Strategies for Excellence in Academic Research and Publication

Kin-Choong Yow, *PhD, PEng, SMIEEE, MACM*

Professor, Faculty of Engineering and Applied Science

May 15th, 2024

University of Regina

Contents

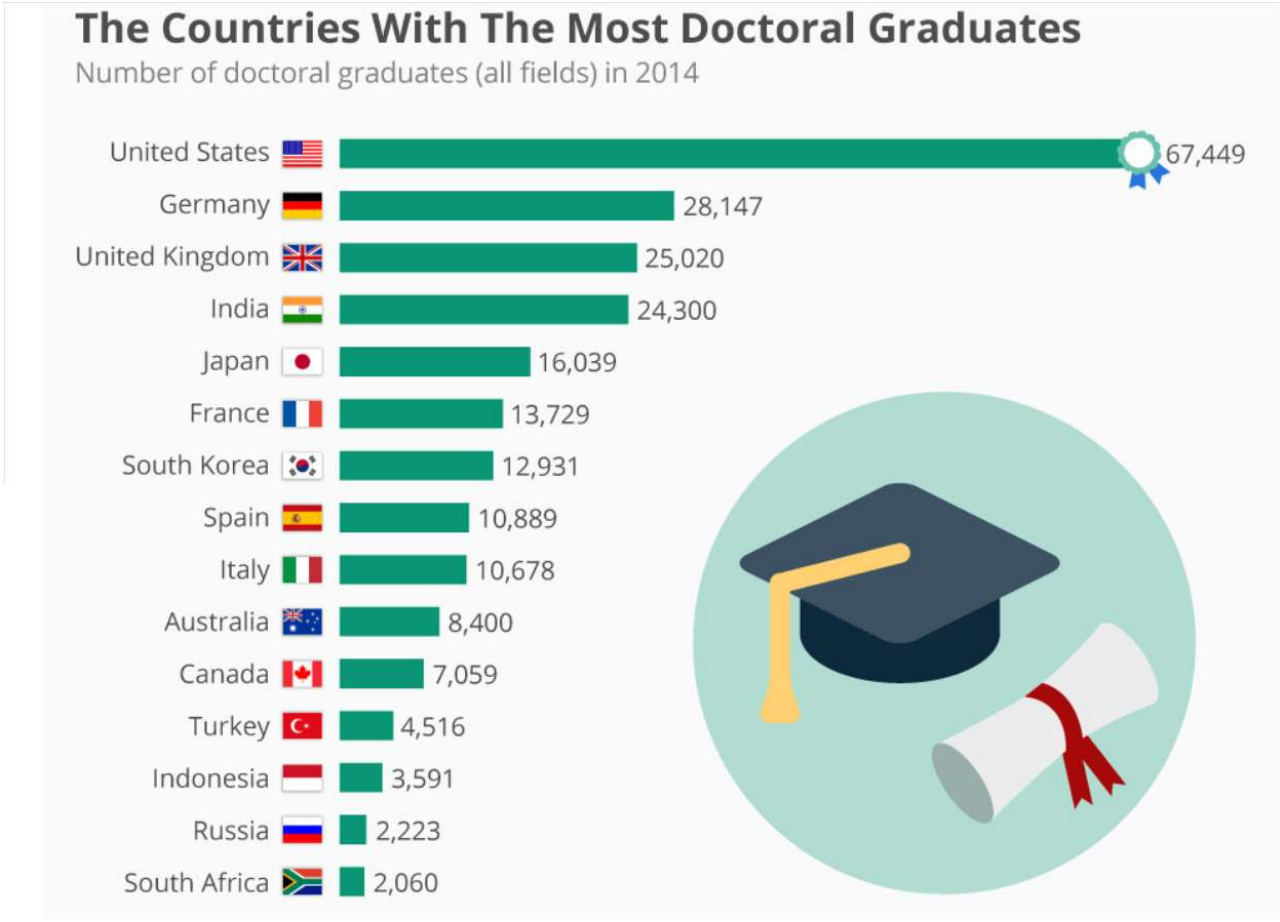
- Introduction
- How to conduct Impactful Research?
- Essentials in Academic Writing
- Tips on making Scholarly Presentations
- Notes on Ethical Academic Practices
- Conclusions

Introduction

Introduction

- Education is **key** to economic growth.
- And **tertiary education**, in particular, is at the heart of the innovation that we see around us. New discoveries such as 5G communication technologies and ChatGPT would never have happened were it not for PhD research.
- Countries are **investing** in their higher education systems, and more people than ever before are completing doctoral degrees.

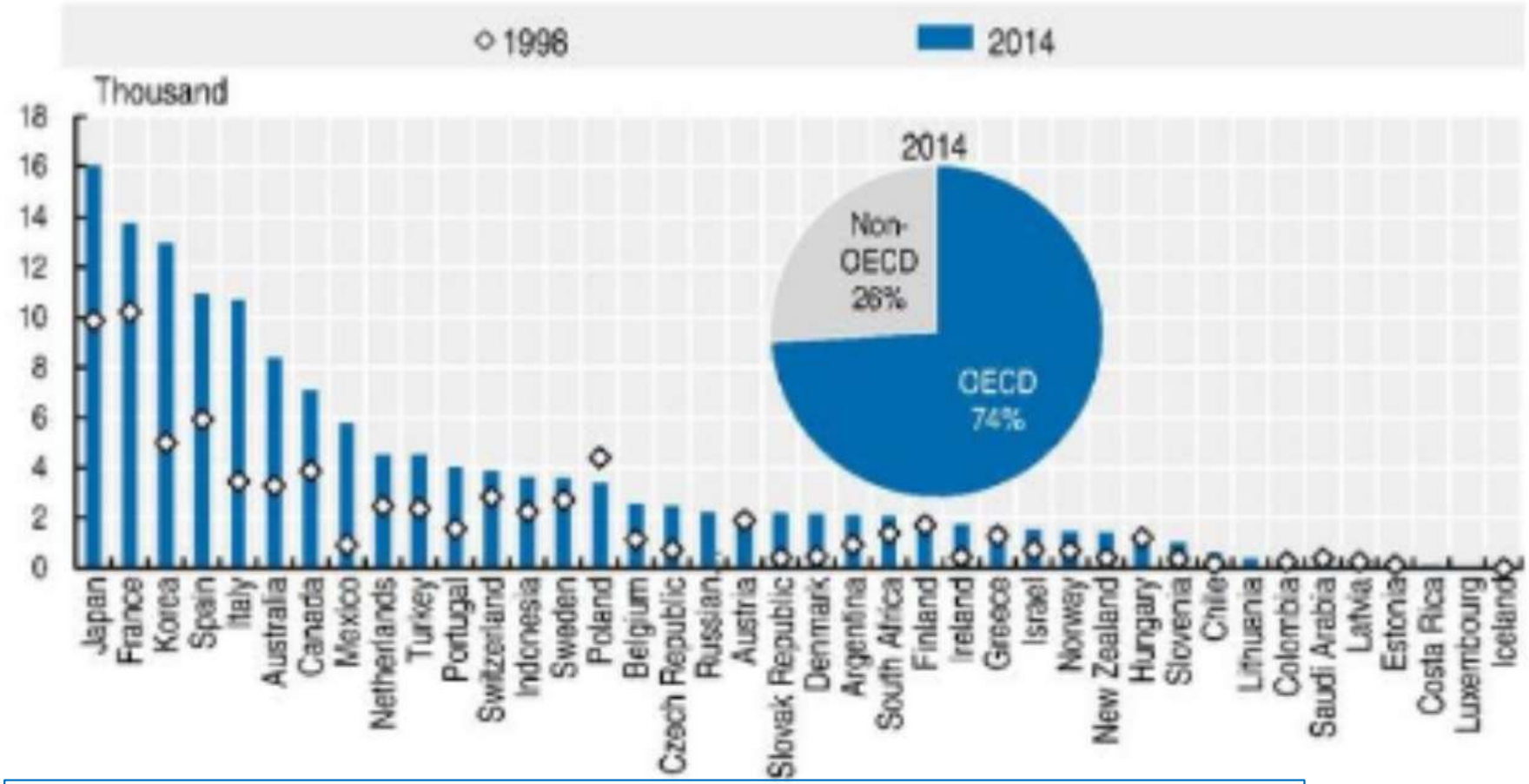
No. of PhD Graduates (2014)



Source: World Economic Forum
https://www.weforum.org/agenda/2017/02/countries-with-most-doctoral-graduates/

University of Regina

Increase in PhD Graduates



Source: World Economic Forum
https://www.weforum.org/agenda/2017/02/countries-with-most-doctoral-graduates/

University of Regina

No. of Scientific Pub. in 2023 (All Subject areas)

	Country	↓ Documents	Citable documents	Citations	Self-Citations	Citations per Document	H index
1	China	1043131	1018423	1094503	768786	1.05	1333
2	United States	714412	609674	654637	270853	0.92	3051
3	India	306647	269183	252299	113231	0.82	858
4	United Kingdom	238568	201255	272435	66276	1.14	1928
5	Germany	202397	179861	202876	56510	1.00	1690
6	Italy	155258	137096	170158	54453	1.10	1333
7	Japan	134358	124330	102180	26579	0.76	1301
8	Canada	128502	113461	137877	27662	1.07	1562
9	Spain	122876	111563	125846	29830	1.02	1215
10	France	122302	110009	121557	26658	0.99	1514

Source: SJR (Scimago Journal and Country Rank)
https://www.scimagojr.com/countryrank.php?year=2023

University of Regina

So how can we make a difference?

- Conduct impactful research
- Write high quality papers
- Make effective presentations

How to conduct Impactful Research?

University of Regina

Five steps of Impactful Research

1. Problem Identification
2. Literature Survey
3. Propose a Solution
4. Experiments and Evaluation
5. Future Work

1. Problem Identification (i.e. What is the problem?)

Example

- ChatGPT-3.5 was released on November 30, 2022 and hit 1 million users in less than one week on December 4, 2022. Today it has about 180.5 million users.
- **So what are the problems?**
  - Hallucination (39% hallucination rate)
  - Limited input sequence length (2k tokens, i.e. ~1,500 words)
  - Does not take image input
  - Does not take voice input
  - Cannot interrupt the output

1. Problem Identification (i.e. What is the problem?)

Possible Solutions

- One method to reduce hallucination involves creating a **deliberately flawed** model trained on data that contradicts the prompt or contain information it lacks. We then subtract the weights of the flawed model from those of the original to create a model which hallucinates less.
- If we increase the input length, it will suffer from the **Serial Position Effect** (i.e. the model remembers the first and the last item the best). To overcome this, we can employ **Retrieval Augmented Generation** (i.e. using embedding search and vector databases to find only relevant context).

1. Problem Identification (i.e. What is the problem?)

GPT-4

- However, GPT-4 (released March 14, 2023) has overcome several of the problems
  - Hallucination rate reduced to 3%
  - Input sequence length increased to 8k tokens, with an additional extended model of 32k tokens (~25,000 words)
  - Able to take images as input
- **So what are the remaining problems?**
  - Does not take voice input
  - Cannot interrupt the output

1. Problem Identification (i.e. What is the problem?)

GPT-4o

- Fortunately (or unfortunately), a new version GPT-4o (released May 13, 2024) can now do the following
  - Able to take voice as input
  - Able to interrupt the output
- See a live demo of GPT-4o https://www.youtube.com/watch?v=RI-BxtCx32s

- **So what are the remaining problems?**

2. Literature Survey (i.e. What have other people done about the problem?)

- Also known as Related Work
- You need to search digital repositories for leading journals and top conferences. Just doing a Google search is not enough
- Journals with Impact Factor (IF) of 10 are outstanding, 3 are good, and less than 1 are not so good
- Journals with Scimago Journal Rank (SJR) score > 1.0 are good, < 1.0 are not so good

2. Literature Survey (i.e. What have other people done about the problem?)

- Your search must include recent articles (within the past 3 years)
- Avoid citing arXiv. arXiv contains preprints (i.e. not peer-reviewed). If it is a good technique, it is likely to be accepted and published in a journal or a conference. Cite those instead.
- Avoid citing Wikipedia. They are also not peer reviewed, and anyone can edit them.

2. Literature Survey (i.e. What have other people done about the problem?)

- As you read the papers, you must also find out **why their techniques are not good enough** (i.e. if their techniques solves the problem, then there is no more problem).
- If you cannot find any papers related to your problem, it can only be one of two reasons:
  1. You did not search widely enough
  2. Nobody is interested in this problem

3. Propose a Solution (i.e. how would you solve the problem?)

- This is probably the most difficult part of the research
- You need to think of a new technique that solves the problem that you have identified
- **Simple rule** – if the technique is simple, other people probably have already tried it.

4. Experiments and Evaluation (i.e. Show that your method is indeed a solution)

- You need to perform extensive evaluation to show that your technique is indeed a solution
- Typically, you need to
  - evaluate your technique on some standard metric, e.g. SNR, RMSE, F1-score, etc.
  - evaluate how your result change with different model parameters
  - comparison with other state-of-the-art techniques
  - analysis of Type I and Type II errors
  - ablation studies

5. Future Work (i.e. What would you do if you have more time?)

- After you have evaluated your technique and shown that it is a solution, you need to identify shortcomings in your technique
- It is unlikely that your solution is so complete that there is nothing more to do.
- You probably have made some assumptions in your technique. So your future work can address situations where those assumptions fail.
- You should give tangible proposed solutions with some actual (but brief) details

Summary

The five steps of Impactful Research

1. **Problem Identification**
   (i.e. what is the problem?)
2. **Literature Survey**
   (i.e. what have other people done about the problem?)
3. **Propose a Solution**
   (i.e. how would you solve the problem?)
4. **Experiments and Evaluation**
   (i.e. show that your method is a solution)
5. **Future Work**
   (i.e. what would you do if you have more time?)

Essentials in Academic Writing

University of Regina

Essentials in Academic Writing

- With hundreds of thousands of paper published per year, your paper has to stand out from the rest in order to be noticed.
- Reviewers and Researchers are looking for two things
  - Novelty
  - Impact
- We will now discuss how you can write up your research so that it will get noticed. We will use the following paper from IEEE T-NNLS (IF: 14.255) as an example:
- Abnormal Event Detection and Localization via Adversarial Event Prediction https://ieeexplore.ieee.org/abstract/document/9346050

Essentials in Academic Writing

1. Abstract

Your abstract is a summary of your **whole paper**. It must include a brief description of each section of your paper, i.e.

- the problem,
- why existing techniques cannot solve the problem,
- the key idea of your proposed technique,
- significant results

Your readers will read your abstract and **decide** whether it is worth spending more time to actually read your paper

Essentials in Academic Writing

2. Introduction

Apart from the usual problem definition and motivation, your introduction should include a section (or paragraph) on **Contributions**. In this section you want to highlight what are the key contributions of your paper (preferably in point form).

Essentials in Academic Writing

3. Related Work

Apart from the usual description of related works of other researchers, it is important to highlight why their techniques are **not good enough**. This can be at the end of your Related Work section so that it naturally leads into your proposed technique.

Essentials in Academic Writing

4. Proposed Method

You should write your proposed method in enough details so that it is **repeatable** by other researchers. It is recommended that you give an overall architecture in the beginning of the section, and then explain in detail each of the components in subsequent sections. Try to use as many **figures** as possible. A picture is worth a thousand words.

Essentials in Academic Writing

5. Results and Evaluation

You need to **explain** what the reader should see in each of your result table. Do not just say “Table xxx shows the result of the experiments” and leave the reader to figure out what does the table says. You should also **highlight** (usually in bold) the highest value (or best value) in each column so that the reader can easily see that your technique is better than other state-of-the-art techniques.

Tips on making Scholarly Presentations

University of Regina

Tips on making Scholarly Presentations

- When making a presentation about your research (usually at a conference, or during your thesis defense), you need to remember that:
  - Your audience are your peers
  - Some of them are experts in your field
  - Some of them are experts in a related field but may not know very much about your field
- Therefore, the common thing that they want to know is the **idea** behind your technique, and **how well** it works

Tips on making Scholarly Presentations

- Usually, at a conference presentation, you are limited to about 15 or 20 minutes. You have to keep to that time otherwise the Session Chair may cut you off.
- Your 15 minutes should be divided as follows:
  - 5 minutes – Introduction, Problem Definition, Related Work
  - 5 minutes – The proposed technique
  - 5 minutes – Results and Conclusion, Future Work
- Remember, the audience wants to know **what is your work**, not other people's work. So keep the problem definition and related work brief (not more than 5 mins) and jump straight to your proposed technique.

Tips on making Scholarly Presentations

- Use lots of **figures** and **tables**. Avoid using long sentences
- If long sentences (> 3 lines) cannot be avoided, **highlight** the keywords in those sentences
- Do not reduce your font size. (Microsoft Powerpoint comes with a default 32 pt font. Do not go below 28 pt font)
- **Rule of Thumb** – It takes about one minute to talk about each slide.

Tips on making Scholarly Presentations

- Animations are fine, but do not do it excessively

Example: My keynote address on Enhancing Artificial Intelligence with Self Attention Mechanisms at the 5th International Conference on Artificial Intelligence, Robotics, and Control (AIRC 2024), April 22-24, 2024, Cairo, Egypt.

Notes on Ethical Academic Practices

University of Regina

Famous cases of Academic Misconduct

- **Diederik Alexander Stapel** (born 19 October 1966) is a Dutch former professor of social psychology at Tilburg University.
- In 2011 Tilburg University suspended Stapel for **fabricating** and **manipulating** data for his research publications.
- This scientific misconduct took place over a number of years and affected dozens of his publications. By 2015, **fifty-eight** of Stapel's publications had been retracted.
- He has been described in coverage by the *New York Times* as "the biggest con man in academic science".

Source: Wikipedia
https://en.wikipedia.org/wiki/Diederik_Stapel

Famous cases of Academic Misconduct

- **Ranga P. Dias**, a physicist at the University of Rochester, was in 2024 found by an investigatory committee to have committed "research misconduct" related to his work on alleged superconducting materials.
- A 2023 report in *Science* noted that at least 21% of Dias's 2013 doctoral thesis had been copied from uncredited sources. As of 2024, Dias has had **four** of his research papers retracted, and four other papers have received an expression of concern.

Source: Wikipedia
https://en.wikipedia.org/wiki/List_of_scientific_misconduct_incidents

Plagiarism Checker

- In all present day journal and conference academic paper review, there is a plagiarism checker that will compute the similarity index

  Example: this paper submitted to CCECE 2024

- IEEE uses CrossCheck which compares submitted manuscripts against a very large database of published technical papers (as well as over 6 billion web pages). IEEE recommends a **30% threshold**. Therefore, any manuscript with 30% or more similarity to previously published content will be flagged for further review

Is using ChatGPT plagiarism?

- No, ChatGPT doesn't plagiarize in the sense that it doesn't copy information and pass it along to you.
- But—and it's a but you can't overlook—because it learns from existing sources to write the information it gives you, it may borrow other writers' ideas without giving them credit, which is a form of plagiarism.
- It may also generate content that closely resembles existing content, which can be mistaken for plagiarism.

Source: Quillbot
ChatGPT and Plagiarism: Academic Authenticity - QuillBot Blog

University of Regina

ChatGPT detectors?

- There are several tools that can be used to detect if a work is created by ChatGPT or not.
  - ZeroGPT (https://www.zerogpt.com/)
  - GPTZero (https://gptzero.me/)
  - AI Text Classifier (https://platform.openai.com/ai-text-classifier)
- These tools detect whether the text is likely to be AI generated together with a confidence value.
- Values can be from 10% to 99% for ChatGPT generated outputs.

Conclusions

Conclusions

- We have discussed the five steps of impactful research
- We have talked about the essentials of writing an academic paper, and some tips on making scholarly presentations.
- We have discussed some points about academic integrity and whether using GenAI tools count as plagiarism.

University of Regina