

#### The Logistics of Scientific Writing Dr. Jim A. Field jimfield@email.arzona.edu

Assistant Dean for Graduate Education, College of Engineering

1

- If your work is not published it essentially does not exist
  - All the hard labor and effort, year after year to do your PhD research, for what?
  - In the end it is of little significance if nothing is published
- Not published = 0 impact
- <u>Video Importance of Writing</u>





Not published = pinning a note on a saguaro cactus

- Expectations for a Competitive Academic Candidate. Number of Peer Reviewed Journal Publications
  - After PhD = 3 to 4 publications (requirement = 2)
  - After first 2 y post doc = 6 to 8 publications
  - Minimum to be considered a serious candidate for assistant professor = 5 to 10

#### Publications

FPGA acceleration of seeded similarity search, Arpith Jacob, Joseph Lancaster, Jeremy Buhler a nd Roger Chamberlain. Book chapter in *Bioinformatics: high performance parallel computer architectures*, edited by Bertil Schmidt, Taylor & Francis/CRC Press. Available on Amazon.com.

Mercury BLASTP: accelerating protein sequence alignment, Arpith Jacob, Joseph Lancaster, Brandon Harris, Jeremy Buhler and Roger Chamberlain. ACM Transactions on Reconfigurable Technology and Systems, 2008.

Hardware technologies for high performance data-intensive computing, Maya Gokhale, Jon Cohen, Arpith Jacob, Craig Ulmer, Andy Yoo, Roger Pearce and Marcus Miller. *IEEE Computer (Cover Feature), 2008.* 

Biosequence similarity search on the Mercury system, Praveen Krishnamurthy, Jeremy Buhler, Roger Chamberlain, Mark Franklin, Kwame Gyang, Arpith Jacob and Joseph Lancaster. Journal of VLSI Signal Processing, 2007.

Design of throughput-optimized arrays from recurrence abstractions, Arpith Jacob, Jeremy Buhler and Roger Chamberlain. IEEE International Conference on Application-specific Systems, Architectures and Processors, 2010. Fondation Michel Métivier Best Paper Award. (Acceptance rate: 27.0%)

Rapid RNA folding: analysis and acceleration of the Zuker recurrence, Arpith Jacob, Jeremy Buhler and Roger Chamberlain. IEEE Symposium on Field-Programmable Custom Computing Machines, 2010. Best Paper Award. (Acceptance rate: 18.2%)

**Optimal runtime reconfiguration strategies for systolic arrays**, Arpith Jacob, Jeremy Buhler and Roger Chamberlain. *IEEE International Conference on Field-Programmable Logic and Applications*, 2009. (Acceptance rate: 25.0%)

Accelerating Nussinov RNA secondary structure prediction with systolic arrays on FPGAs, Arpith Jacob, Jeremy Buhler and Roger Chamberlain. IEEE International Conference on Application-specific Systems, Architectures and Processors, 2008. (Acceptance rate: 31.8%)

Language classification using n-grams accelerated by FPGA-based bloom filters, Arpith Jacob and Maya Gokhale. Supercomputing (Workshop on High-Performance Reconfigurable Computing Technology and Applications), 2007.

FPGA-accelerated seed generation in Mercury BLASTP, Arpith Jacob, Joseph Lancaster, Jeremy Buhler and Roger Chamberlain. IEEE Symposium on Field-Programmable Custom Computing Machines, 2007. (Acceptance rate: 27.4%)

#### Invited Talks

Automatic parallelization of dynamic programming recurrences in computational biology Custom Computing Research Group, Imperial College London, England. July 2010. IRISA, Campus de Beaulieu, Rennes, France. September 2009.

FPGA acceleration of seeded similarity search IRISA, Campus de Beaulieu, Rennes, France. September 2009.

#### "Publications are the most important beans in academic bean counting"



• Role of Writing in the Life-Cycle of the Scientific Endeavour



#### • Research without writing = tip of iceberg



# **3 THINGS THAT YOU NEED TO GET READY**



- 1. What is your narrative?
  - Does the story line make sense?
  - Should be a simple message (1 or 2 objectives with 1 or 2 simple take home messages)
  - Complex messages from multiple objectives should be avoided
- 2. Identify the data you want to include in the article
  - Before writing make the Figures and Tables for the article
- 3. Complete Literature Research
  - Must do extensive literature search before writing
  - Graduate Students should read an article a day



# NARRATIVE BLOCKBUSTER MOVIE



• *Hidden Figures* tells the incredible untold story of three brilliant African-American women working at NASA who served as the brains behind the launch into orbit of astronaut John Glenn, a stunning achievement that turned around the Space Race. The visionary trio crossed all gender and racial lines and inspired generations.

## **EXAMPLE: 1 SENTENCE NARRATIVE MANUSCRIPT**



• A microbial culture accelerates the corrosion of zero valent iron (ZVI) forming biogenic reactive secondary minerals that greatly increase the uranium remediation rate in groundwater by more than 20-fold with a mechanism involving reductive precipitation U<sup>VI</sup> to U<sup>IV</sup>.

# OUTLINE

#### • Framework to follow while writing article

- 1) Introduction
  - a) Background
    - i) Compound A is an important breakthrough that can improve environmental performance
    - ii) Indications that Compound A can react with catalyst x
    - iii) Hypothesis
  - b) Objectives
    - i) Determine reactions products of compound A by catalyst x
    - ii) Elucidate mechanisms accounting for observed products

#### 2) Material and Methods

- a) Materials
  - i) Source chemicals
  - ii) Synthesis catalyst x
- b) Experimental
  - i) Reaction vessels
  - ii) Addition of chemicals
  - iii) controls
- c) Analytical
  - i) Measurement of products in GC
  - ii) Measurement of products in ICP-MS
  - iii) Mass spectrometry

#### MASTERPIECE: LAS MENINAS BY VELÁZQUEZ



10

# HOW DOES A MASTER PAINTER PAINT A MASTERPIECE?

- Starts in one corner completes everything in detail and moves on to next segment of the canvass?
  - -grid approach

![](_page_10_Figure_3.jpeg)

- Makes a crude sketch over entire canvass, then goes back and each time layers on more detail over the whole canvass?
  - -outline approach

![](_page_10_Figure_6.jpeg)

## PICASSO'S MASTERPIECE, GUERNICA

![](_page_11_Picture_1.jpeg)

#### EVOLUTION OF GUERNICA

![](_page_12_Picture_1.jpeg)

# BRAINSTORMING AN OUTLINE

## • What are all ideas about article

• Write them down (don't worry about logic or order)

### • Organize ideas

- Start thinking up schemes for organizing ideas
- Keep improving organization

#### • Make an outline at all levels

- First the whole manuscript
- Than a section in the manuscript
- Than a paragraph in the section

![](_page_13_Picture_10.jpeg)

![](_page_13_Figure_11.jpeg)

• Title Page (title, authors, affiliation, contact info)

• Abstract (summarize, stand alone, no new info)

- 1-2 sentences background
- 1 sentences objectives
- 1-2 sentences approach/methodologies
- 2-6 sentences results (provide quantitative info)
- 1-2 sentences discussion/implication & conclusion

• Introduction (background, context, objectives)

- Background and description general problem
- Description specific research question (hypotheses)
- Objectives

#### o Material and Methods

- Materials Used
- Experimental Set Up
- Methodologies
- Analytical Protocols

• **Results** (describe data in Figure and Tables)

- Walk readers through the data
  - Call each Table/Figure
  - Explain logic or specific goal of the experiment
  - Describe to readers the important trends
  - what is main message from the Figure or Table
  - repeat again on next Figure or Table
- Write like you talk when showing a data slide during a seminar
- Not necessary to repeat data in Figures/Tables in text
- Leave big picture interpretation for discussion

# EXAMPLE OUTLINE OF RESULTS

#### 3) Results

- a) Time course of reaction (Figure 1)
  - i) Compound A rapidly decreases
  - ii) Not different from control
- b) Time Course Product formation (Figure 2)
  - i) C and D were formed based on GC evidence
  - ii) Structures of C and D confirmed Mass Spectrometry
- c) etc. etc. etc.

- **Discussion** (interpret results in broader context and answer hypothesis and objectives statements)
  - <u>Main findings:</u>
    - Short overview of the key findings obtained from results
  - <u>Compare and Contrast:</u> How does previous work support the findings observed
    - Data from literature that agrees/disagrees with findings and why
  - <u>Mechanisms:</u> Intellectual Analyses of Information (your results and literature data)
  - Support Hypothesis / Modify Hypothesis
    - Based on all information and its analysis is the hypothesis correct? if not, how does it need to be modified or studied further
  - <u>Implications</u> for discipline and/or industry

- [Conclusions](some journals require 2-4 bullets as concluding remarks)
- Acknowledgements (funding source, analyst)

#### • References

- List of References cited according to format of journal
- Tables
- Figure Captions

#### • Figures

# Quality Graphs and Tables: Two Considerations

#### • Design for Data Visualization

• How the graph or table conveys information most efficiently to the reader

#### Format and Drafting Quality Graph and Tables

- Following guidelines (elements of figures and tables)
- Graph drafting details (font size, marker size, line thickness, scaling etc)

![](_page_20_Picture_6.jpeg)

![](_page_20_Picture_7.jpeg)

![](_page_20_Picture_8.jpeg)

# GUIDELINES FOR VISUALIZING DATA

#### • Edward R. Tufte (Professor Yale)

- pioneer in the field of data visualization, authored 3 important books
  - $\rightarrow~$  The Visual Display of Quantitative Information
  - $\rightarrow$  Envisioning Information
  - $\rightarrow$  Visual Explanations

#### o Important Ideas of Tufte

- Scientific illustrations should convey the data more efficiently than text
- Strive for highest ratio data to ink
- Graphics simplify comparisons between many variables, encouraging deep thinking about the data (not just observation)
- Graphics must enlighten not mislead

![](_page_21_Picture_11.jpeg)

![](_page_21_Picture_12.jpeg)

![](_page_21_Picture_13.jpeg)

# GUIDELINES FOR VISUALIZING DATA • Historical Examples

![](_page_22_Picture_1.jpeg)

## Charles Joseph Minard- Napoleon's Campaigns

23

# GUIDELINES FOR VISUALIZING DATA

#### • Historical Examples

![](_page_23_Picture_2.jpeg)

Pump sites Deaths from cholera OXFORD STREE meters

John Snow, Cholera Outbreak London and hypothesis source disease is from drinking water

#### In Figures, captions are at bottom (legends should be in caption)

![](_page_24_Figure_1.jpeg)

**Figure 1.** Formation of carbonyl groups in the protein bovine serum albumin (BSA) after reaction with nanoparticles (NP). Legend: reaction with BSA and NP, ( $\bullet$ ); BSA only, (o); NP only ( $\blacktriangle$ ); maximum oxidation of BSA (-----).

Bad graphs will look like this in publication:

![](_page_24_Figure_4.jpeg)

25

#### Example of Improved Graph

![](_page_25_Figure_1.jpeg)

**Figure 1.** Oxidation of protein by nanoparticles. Formation of side chain carbonyl groups as indicators of the oxidation of bovine serum albumin (BSA) after reaction with  $Mn_2O_3$  nanoparticles (NP). Legend: reaction with BSA and NP, ( $\bullet$ ); BSA only, (o); NP only ( $\blacktriangle$ ); maximum oxidation of BSA (-----).

Good graphs will look like this in publication:

![](_page_25_Figure_4.jpeg)

26

# IDEAL GENERIC PARAGRAPH

- Introductory Sentence (sets the stage for topic)
- The "Claim". Statement of Hypothesis or Identify depice a Trend or Statement of Position
- Supporting Evidence(s) of Claim.
  - Evidence 1
  - Evidence 2
  - Evidence  $n_f$
- Conclusion of Claim or Summarize Trend
- Transition Sentence to Introduce Next Topic

# EXAMPLE GENERIC PARAGRAPH, RESULTS SECTION

![](_page_27_Figure_1.jpeg)

**Figure 1.** Time course of the reduction of uranium in the presence of *Eerbeek* sludge. Legend: --- $\diamond$ ---, Not inoculated; --- $\diamond$ ---, Heat killed sludge with H<sub>2</sub>; —  $\blacktriangle$  —, Live sludge (no electron donor added); — $\bullet$ —, Live sludge with H<sub>2</sub>.

# EXAMPLE GENERIC PARAGRAPH, RESULTS SECTION

The ability of anaerobic sludge to catalyze the reductive Introduction precipitation of uranium was investigated in batch experiments. An example of a typical experiment is shown in Figure 1, Claim demonstrating that microorganisms in anaerobic sludge reduce <u>uranium</u>. The uranium concentration did not change in noninoculated and heat-killed inoculum controls. Rapid decreases in the soluble uranium concentration were only observed in Evidence treatments with active sludge. The rate of reduction was stimulated approximately two-fold when H<sub>2</sub> was added as an electron donor. Nonetheless it should be noted that a substantial rate was also observed in treatments lacking  $H_2$ . The results indicate that uranium removal is a reaction catalyzed by intrinsic Conclusion microorganisms which utilize both endogenous substrates as well as exogenously supplied  $H_2$  to provide electrons to support the reduction process. The ability of different anaerobic sludge samples to catalyze uranium reduction was evaluated in several Transition additional experiments. 29

# CONCLUSIONS 1

#### • Prepare Before Writing

- Select and Prepare Optimized Figures /Tables
- Literature Research
- Create a Narrative (story-line) and a Descriptive Title

#### • Outline

- Brainstorm ideas
- Organize them into a framework and logical narrative
- Keep layering on more detail until you have an entry for each sentence

#### • Main Sections

- Title page, Abstract
- Introduction
- Material & methods
- Results
- Discussion
- Tables
- Figures

# CONCLUSIONS 2

## o Ideal Paragraph

- Introduction
- Claim (Topic Sentence)
- Supporting Evidence(s), 1, 2, ...  $n_f$
- Conclusion
- Transition to next paragraph