M1 MSc DataScience - Methods and tools for technical and scientific writing

Bibliographical research

Semestre 1, 2021-2021

Lecturer: Marco Winckler (UCA, I3S)

winckler@unice.fr

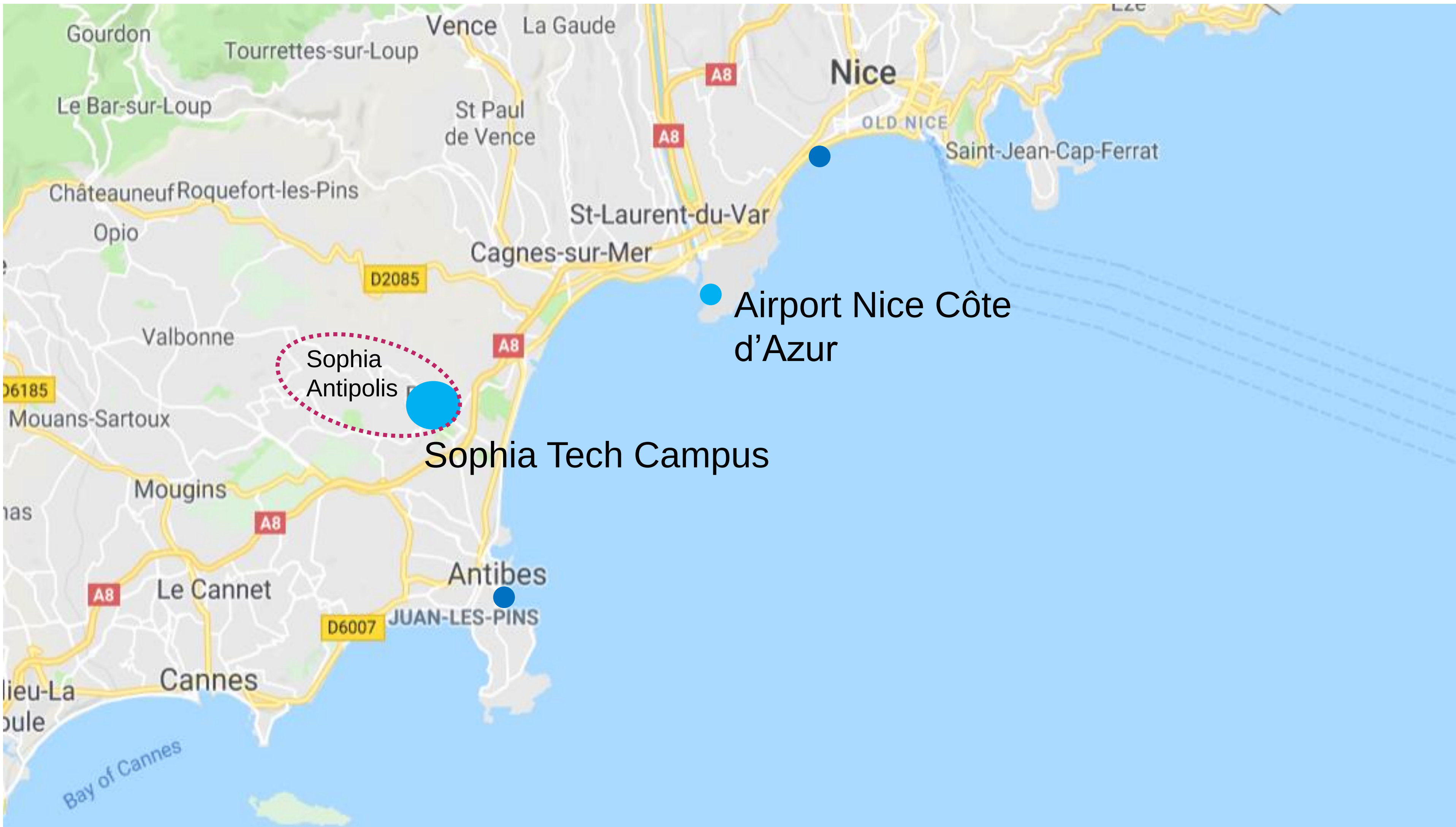
UNIVERSITÉ CÔTE D'AZUR

About me…

- Born in Nonoai, RS, Brazil
- Research domain:
  - Human-Computer Interaction (HCI) Interaction Home-Machine (IHM)
  - Interactive Systems Engineering
  - Web Engineering

- Before…
  - From 2009-2017 responsible internship program of Master 2 IHM (Université Paul Sabatier/ENAC)

About the Université Nice Sophia Antipolis



Why Conduct Research?

- To develop knowledge for professions.
- To develop effective policies.
- To solve practical problems.
- To make informed decisions.
- To increase the knowledge base of larger society.

- Huge amounts of daily life and experience in our society are based on what we have learned using the logic and evidence involved in scientific research.

BRETAM model of the development of science technology

- Gaines, B.: Modeling and forecasting the information sciences. Inf Sci 57/58: (1999) 13-22

Modeling and Forecasting the Information Sciences

Brian R. Gaines
Knowledge Science Institute
University of Calgary
Alberta, Canada T2N 1N4

Abstract

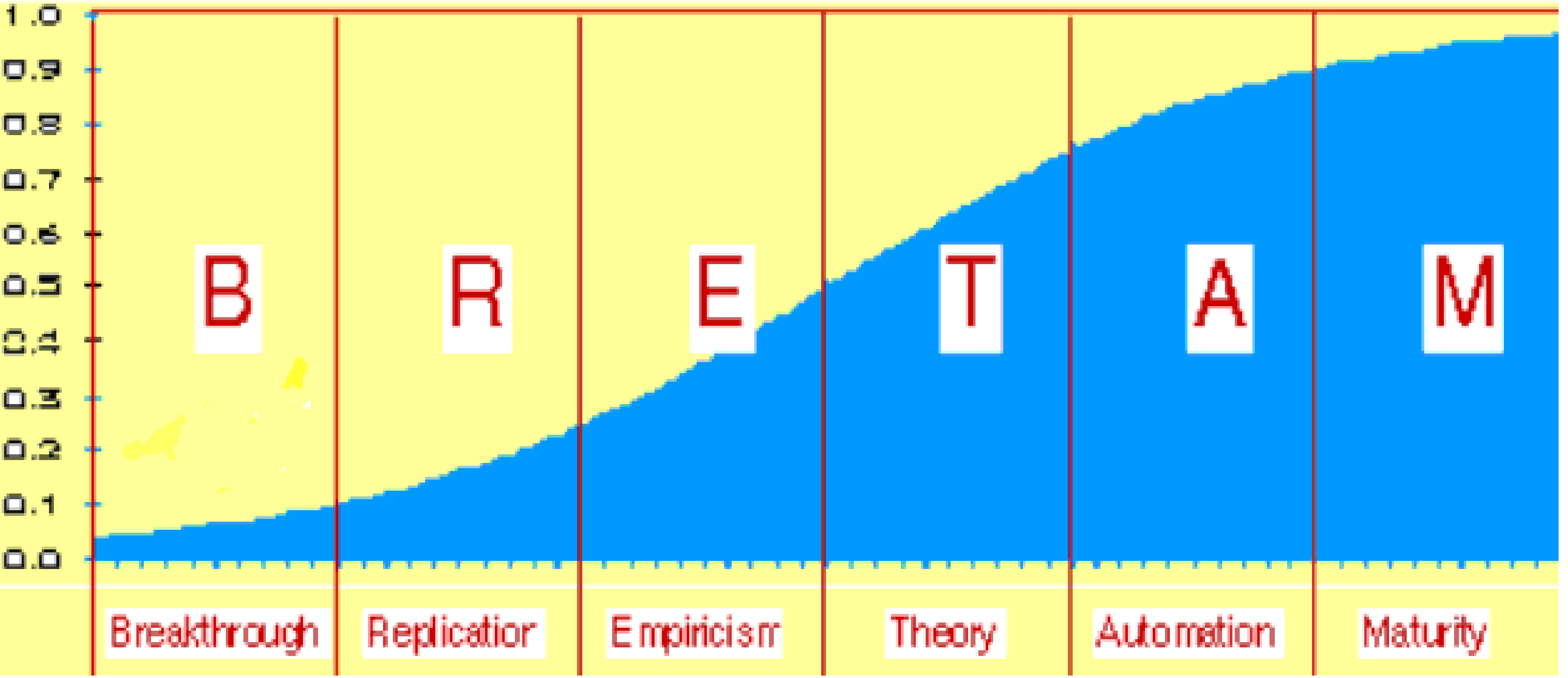
A model of the development of the information sciences is described and used to account for past events and predict future trends, particularly fifth and sixth generation priorities. The information sciences came into prominence as electronic device technology enabled the social need to cope with an increasingly complex world to be satisfied. Underlying all developments in computing is a tiered succession of learning curves which make up the infrastructure of the computing industry. The paper provides a framework for the information sciences based on this logical progression of developments. It links this empirically to key events in the development of computing. It links it theoretically to a model of economic, social, scientific and individual development as related learning processes with a simple phenomenological model. The fifth generation development program with its emphasis on human-computer interaction and artificial intelligence, and the sixth generation research program with its emphasis on knowledge science are natural developments in the foci of attention indicated by the model.

1 Introduction

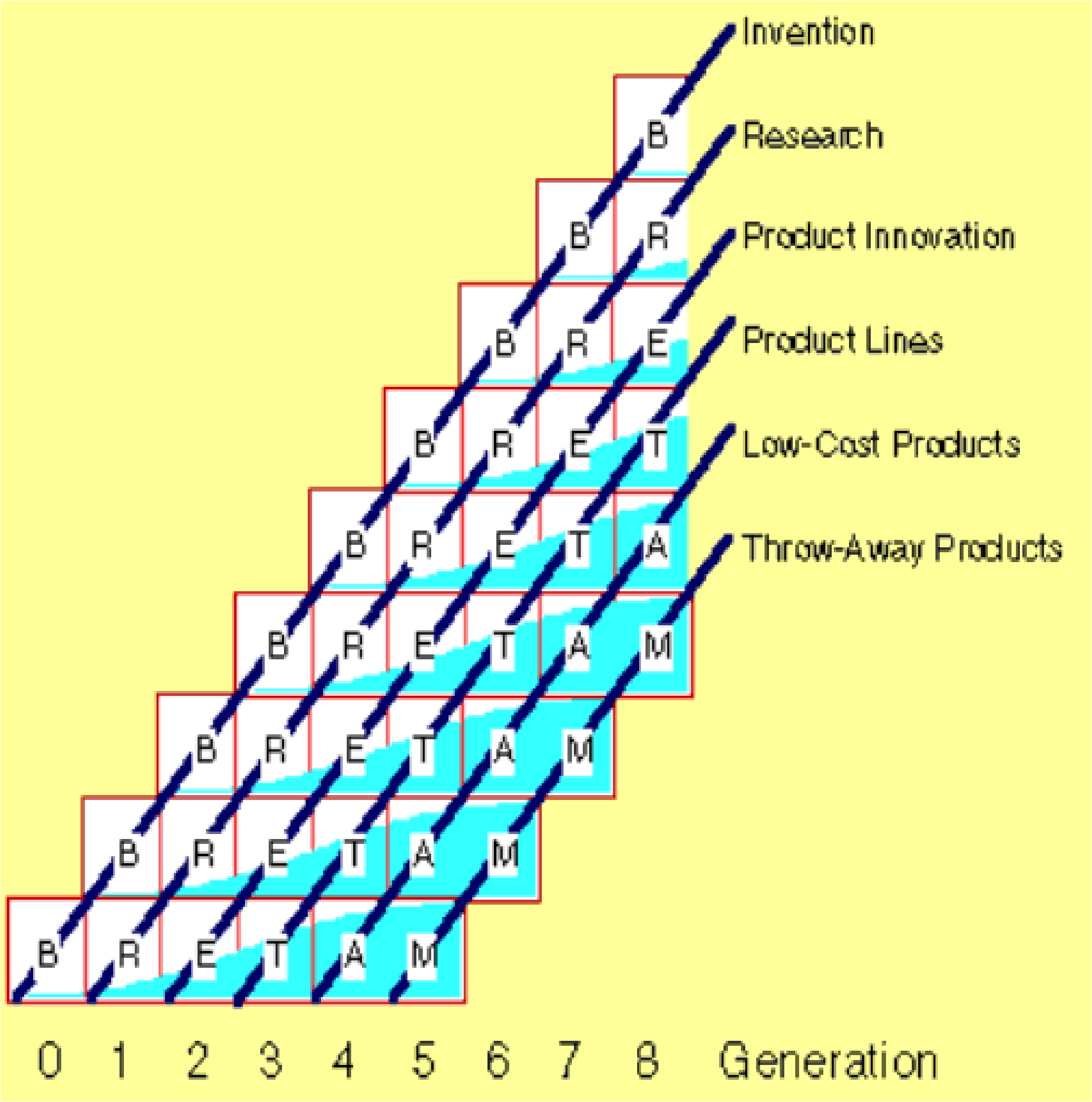
Forecasting advances in technology and their impact has a track record of making fools of the forecasters (Schnaars, 1989). However, the game of life is one of anticipating the future. We model the past that we may learn the lessons of experience, and we extrapolate our models into the future. That the future reserves the right not be anticipated is a meta-lesson that Hume taught many years ago. Nevertheless, our individual behaviors and our civilizations are founded on the assumption that anticipation is possible and, at some level of modeling, history does repeat itself. Our technological civilizations go one step further and change the universe to reify our anticipations. The information sciences, in particular, structure a wholly artificial reality composed from the ultimate abstractions of the human mind. Paradoxically, they should be readily modeled because they are artifacts of our own mentation, but they may also be beyond modeling because to do so fully may involve ultimate understanding of ourselves.

This paper presents an integrative model of the information sciences that shows them as a tightly coupled, mutually supportive system. It first presents the underlying electronic device technology which provides both the basic support and physical constraints on information technology. It then presents a model of the learning curves of scientific and technological knowledge acquisition that underlies the development of the information sciences. The historic opportunities triggering successive advances in the information sciences are then analyzed and their learning curves superimposed to provide a model of the infrastructure of the information sciences. This model is extrapolated to provide forecasts of future directions, and fitted to various scientific and technological developments. The interactive synergies between levels are analyzed to show the basis of the positive feedback phenomena which continue to support the exponential growth of the information sciences and technologies.

Evolution of domain



The process ladder



Keep updated

Sources of information about

- Attending conferences, symposiums, workshops, …
- Digital libraries
  - ACM Digital Library: http://dl.acm.org/
  - Springer Link: https://link.springer.com/
  - IEEExplorer: https://ieeexplore.ieee.org/
  - HAL Archives Ouvertes (France): https://hal.archives-ouvertes.fr
- Collaborative Platforms
  - Interaction Design Fondation: https://www.interaction-design.org/
  - Research Gate https://www.researchgate.net/
  - Mendely: https://www.mendeley.com/
- Associations
  - Association for Computing Machinery (ACM)
  - International Federation for Information Society (IFIP)
  - Institute of Electrical and Electronics Engineers (IEEE)

How to read a scientific paper ?

(usual) Anatomy of a scientific paper

- Title and authors
- Abstract/summary
- Introduction
- Materials and Methods
- Results
- Discussion
- Acknowledgements
- References
- Figures/Tables

Title and authors

- Title is very descriptive (often states the main finding) and is not about being creative and “catchy”!
- Order of authors is important. What can you tell from it?

Example:

VEGF, a prosurvival factor, acts in concert with TGF-beta1 to induce endothelial cell apoptosis.
Ferrari G, Pintucci G, Seghezzi G, Hyman K, Galloway AC, Mignatti P.

Abstract/Summary

- Brief background of subject
- Purpose for the study
- Major findings of the study
- Relationship between these findings and the field

This is what you see when you do a pubmed search. You can decide if the paper is worth reading based upon this.

Introduction

- Presents the background information for a fellow scientist (possibly in another field) to understand why the findings of this paper are significant.
- Structure is usually:
  - Accepted state of knowledge in the field
  - Focus on a particular aspect of the field, often the set(s) of data that led directly to the work of this paper
  - Hypothesis being tested
  - Conclusions (scientists don’t really like surprise endings!)

How to approach the introduction

- **Grab a blank piece of paper:**
  - Take notes
  - Draw mini figures
  - Define vocabulary
    (wikipedia is a quick reference)

- **Answer these questions:**
  - What is the accepted state of knowledge?
  - What data led directly to the work of this paper?
  - What is the hypothesis being tested?
  - What are the basic conclusions? (Scientists don't really like surprise endings and this is usually stated in the last paragraph.)

Notes allow you to take a break (hours to days) and come back to your thoughts…you won't have to re-read the parts you completed.

Materials and Methods

- Should be detailed enough for another scientist to replicate the work (volumes, times, company material was purchased from etc.)
- In reality, often compressed and you may need to look up another paper that is referenced for more detail.

Should you read the materials and methods?

- Often you can skim over them before the results.
- However, when you get to the results, you will need to flip back to them often to clarify how experiment was done.
  - Sample number? (Did they do this more than once?)
  - Conditions? (Am I looking at a reduced or non-reduced protein gel?)

Results

- While the introduction poses the questions being asked, the results describes the outcome of the experiments that were done to answer the questions.
- Results are often simply stated with *interpretation* of them coming later in the discussion.
- Figures and tables allow the reader to see the outcomes of the experiments for themselves!

How to read the results:

- Read the text straight through, but as a figure is referred to, examine the figure.
- Take notes, giving yourself a place to refer to about each figure.
- With each experiment/figure you should be able to explain :

1) the basic procedure
2) the question it sought to answer
3) the results
4) the conclusion
5) criticisms

Discussion

- Data is analyzed to show what the authors believe the data show. (You don’t have to agree with their interpretations!)
- Findings are related to other findings in the field (contribute to knowledge, correct errors, etc.)– How is this work significant?

How to read a discussion

- Take notes and answer these questions:
  - What conclusions do the authors draw? Be sure to separate fact from their opinion/interpretation?
  - Describe for yourself why these data significant. (Does it contribute to knowledge or correct errors?)

By now, you may be tired of this paper…
but don’t relax yet.

Save energy for the overall reflection and criticism.

Acknowledgements

- Thank people who contributed materials.
- Thank people who contributed technically but maybe not intellectually (would not be authors).

References

- Papers cited in the text
- What parts of the paper cite other papers?
  - Introduction
  - Materials and Methods
  - Discussion
  - (Maybe a few in Results)

Question: How should we READ a scientific paper?

Answer: not necessarily in order!

A four-step method based on: **Ann McNeal, School of Natural Science, Hampshire College, Amherst MA**

http://hampshire.edu/~apmNS/design/RESOURCES/HOW_READ.html

Step 1: Skim the entire paper

- Look at the major headings (do they follow the “anatomy” we just described?)
- How many figures are there, what kinds of figures are they (gels, graphs, microscopic images?)
- What is the conclusion of the paper?
  – (It may not make sense to you at the moment, but note what it is.)

Step 2:Vocabulary

- Go through the paper as a whole simply underlining words and phrases you do not understand.
- You are not reading the paper for comprehension of the whole paper yet, just making sure you have understanding of the words to then comprehend it.

Step 2: Vocabulary continued

- Look up simple words and phrases, where?
  - Biology textbooks
  - Online at biology dictionaries or encyclopedias (www.wikepedia.com seems to be a good resource for basic definitions and procedures)
  - Look up methods that you are not familiar with

  (i.e. what is an immunoprecipitation or a transformation?)

- Note important phrases that are part of a major concept and are bigger than just vocabulary (i.e. “risk reduction”). You will come back to them in context while reading for comprehension.

Step 3: Read for comprehension, section by section

- Introduction
  - What is the accepted state of knowledge in the field (take notes and even draw your own figures)?
  - What data led directly to this work?
  - What question are they answering? (Is there a clear hypothesis?)
  - What are their conclusions?

Step 3: Read for comprehension, section by section

- Materials and Methods and Results:
  - Read the methods first or read them as you read the results. (I prefer the latter)
  - With each experiment/figure you should be able to explain
    - 1) the basic procedure
    - 2) the question it sought to answer
    - 3) the results
    - 4) the conclusion

- You should be able to explain all of these (1-4) to another classmate clearly!

Step 3: Read for comprehension, section by section

- Discussion
  - What conclusions do the authors draw? Be sure to separate fact from their opinion/interpretation?
  - Describe for yourself why these data significant. (Does it contribute to knowledge or correct errors?)

Step 4: Reflection and criticism

- Do you agree with the authors' rationale for setting up the experiments as they did?
- Did they perform the experiments appropriately? (Repeated a number of times, used correct control groups, used appropriate measurements etc)
- Were there enough experiments to support the one major finding they are claiming?
- Do you see patterns/trends in their data that are problems that were not mentioned?
- Do you agree with the authors' conclusions from these data? Are they over-generalized or too grand? Or are there other factors that they neglect that could have accounted for their data?
- What further questions do you have? What might you suggest they do next?

Tips for success:

- Spend a lot of time on each paper NOW look up every detail that you are unsure of. (Time you invest now will payoff in the long run). Discovering the answers for yourself is one of the best ways to learn and have the information be retained.
- Imagine yourself teaching the paper or figures to classmates—teaching something to others is also another great way to learn.

Tips for success:

- Start a database of procedures that you take the time to look up and teach to yourself. What are some of the common procedures that are used in various papers? (e.g. western, immunoblots, RT-PCR, apoptosis assays, yeast two hybrids, etc.)
- Watch others in your lab experiences and find out what they are doing…you may never get the opportunity to do RT-PCR, but the more you understand the procedure, the more critical you can be of data you need to interpret.

Tips for success:

- Read papers when you are awake and interested in reading. If you are going to break up a paper and read it over several days be sure to summarize before continuing each day.
- If you are already in the field you plan to stay in, consider starting a database on papers that relate to your lab/project. You will want to be able to impress your P.I. with your quick analysis and summary of a monumental paper from another lab!

Exercise

- Take a quick look of a scientific paper (sample provided)
- Take notes and answer the following questions:
  - Identify the main sections and sub-section, how can you identify them ?
  - How figures and tables are presented and cited?
  - How references to other works are cited ?
  - How can you contact the authors ?
  - When and where the paper was published?
  - What is the subject of the paper ?
  - What are the main claims ?
  - In your opinion, are the results valid ?
  - Which kind of evidence to the claims are provided?