

Academic / Technical Writing

in Computer Science

Topic 4:

Writing a Technical Report



Vadim Ermolayev

Zaporizhzhya National Univ., Ukraine

vadim@ermolayev.com

Ukrainian Catholic Uni., Lviv, Fall 2018
M.Sci Programme in CS - DS



Plan

- What is (**specific** in writing) a **Technical Report**?
 - What to **put in** and what to **leave out**?
- How to:
 - **Structure**
 - **Narrate** (telling your story)
 - **Illustrate**
- How (and why) to make it **reproducible**?
- Why **teamwork** matters
 - How it should be **arranged**

Credits

- To **all** the **colleagues** whose technical reports I read
- To **learn** to do myself

What is a Technical Report?

A Technical Report

- A **basic component** of R&D output
- Written on a **project**, as a partial fulfilment...,
 - To document your **accomplishment(s)**
 - To mark up a **milestone**
- **Motives** to write (wish to):
 - **Inform the community** about my progress
 - **Tame a land**
 - **Share important results** with my community
 - Have a **fully detailed backup** for my future publications
 - Get **feedback** from the community
 - Later, **reuse** it for my **thesis**
 - ...
 - **Report** to your funding body

A Technical Report

- A **basic component** of R&D output
- **Purposes** to write (have to):
 - Contractual **commitment** = otherwise I will not finish my M.Sci / Ph.D **program**
 - I **promised** to my mentor = **ethical** commitment
 - I have to do it for my **career** = your **Hall of Fame**
- **Important:**
 - **Take care** about (beware of) those who **may wish** to read it
 - Do **not disappoint** people ...

A Technical Report

- A written **manuscript**
 - Formatted as requested by your funding body
- Made **publicly available**
 - If your funding body allows
- **Self-contained** – for a potential reader
 - All the bits s/he may wish to have to **fully understand** your story
- Assembled in a way to **allow** a reader to **reproduce** the reported accomplishments / **results**
 - Containing all necessary appendices / addendums / references to Your:
 - **Solution**
 - **Instruments**
 - **Data**
 - **Experimental results**

Specifics

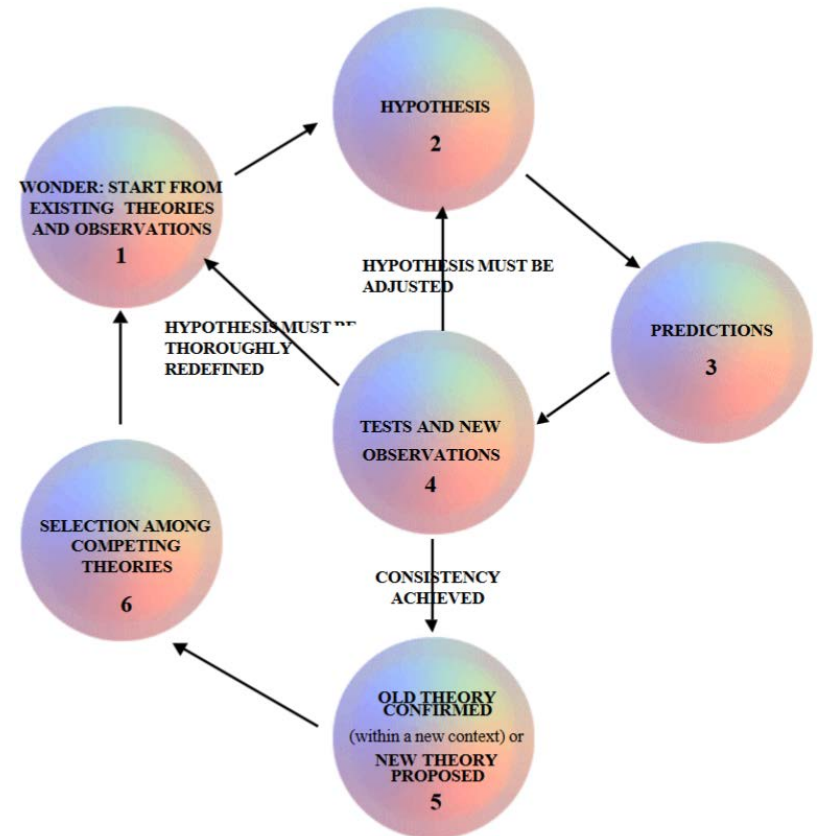
- A primary source to **evaluate your work**
 - For your **peers**
 - For your **supervisors**
- Not published
 - You, the author(s), bare **full responsibility** for the quality
- To be as **short** (and **sweet**) as **possible**
 - Any **beautifiers** to be brutally **removed**
 - Keep it **technical**
- To be as **long** (and **sweet**) as **needed**
 - Use **as much space** as you find **necessary** to convey **all the detail** to your reader

Tips on Structuring

Methodology of Science

An **outline*** :

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

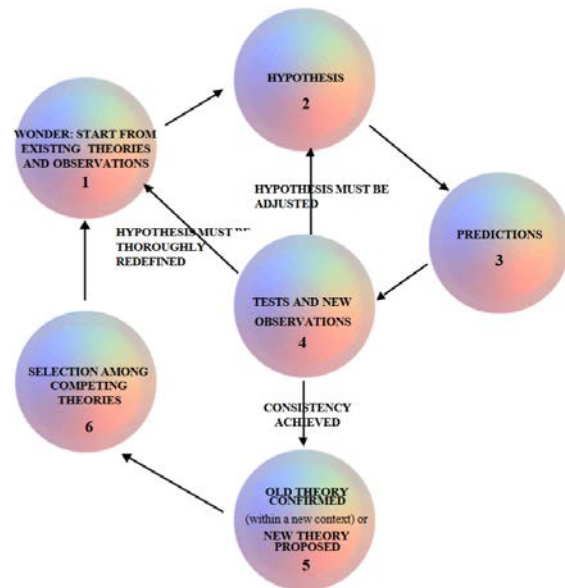


* Dodig-Crnkovic, G.: Scientific Methods in Computer Science. Proc Conf for the Promotion of Research in IT at New Universities and at University Colleges in Sweden (2002)

Scientific Method \Rightarrow Structure

- **Iterative**

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



- Introduction
- Motivation
- Problem Setting
- Related Work
- Approach to Solution (background/transition/foreground)
- Solution (and Discussion)
- Evaluation (Plan, Data, Tool, Results, Discussion)
- Conclusions, Recommendations, and Future Work

A Question ...

- **What is a research question?**
 - Question asked in the **context** of **what is known** in your research field
 - **Existing** knowledge: **theory** and/or **observations**
 - A question that **existing knowledge can answer** (usually the case)
 - A question that **calls for a new theory** (rarely)
- Hence, **to pose a question** you have to have:
 - Your knowledge of **existing knowledge**
 - Your knowledge about the **gaps in the existing knowledge**
 - Your motive to **fill the gap**
- Introduction
- **Motivation**
- **Problem Setting**
- **Related Work**
- Approach to Solution (background/transition/foreground)
- Solution (and Discussion)
- Evaluation (Plan, Data, Tool, Results, Discussion)
- Conclusions, Recommendations, and Future Work

A Hypothesis ...

- What is a **hypothesis**?
 - An **envisioned answer** to your research question
 - Could be
 - Wrong / Correct
 - Incomplete / Complete
 - Biased, Subjective / Objective
 - An **educated insight**
 - Based on **what**?
 - Your knowledge of **existing knowledge**
 - Your knowledge about the **gaps in the existing knowledge**
 - Your insight on how to **fill the gap**
- Introduction
- Motivation
- **Problem Setting**
- **Related Work**
- **Approach to Solution**
(background/transition/foreground)
- Solution
(and Discussion)
- Evaluation
(Plan, Data, Tool, Results, Discussion)
- Conclusions,
Recommendations,
and Future Work

A Prediction ...

- **What to predict?**
 - Suppose your **hypothesis** was **right**
 - Deduce **consequences**
 - Predict if the **goal state** achieved
- **What do you need to do?**
 - Assess **potential impact**
 - **Breadth** and **depth**
 - Check against the **goal**
 - What happens to the **gap**?
- Introduction
- Motivation
- **Problem Setting**
- Related Work
- **Approach to Solution**
(background/transition/foreground)
- **Solution**
(and Discussion)
- Evaluation
(Plan, Data, Tool, Results, Discussion)
- Conclusions,
Recommendations,
and Future Work

Test Hypothesis ...

- **How to test?**
 - Prove, **formally**
 - Evaluate, **experimentally**
 - Plan, Method
 - Data
 - Tool
 - Results
 - Analysis (Discussion)
 - Check:
 - **Background conformance**
 - **Validity bounds**
 - **Correctness of transition**
- Introduction
- Motivation
- Problem Setting
- Related Work
- **Approach to Solution**
(background/transition/foreground)
- **Solution**
(and Discussion)
- **Evaluation**
(Plan, Data, Tool, Results, Discussion)
- Conclusions,
Recommendations,
and Future Work

Revision / Refinement

- **Why?**
- **Hypothesis**
 - Wrong
 - Incomplete
 - Biased, Subjective
- Your **educated insight**
 - Superficial
- Some **impacts**
 - Overestimated
- Some **tests**
 - Negative
- **Ways to improve?**
 - **Revise** – do more work, rewrite report
 - **Refine** – do more work, add (a) section(s)
 - **Plan improvement** – discuss, plan future work
- Introduction
- **Motivation**
- **Problem Setting**
- **Related Work**
- **Approach to Solution**
(background/transition/foreground)
- **Solution**
(and Discussion)
- **Evaluation**
(Plan, Data, Tool, Results, Discussion)
- Conclusions, Recommendations, and **Future Work**

Priorities in Structuring

- I – Logic & Methodology
 - Why **I-st**?
- II – Intent (story)
 - A pattern to **convey your idea** through the narrative
 - **General** to **Specific**
 - **Specific** to **General**
 - **NEVER** be **mixed** in one paper
 - Throughout the **document**
 - Make your parts **intentionally linked**
 - E.g. by offering a **run-through example**
 - At a **section** level
- III – Partitioning – as discussed

Section Structuring

Introduction

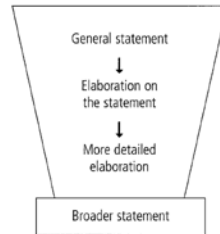
- Explain what is your **field**
- Explain what is your **focus** in the field
- Outline what is the **problem** you address
- Explain why **solving** the problem is **important**, for **whom**
- Outline the **parts / sub-problems** that have **not** been **solved**
- Explain the **organization** of the remainder of your report
 - How you **structure the story** about your **solution** to **THE problem**

Introduction

- **General to Specific**

- E.g.

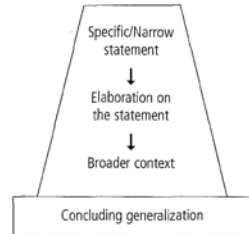
- ATE is an important and vibrant research area in Text Mining
- A lot of effort is being spent to improve its quality
- Quality is still unsatisfactory
- One important issue – the completeness of your text collection
- This report presents an approach to measure the completeness
- The remainder of the document is structured as follows
 - Section 2 presents our motivation to ...
 - ...
 - Finally, Section N concludes the report ...



- **Specific to General**

- E.g.

- ... disk drive capable of storing all the world's music is now worth US \$600; 30 billion of content pieces are shared monthly at Facebook ...
- ... a dramatic increase in social networking applications that allow ...users create a huge amount of content easily
- Society is, however, left bewildered about how to use these data efficiently and effectively
- ...



Motivation (RW – Topic 3)

- Voila: this is THE **problem**; in this particular **context**
- It is **important** to have a **solution**
 - This is your **vision** that the **impact** will be extraordinary / substantial / noticeable
- In **addition to the vision**, we are **prepared** to develop a solution
 - Here is our **background**
 - Here is the **idea**
 - We expect that **following** the idea **this way**
 - will bring us about **THIS foreground**
- We expect that this **foreground** will allow the **early adopters** to have **THIS impact ... there and there**
- Could be also done
 - **General to Specific**
 - **Specific to General**

Problem Setting

- Based on the **background** and **motives**
 - The **preliminaries** ...
- Here we go: our **research questions** are...
- We've got an **idea** that the **answers** might be ...
 - The **hypotheses**
- We now **formalize** the hypotheses arriving at **statements**
 - A **Theorem** ...
 - A Set of **Requirements** ...
 - ...
- Statements to be further **verified**

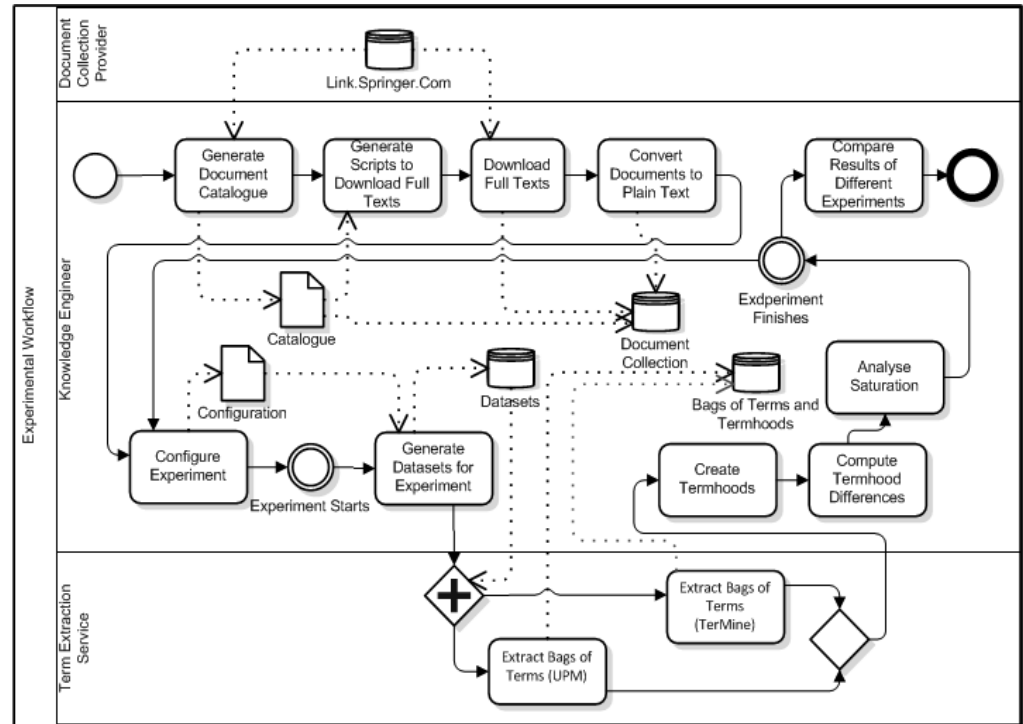
Approach to Solution

- We have:
 - The **Background**
 - The **Problem Statement**
- Now we explain how we go through the **Solution Space**
 - Using the **Background**
 - Collecting the bits of the **Foreground**
 - Assembling the **Solution**
 - **Measuring** our **proximity** to the **Goal**
- E.g.
 - Background: We have a baseline algorithm to measure THD
 - Problem: low quality; slow terminological saturation
 - Idea: do similar terms grouping (STG)
 - Approach:
 - 1: Choose proper measures
 - 2: Choose reasonable similarity threshold(s)
 - 3: Develop STG algorithm
 - 4: Refine baseline THD
 - 5: Cross-evaluate

Approach to Solution

- Or **show** and **explain** your research **workflow***

- Right place to present also your
 - **Tools**; and
 - **Resources**



* Kosa, V., Chugunenko, A., Yushcenko, E., Badenes, C., Ermolayev, V., Birukou, A.: Semantic Saturation in Retrospective Text Document Collections. . In: Mallet., F. and Zholtkevych, G. (Eds.): Proc ICTERI 2017 PhD Symposium, Kyiv, Ukraine, May 16-17, 2017, CEUR-WS vol. 1851, p. 1-8, online

Solution and Discussion

- You followed the **Approach**
- You found out that the **goal** has been (somewhat) **achieved**
- Now **share** your **feelings** in the goal state
- Voila:
 - The **solution** to the **stated problem**
 - And its **discussion**
 - I **claim** that it **does this**, in this **context**
 - However, it does **NOT do that**
 - Yet however, it does **not do this**
 - » If the **context** is **different**
 - I will now **evaluate** my **claim**, e.g. **experimentally**

Evaluation

- No time to go into the specifics of **experimental method**
 - Suppose **you are informed**
 - Ample **sources**
 - E.g. <https://www.inf.unibz.it/~calvanese/teaching/2016-01-PhD-RM/RM-2016-M5-gamper.pdf>
- For **documenting***:
 - The **Plan** of your experiment
 - The **Data** you use in your experiment
 - The **Tools** and **Environment** in your experiment
 - The **Results**
 - The **Analysis** of your **Results**

* Kosa, V., Chaves-Fraga, D., Naumenko, D., Yuschenko, E., Badenes-Olmedo, C., Ermolayev, V., and Birukou, A.: Cross-Evaluation of Automated Term Extraction Tools. Technical Report TS-RTDC-TR-2017-1, 30.09.2017, Dept of Computer Science, Zaporizhzhia National University, Ukraine, 61 p.

Conclusive Remarks

A **Summary** in 3 interlinked facets

- **Conclusions**

- What has been done – **result(s)**
- **Highlights**
 - Why the result is better of than the related / prior work
- **Lowlights**
 - What is, still, **improvable**
 - What is an **inherent weakness** and cannot be improved
- What are the **boundaries** of effective **applicability**

- **Recommendations**

- Would you recommend a **use** to your community?
- **What** and **how** exactly?

- **Future Work**

- Which **improvements** / further **research** are planned
 - In **short-term** / **long-term** time

Illustrations

Types of ...

- 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



1913 newspaper advertisement

Reproducibility

Reproducibility?

- What is **reproducibility**?
 - The **feature** of your report
 - Allows the **others** to **reproduce** what **you did** and report
- Is it **important**? For **whom**? **Why**?
 - **Important**: a more trusted result / solution, if reproduced
 - For **whom and why**:
 - For **you** – you may become more trusted, hence reputable
 - For **your peers** – they may use your result to cross-evaluate with their own
 - For the **early adopters** – they normally (cross-)check before considering to use
- **Reproducibility** – one of the **signatures** of **engineering**, not **craft**

What to Put In?

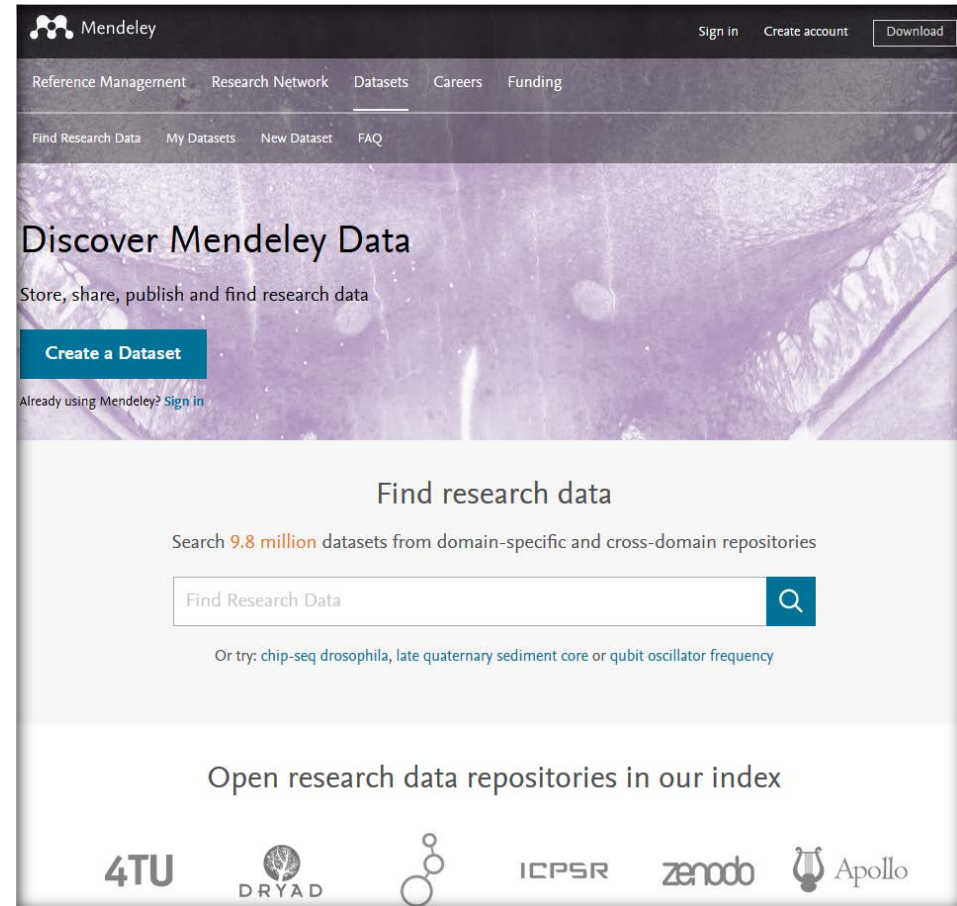
- **Opinions? ...**
 - All your **peers** may need to **reproduce** your result
 - All your **adopters** may need to **use** your result
 - In their **environment**
- In particular:
 - **Solution** (algorithm + code)
 - **Datasets** you used
 - **Additional tools** (instruments) you used
 - **Plan** of your experiments
 - **Requirements** to operating environment
 - Hardware + software
- All in a **directly usable** form

Code / Tool Sharing

- Which you all know well, e.g.
 - GitHub
 - Ample other code Hubs

Data Sharing

- One may use:
 - **Code sharing** hubs
 - **File sharing** hubs
- There are also **data sharing** hubs
 - E.g. **Mendeley Data**
 - <https://data.mendeley.com/>
- Open Science Initiative
- https://en.wikipedia.org/wiki/Open_science



Teamwork

Why Important?

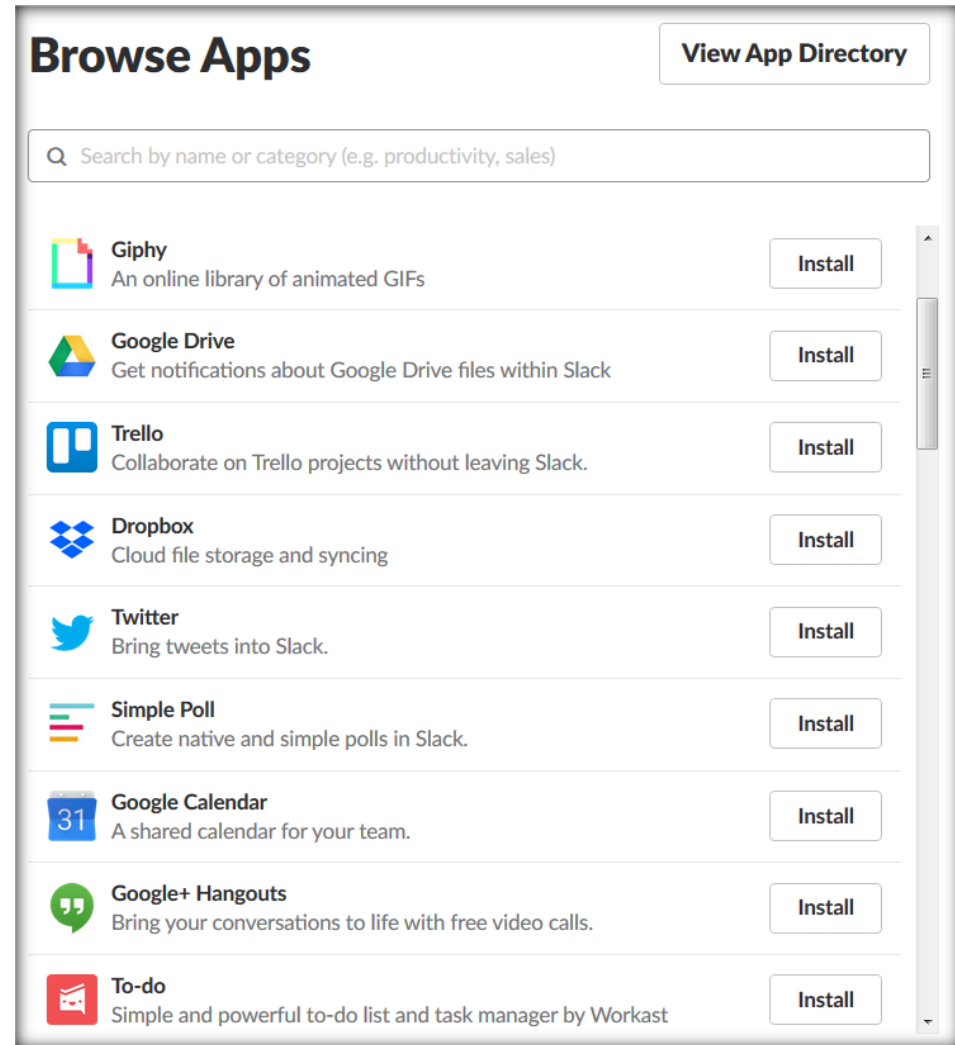
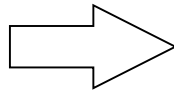
- Have you ever been on a project **alone**?
 - Presumably **not**
 - Even if you think that yes
 - At least you had your **supervisor**
- Often, you are **dispersed**
 - So, E-TeamWork

Patterns

- **Supervised** editing
 - One author supervises the work and contributions of the others
 - E.g. Google Docs
- **Peer-to-Peer** editing
 - There is no boss among them
 - E.g. Google Docs
- **Supply-Write-Edit** chain
 - Teammates provide the bits of project outputs (workers)
 - Other teammates write fragments (drovers)
 - The **editor** puts all together and harmonizes (owner)

Tools / Environments

- Communication & collaboration
 - Like Slack.com
- Collaborative Editing
 - Like Google Docs ...
- Document / File Sharing
 - Like Google Drive ...
- Social Networking
 - Like ResearchGate ...



**Oh, Yeah ...
What
to Leave Out?**

It was Trivial ... 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 reach
THIS?

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



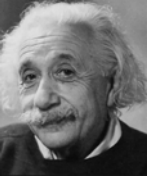
Rodin and ... the Gang



- François Auguste René **Rodin**
 - A.k.a. **Auguste Rodin**
 - A French **sculptor**
 - Generally considered the **progenitor** of **modern sculpture**
- A **2 step** procedure: ...
 - 1. **Take a rock** of marble
 - 2. **Remove** all the **unnecessary**



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 ...



The equation $E=mc^2$ is written in white chalk on a dark, textured blackboard background.

Oh, ... 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** reliable **way** to master
 - Also **Academic Writing**

Sorry, that
ABSTRACT

Not to Forget

Acknowledgements

- Paying your **credits** to those, **who ...**
 - **Fund** your work
 - Are not on the authors list, but **helped** with the reported work
 - Anyone else, whom you do not directly cite, but feel **grateful** to
 - E.g. for a valuable **advise**

Acknowledgements

The first author is funded by a PhD grant provided by Zaporizhzhia National University and the Ministry of Education and Science of Ukraine. The research leading to this report has been done in part in cooperation with the Ontology Engineering Group of the Universidad Politécnica de Madrid in frame of FP7 Marie Curie IRSES SemData project (<http://www.semdata-project.eu/>), grant agreement No PIRSES-GA-2013-612551. A substantial part of the instrumental software used in the reported experiments has been developed in cooperation with BWT Group. The data collection of Springer journal papers dealing with Knowledge Management, including DMKD, has been provided by Springer-Verlag.

Citations

- **Use** without **mentioning** the **source** is **plagiarism**
- Any **mention** of the others **result** has to be
 - **Cited**, by giving a reference
 - E.g. (Allen, 1984), or [13]
 - **Referenced** by putting the paper in the References section. e.g.:
 - Allen, J. (1984): Towards a General Theory of Action and Time. Artificial Intelligence, **23**(2), pp. 123–154

Revision History

- It is worth showing how much you **suffered** ...
- Your funding body may use it for **checking** against your work plan
- E.g.:

Revision Log:				
Version/Revision	Date	Change(s)	Submitted by	Comment
0.2, Draft	31.07.2017	Initial draft containing experimental set-up and partial results	Victoria Kosa	
0.3, Draft	15.09.2017	Intermediate draft: SOTA-methods, all experimental results, updated references	Victoria Kosa	
0.4, Draft	17.09.2017	Intermediate draft. Added: - Sect. Executive Summary - Sect 8: Conclusions and Recommendations A few language and style corrections done.	Vadim Ermolayev	
0.5, Draft	20.09.2017	Pre-final Draft. Added: - Sect 2.2 Available Software Implementations (draft) Corrected: Experimental W-flow in Fig. 4 Layouts and formats checked and corrected	Vadim Ermolayev, Victoria Kosa	
0.6, Final	22.09.2017	Final. Added: - Links to instrumental software modules	Dmitriy Naumenko, David Chaves-Fraga, Vadim Ermolayev	

Reading

Basic Reading

- Sorry, did not find a proper textbook for this topic

Additional Reading

- **Scientific method:**
 - **Dodig-Crnkovic, G.:** Scientific Methods in Computer Science. Proc Conf for the Promotion of Research in IT at New Universities and at University Colleges in Sweden (2002)
 - <http://www.mrtc.mdh.se/publications/0446.pdf>
- Several reports mentioned in the slides – as **examples**

Final Remarks

What We Learned ...

- A technical report is a **primary document** to present any **scientific output**
 - A **deliverable** in any R&D project
 - Structure and narration have to reflect the **methodology** followed in the project
 - The tips for **structuring** the sections
 - The suggestions on **illustrations**
 - The importance and tips on reproducibility
 - Advise on **collaboration** and **teamwork** in report writing
 - What needs to be **filtered out**
 - What needs to be **not forgotten**
-
- **Next topic: Part 2: Writing a Thesis**

**Will be happy
to answer
your questions ...**

Will be also happy to continue discussions

vadim@ermolayev.com

