendations
  - Statements
  - Practices
  - Political slogans
  I used or cited while preparing this tutorial
• All the bits are carefully referenced throughout
  this slide set
Now happy to take your questions ...

also happy to have a coffee and chat ...

vadim@ermolayev.com