MONASH University

MONASH BUSINESS SCHOOL

Writing for Researchers

Rob J Hyndman

ACEMS Mentoring Seminar

Outline

1. References
2. Languages and editors
3. Reproducibility & version control
4. Writing a paper or thesis
5. Getting published
6. Reviewer reports
7. Writing grant applications
8. Writing industry and media reports

Managing references

Mendeley

- Free and on all operating systems
- Web-version and local version synced
- Browser extension for adding papers/books
- Attach notes and annotations to papers.
- Works with Word, LibreOffice or LaTeX.
- Generate bibliography automatically
- Handles all formatting for you.

To install:

- Set up account at www.mendeley.com
- Download from www.mendeley.com

Managing references

Zotero

- Free and on all operating systems
- Web-version and local version synced
- Browser extension for adding papers/books
- Attach notes and annotations to papers.
- Works with Word, LibreOffice or LaTeX.
- Generate bibliography automatically
- Handles all formatting for you.

zotero

To install:

- Set up account at www.zotero.org
- Download from www.zotero.org

Managing references

Paperpile

- $3 per month and runs on Google Chrome
- Papers stored on Google Drive
- Browser extension for adding papers/books
- Works with Google Docs or LaTeX.
- Generate bibliography automatically
- Handles all formatting for you.
- Amazingly fast

To install:

- Set up account at paperpile.com
- Download Google chrome browser extension

What to cite?

- Cite what is important.
- Cite (only) what is relevant.
- Avoid lists of gratuitous references.
- Include proper citations for all packages and software you use.

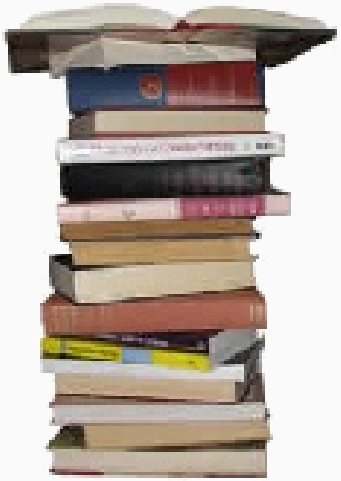
What to cite?

- Cite what is important.
- Cite (only) what is relevant.
- Avoid lists of gratuitous references.
- Include proper citations for all packages and software you use.

```
citation("packagename")
```

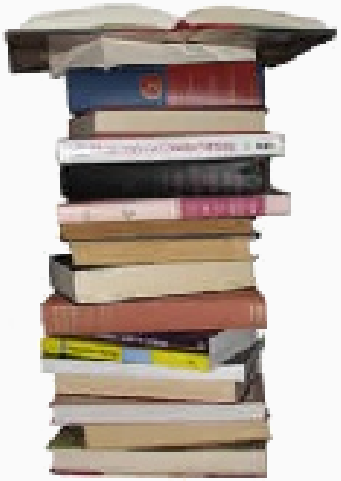
Sight what you cite

- Every article cited should be sighted, & preferably read.
- At the very least, check that the article cited really does say what you think it says.
- Type the reference information yourself.
- Don’t just cite what other people say about citations.
- Store accurate reference info from the start.
- Give credit where it is due.

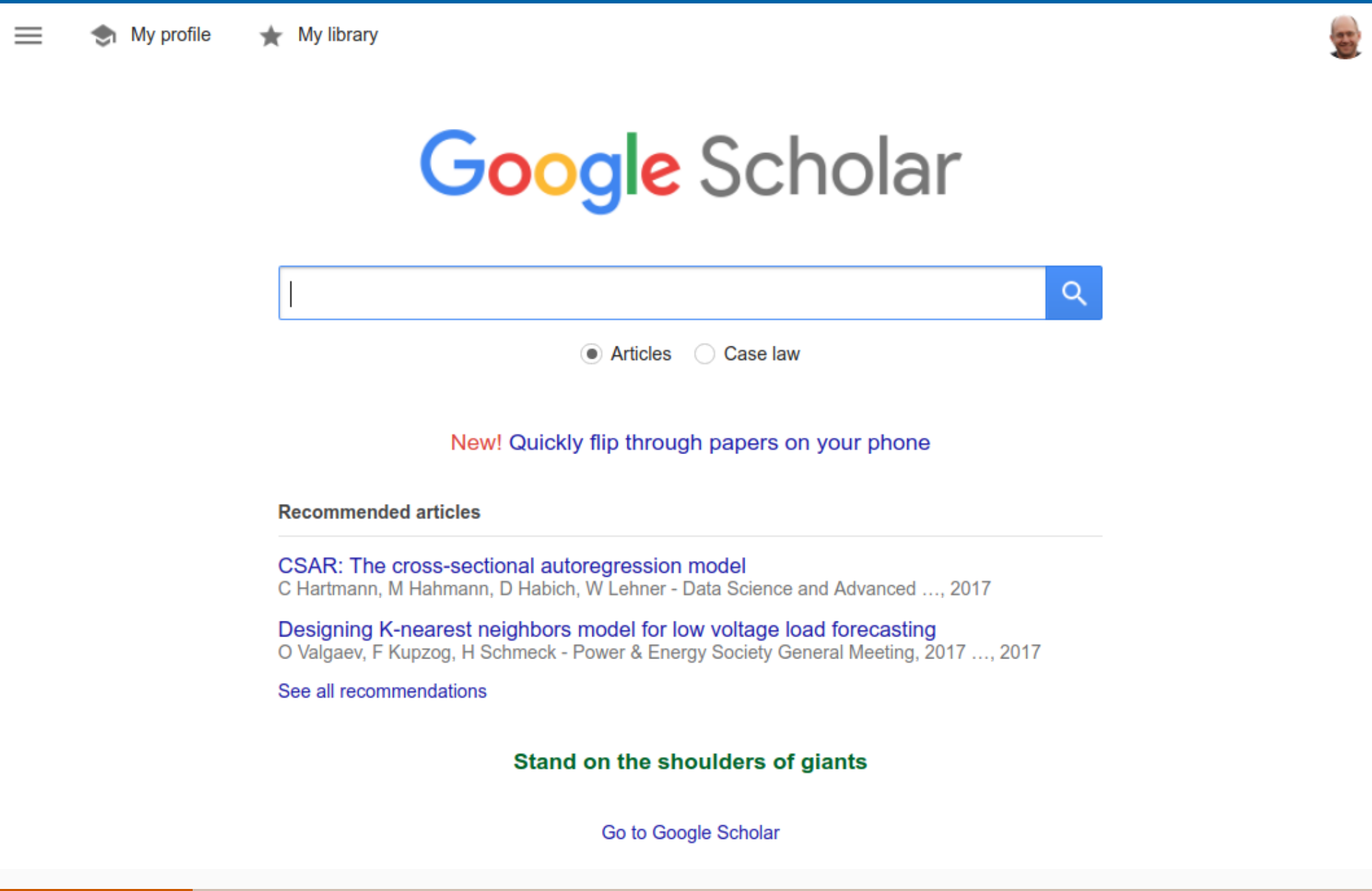


Sight what you cite

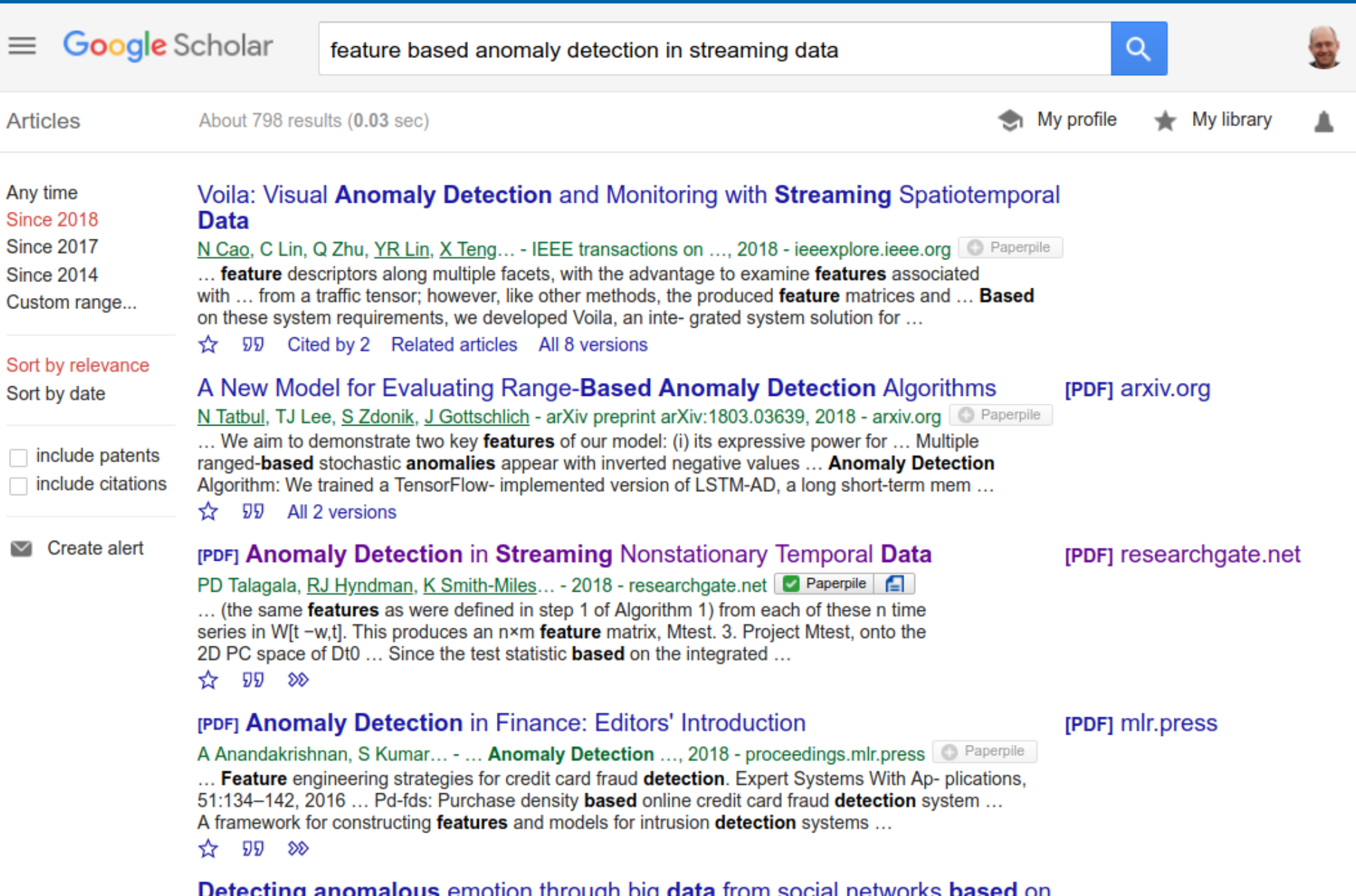
- Every article cited should be sighted, & preferably read.
- At the very least, check that the article cited really does say what you think it says.
- Type the reference information yourself.
- Don’t just cite what other people say about citations.
- Store accurate reference info from the start.
- Give credit where it is due.
  - Diebold did not invent PITs.
  - Hyndman did not invent exponential smoothing or ARIMA models.



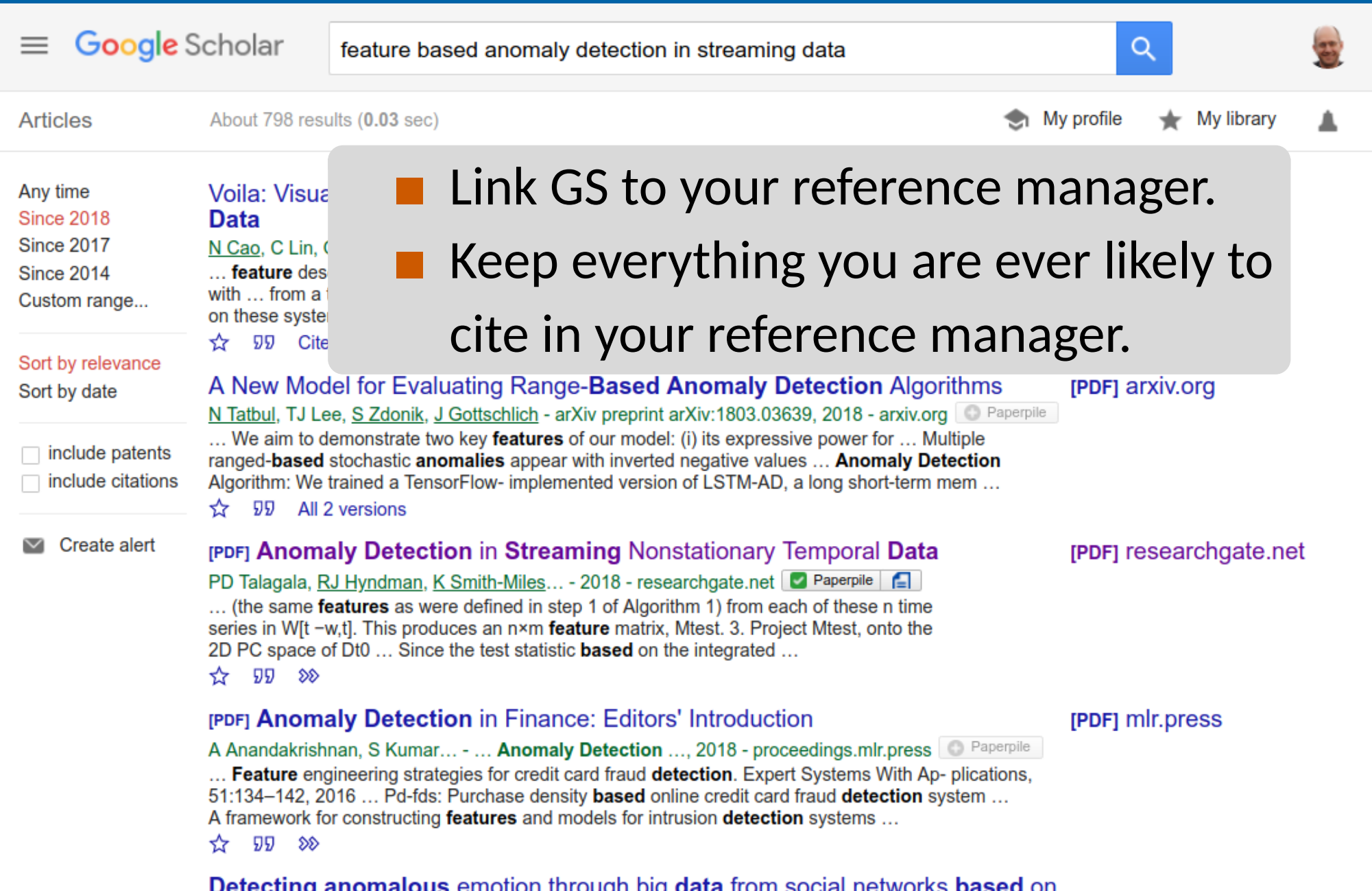
Use Google Scholar for more than searching



Use Google Scholar for more than searching

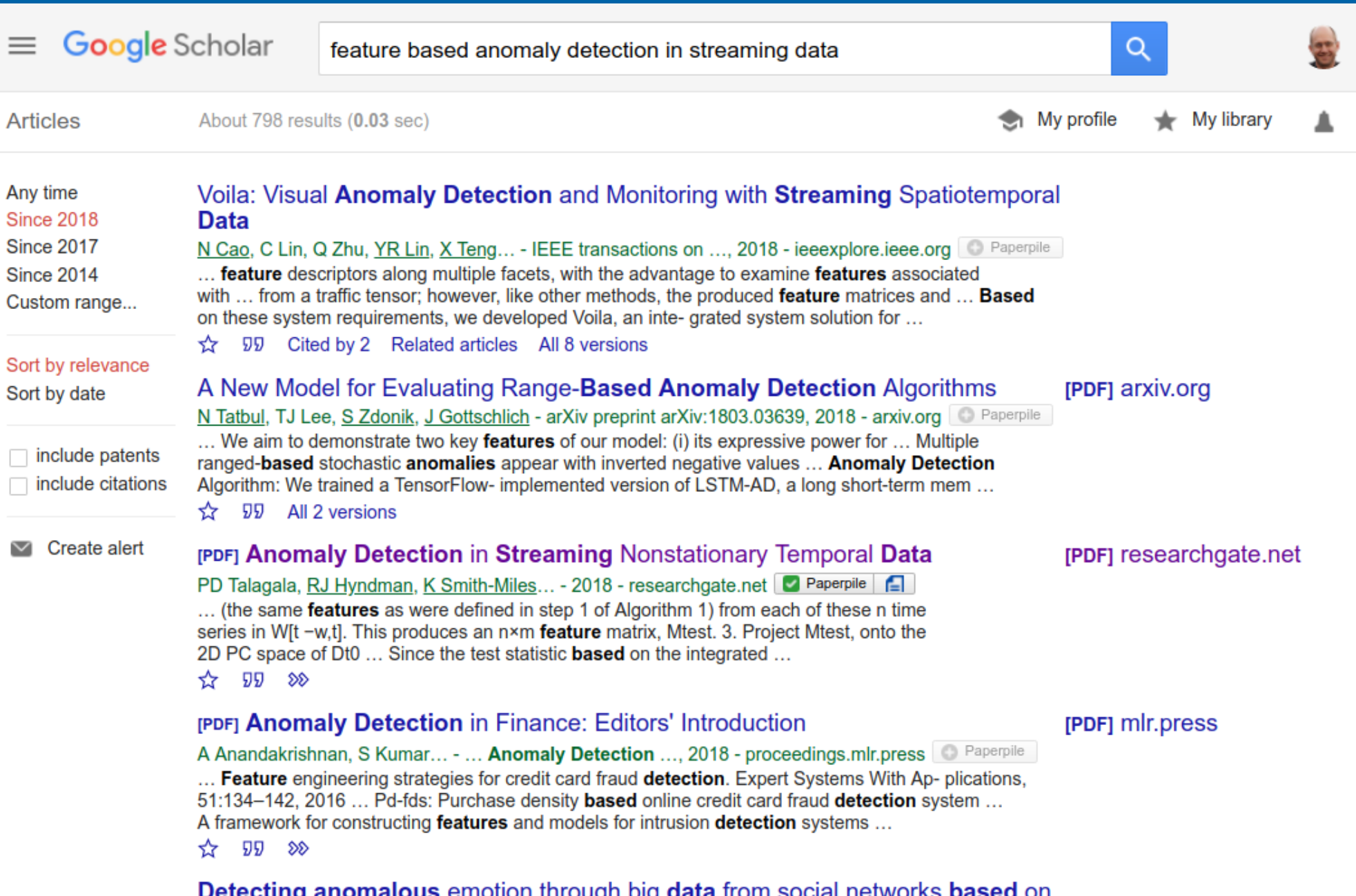


Use Google Scholar for more than searching

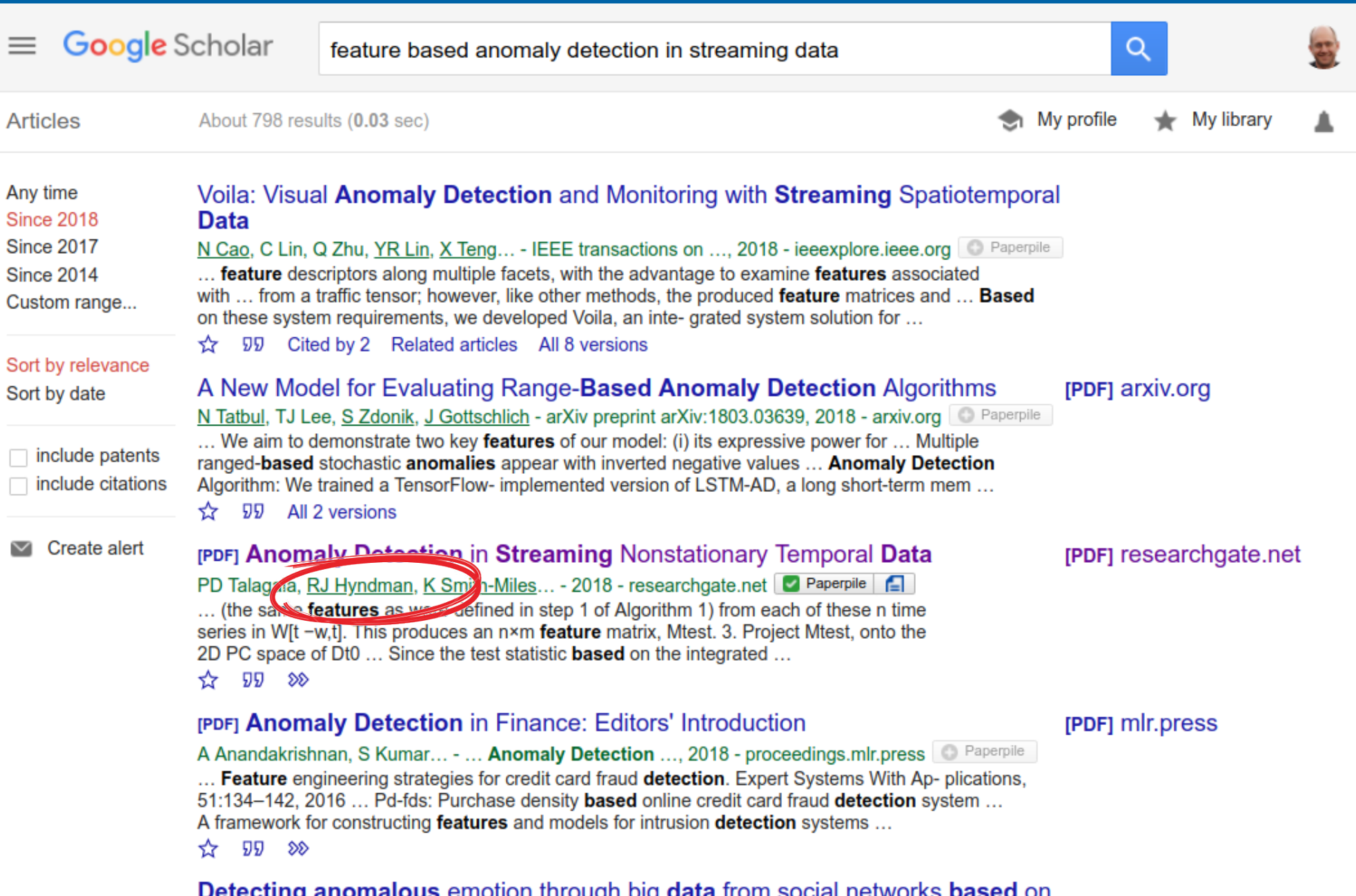


- Link GS to your reference manager.
- Keep everything you are ever likely to cite in your reference manager.

Use Google Scholar for more than searching



Use Google Scholar for more than searching



Use Google Scholar for more than searching

Google Scholar

Rob J Hyndman

FOLLOW

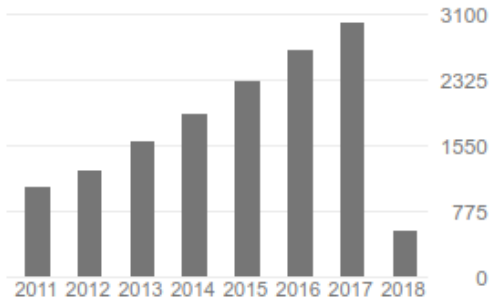
Professor of Statistics, Monash University
Verified email at monash.edu - Homepage

Forecasting Time series Applied Statistics Machine learning Exploratory data analysis

TITLE	CITED BY	YEAR
Forecasting methods and applications S Makridakis, SC Wheelwright, RJ Hyndman John Wiley & Sons	4798 *	1998
Another look at measures of forecast accuracy RJ Hyndman, AB Koehler International journal of forecasting 22 (4), 679-688	1815	2006
Automatic time series for forecasting: the forecast package for R RJ Hyndman, Y Khandakar Journal of Statistical Software	1149	2007
Forecasting: principles and practice RJ Hyndman, G Athanasopoulos OTexts	776	2014
25 years of time series forecasting JG De Gooijer, RJ Hyndman International journal of forecasting 22 (3), 443-473	682	2006
Forecasting with exponential smoothing: the state space approach RJ Hyndman, AB Koehler, JK Ord, RD Snyder Springer Verlag	675	2008
Sample quantiles in statistical packages RJ Hyndman, Y Fan The American Statistician 50 (4), 361-365	636	1996
Detecting trend and seasonal changes in satellite image time series	581	2010

Cited by VIEW ALL

	All	Since 2013
Citations	19636	12091
h-index	54	43
i10-index	113	91



Co-authors EDIT

- Ralph Snyder — Associate Professor, Department…
- George Athanasopoulos — Monash University
- Han Lin Shang — Research School of Finance, Act…
- Keith Ord — Professor of Business Statistics, …
- Bircan Erbas — La Trobe University

Use Google Scholar for more than searching

Google Scholar

Rob J Hyndman

FOLLOW

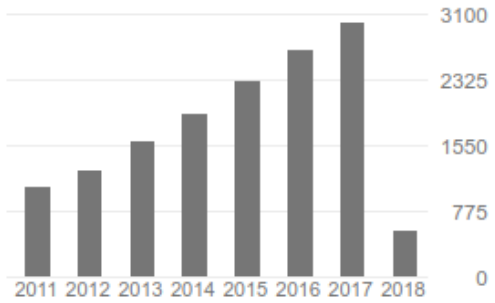
Professor of Statistics, Monash University
Verified email at monash.edu - Homepage

Forecasting Time series Applied Statistics Machine learning Exploratory data analysis

TITLE	CITED BY	YEAR
Forecasting methods and applications S Makridakis, SC Wheelwright, RJ Hyndman John Wiley & Sons	4798 *	1998
Another look at measures of forecast accuracy RJ Hyndman, AB Koehler International journal of forecasting 22 (4), 679-688	1815	2006
Automatic time series for forecasting: the forecast package for R RJ Hyndman, Y Khandakar Journal of Statistical Software	1149	2007
Forecasting: principles and practi RJ Hyndman, G Athanasopoulos OTexts		
25 years of time series forecastin JG De Gooijer, RJ Hyndman International journal of forecasting 22 (3),		
Forecasting with exponential smo RJ Hyndman, AB Koehler, JK Ord, RD S Springer Verlag		
Sample quantiles in statistical packages RJ Hyndman, Y Fan The American Statistician 50 (4), 361-365	636	1996
Detecting trend and seasonal changes in satellite image time series	581	2010

Cited by VIEW ALL

	All	Since 2013
Citations	19636	12091
h-index	54	43
i10-index	113	91

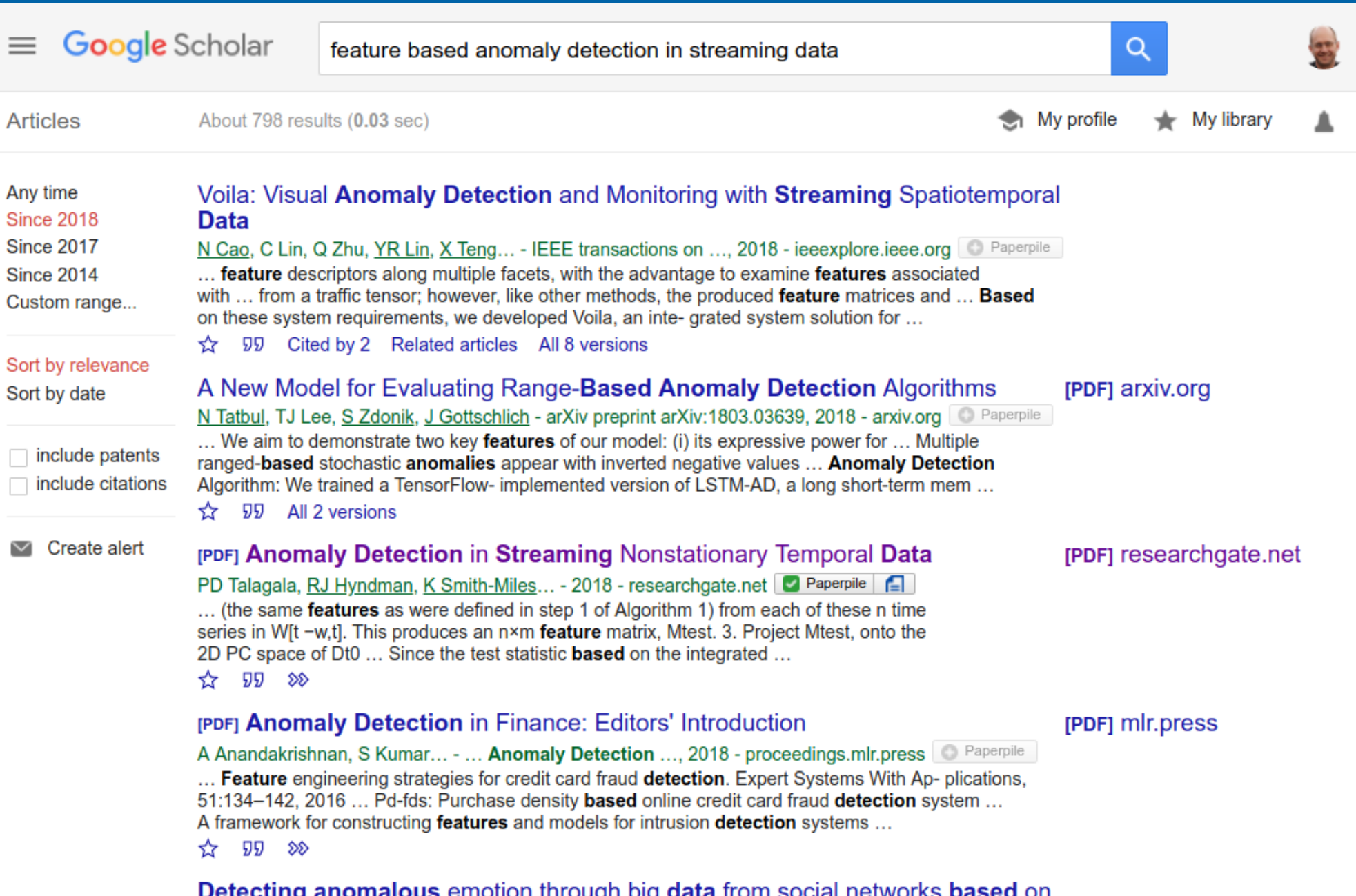


Co-authors EDIT

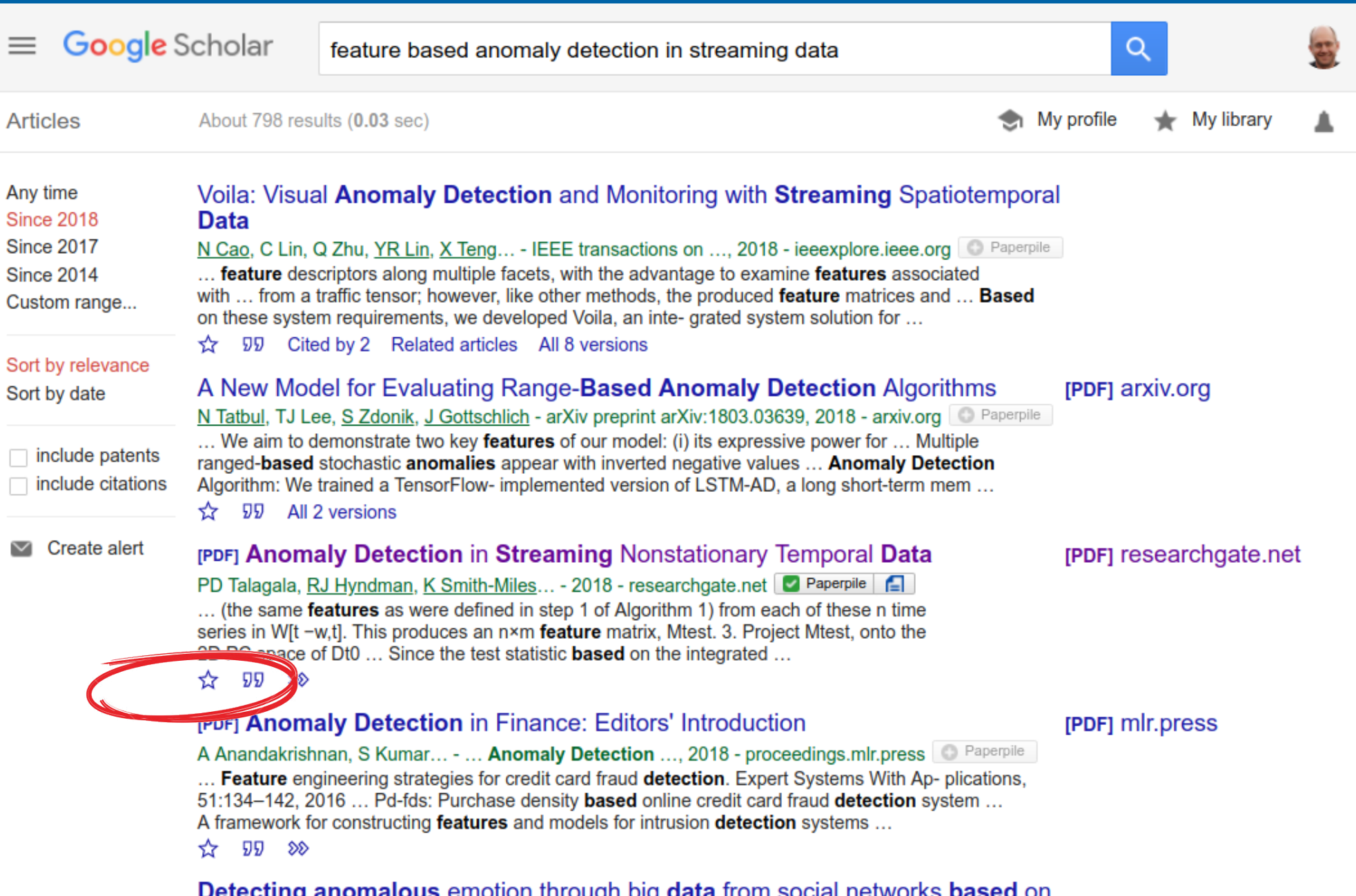
Bircan Erbas
La Trobe University

- Create your GS profile once you have (at least) one paper.
- Follow key authors in your area.

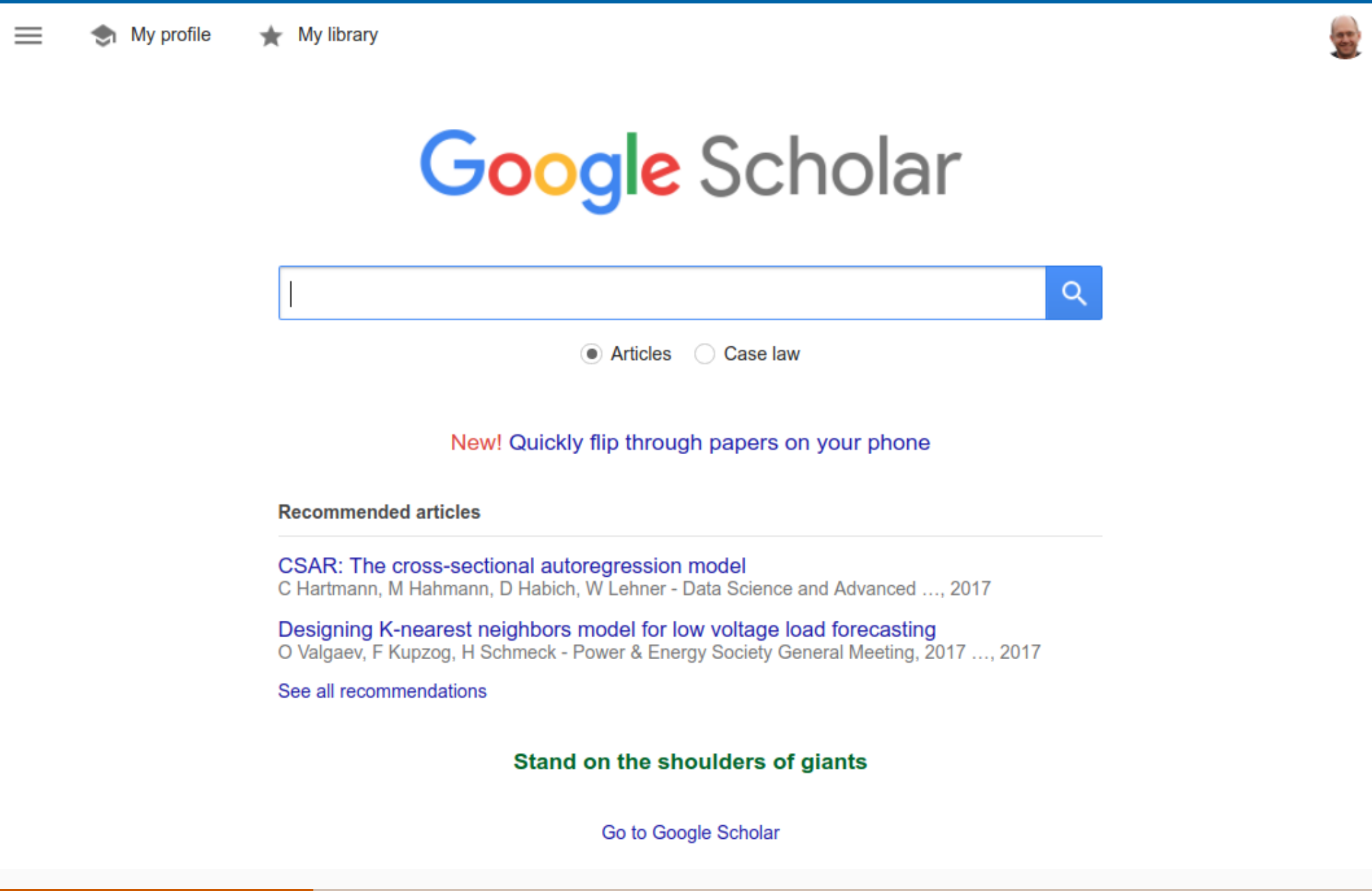
Use Google Scholar for more than searching



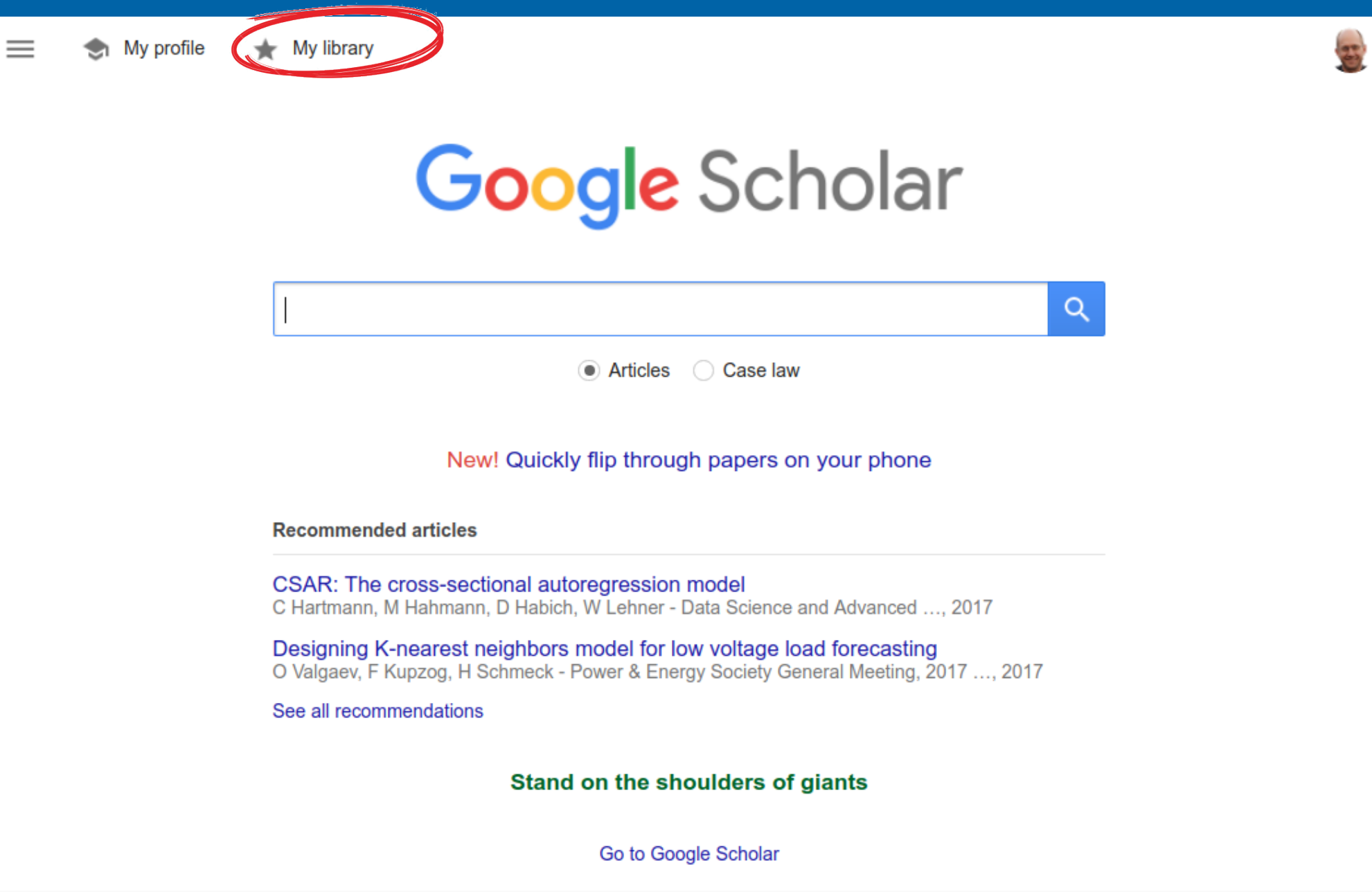
Use Google Scholar for more than searching



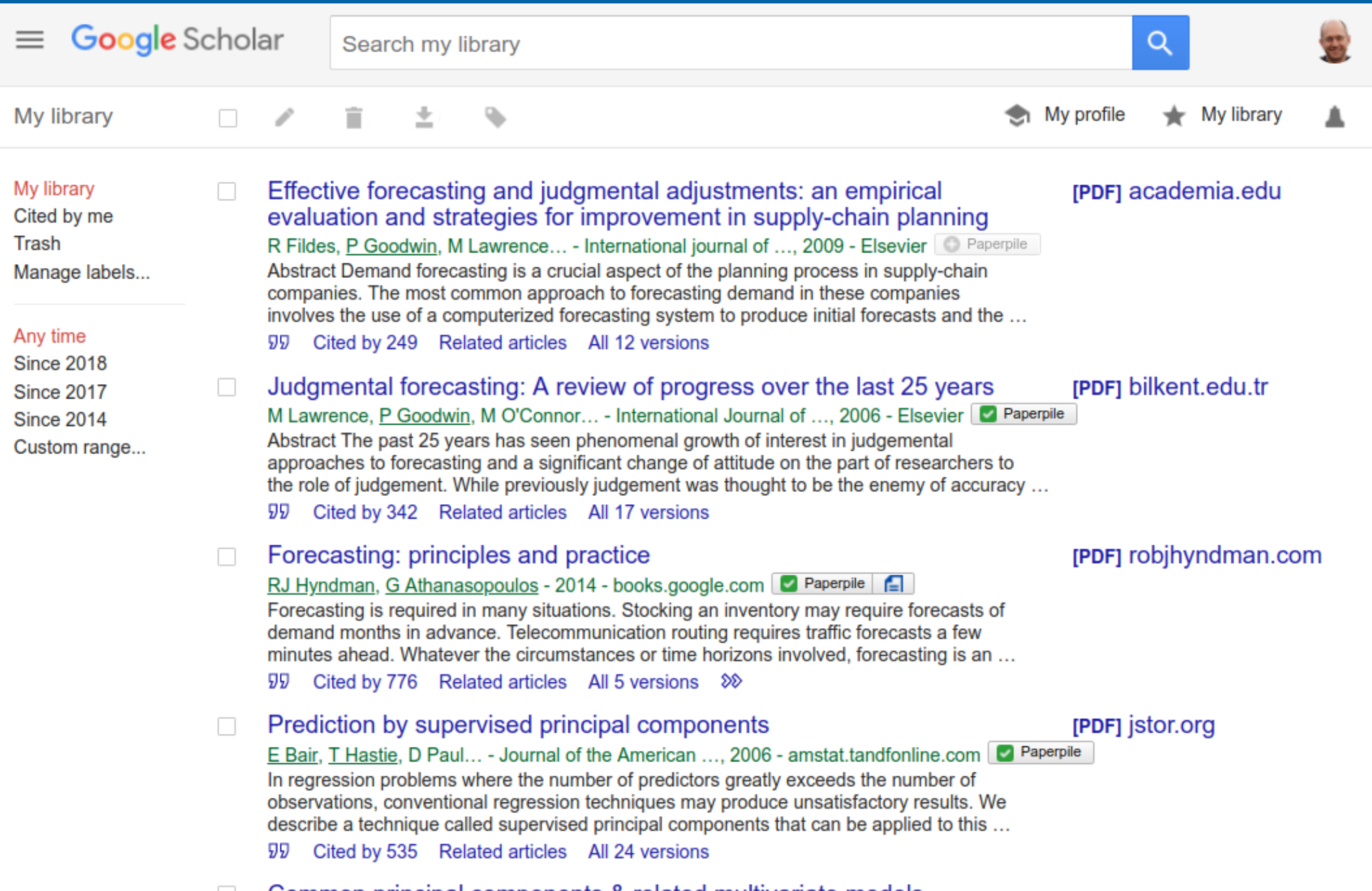
Use Google Scholar for more than searching



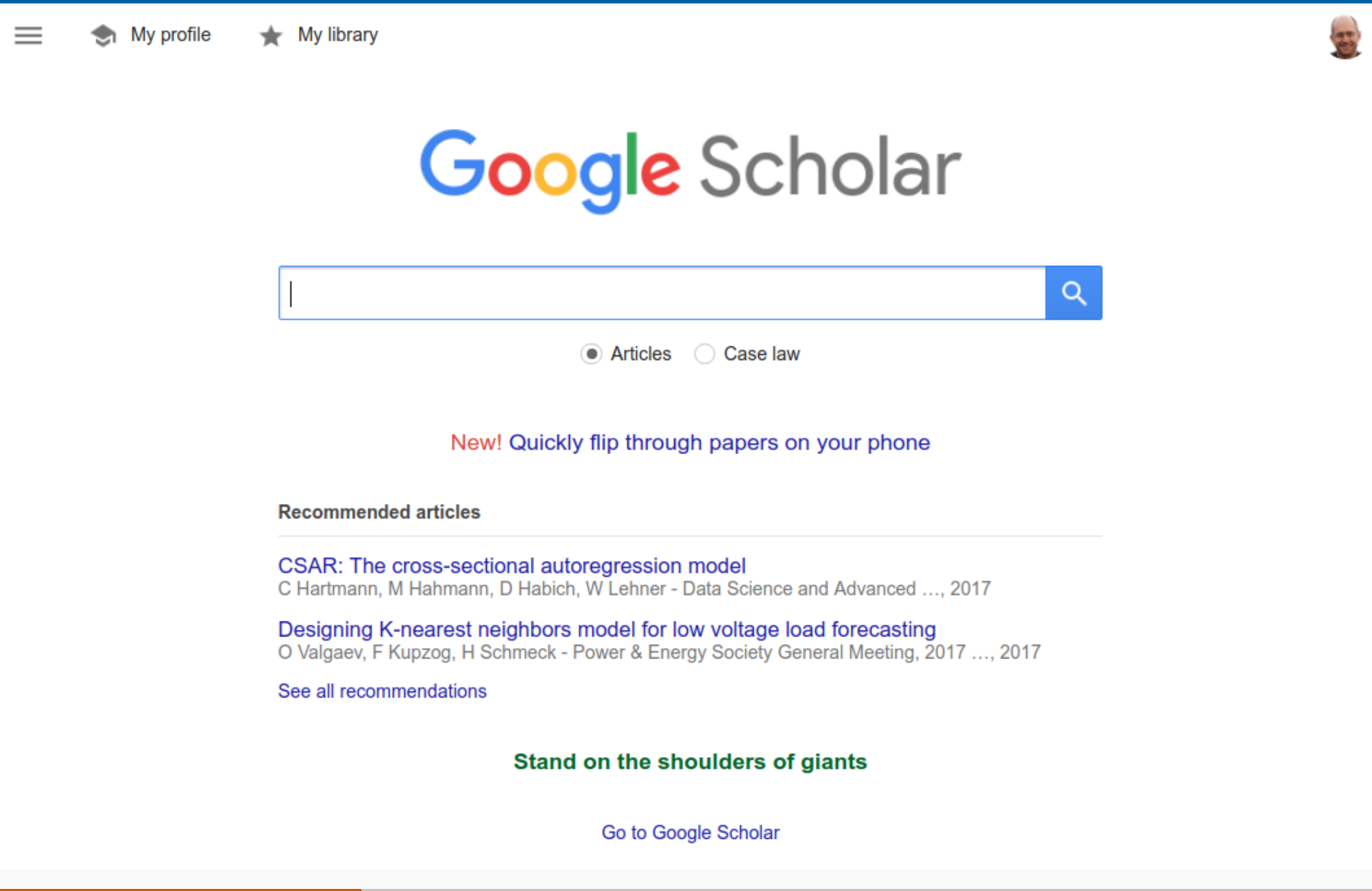
Use Google Scholar for more than searching



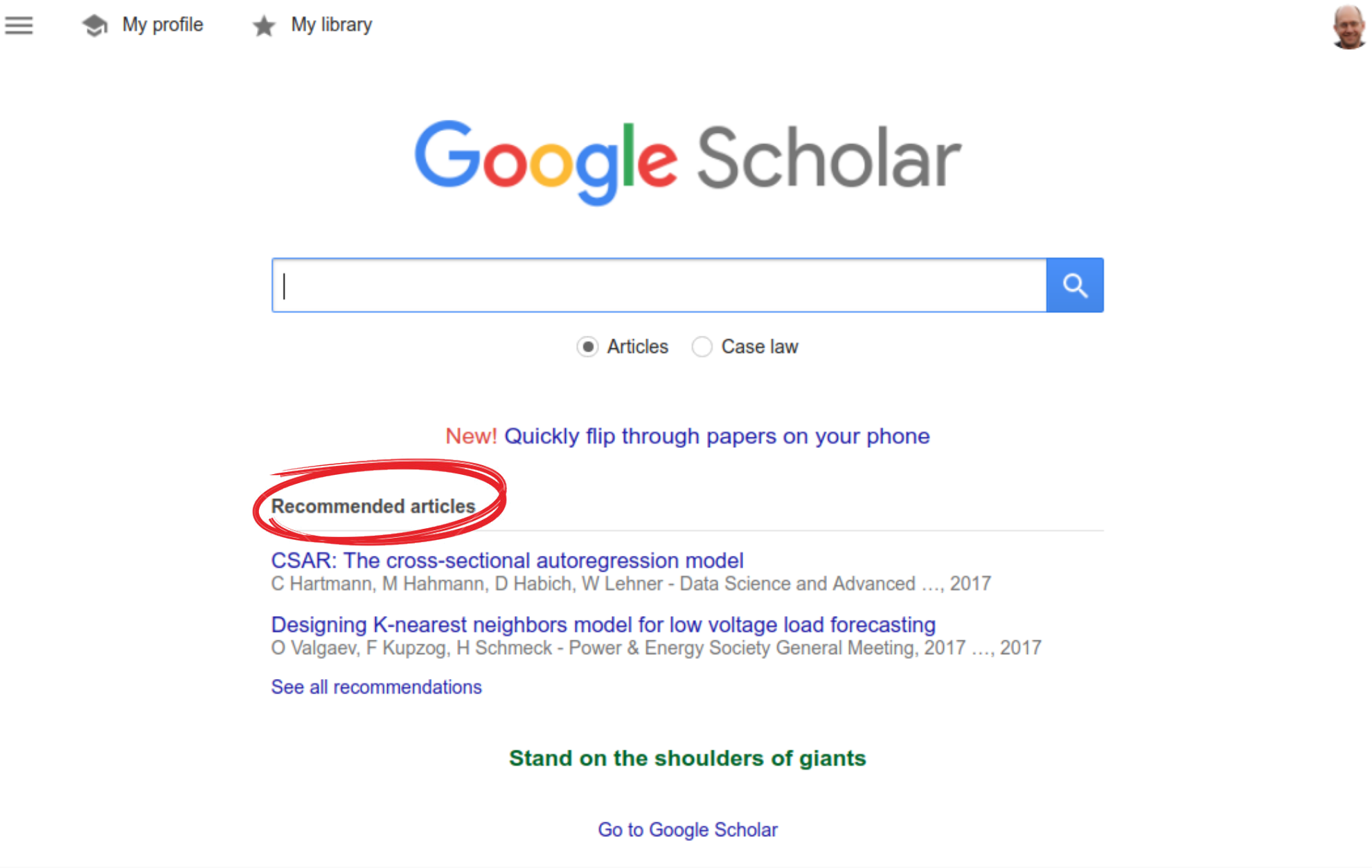
Use Google Scholar for more than searching



Use Google Scholar for more than searching



Use Google Scholar for more than searching



Outline

LaTeX

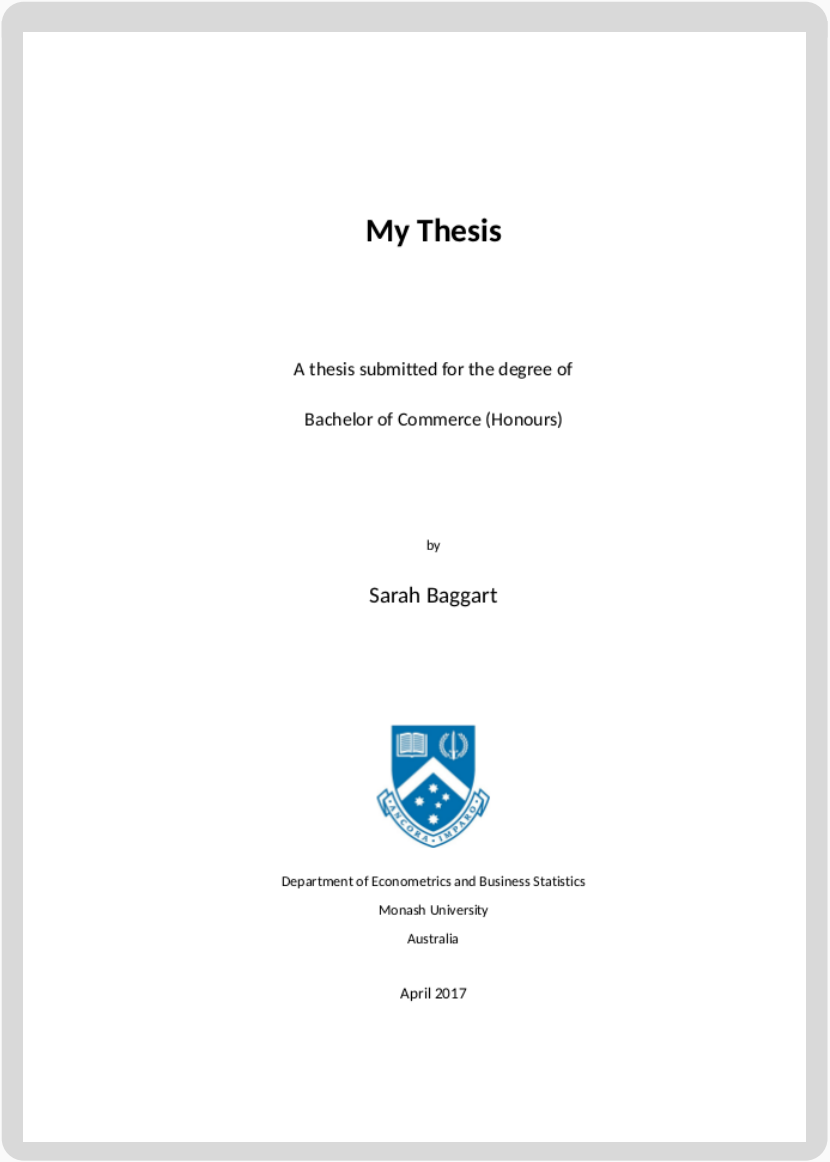
Document processing

- Free and open-source
- Available on all operating systems
- Used by every mathematical publisher
- Separate content from style
- Format complex equations
- Automatic numbering
- Automatic bibliography
- Monash thesis template available

To install:

- Download MikTeX, MacTeX or TeXlive.

RMarkdown



My Thesis

A thesis submitted for the degree of

Bachelor of Commerce (Honours)

by

Sarah Baggart

Department of Econometrics and Business Statistics

Monash University

Australia

April 2017

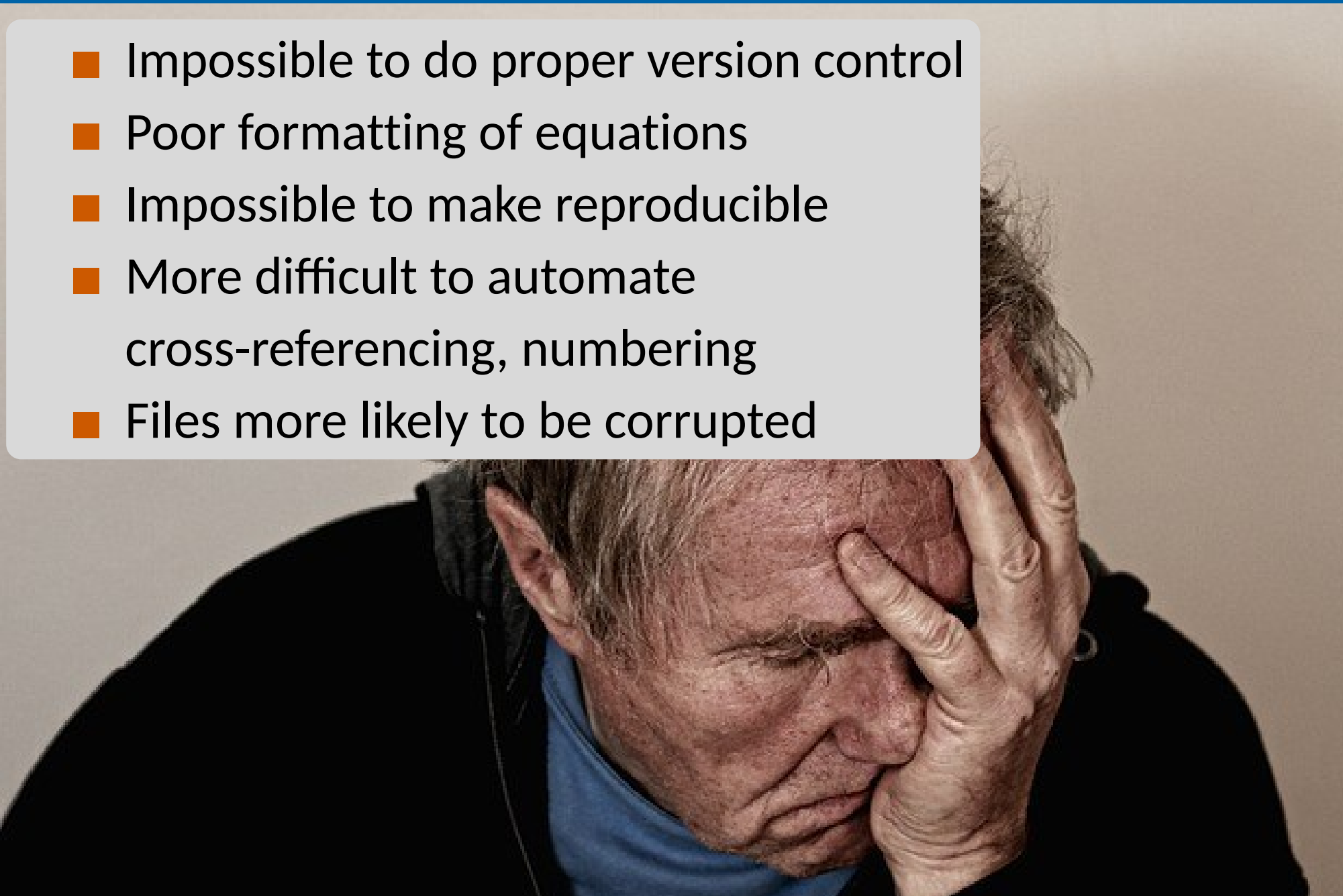
RMarkdown

- Combines R and LaTeX into one system
- Reproducible research
- Monash PhD template
- Monash working paper template
- Templates for most statistics journals
- Mandatory for my students

Microsoft Word

Microsoft Word

- Impossible to do proper version control
- Poor formatting of equations
- Impossible to make reproducible
- More difficult to automate cross-referencing, numbering
- Files more likely to be corrupted



Editors

Find a good text editor and learn how to use it.

- regex search and replace
- spell checking
- syntax highlighting
- linting

Some good editors

- Sublimetext
- Atom
- RStudio
- TeXstudio

Outline

Reproducibility

Not reproducible:

- Data edited in a spreadsheet
- Click and point analysis
- Copy and paste graphs and tables
- Tables typed by hand

Reproducible

- All data edits scripted
- All analysis scripted
- Graphs and tables automatically pulled in to the document
- Tables generated with scripts

Reproducibility

Someone should be able to reproduce your document without having to guess what software you had installed, what versions, which files do what, etc.

Reproducibility

Someone should be able to reproduce your document without having to guess what software you had installed, what versions, which files do what, etc.

- Stay organized.
- Your documents should be in *self-contained* repositories including bib files, Rmd/tex files.
- Be kind to future you. You will have amnesia.
- Track software versions

Reproducibility

Someone should be able to reproduce your document without having to guess what software you had installed, what versions, which files do what, etc.

- Stay organized.
- Your documents should be in *self-contained* repositories including bib files, Rmd/tex files.
- Be kind to future you. You will have amnesia.
- Track software versions

```
{r setup, include=FALSE}
knitr::opts_chunk$set(echo=FALSE)
writeLines(capture.output(sessionInfo()),
     "sessionInfo.txt")
```

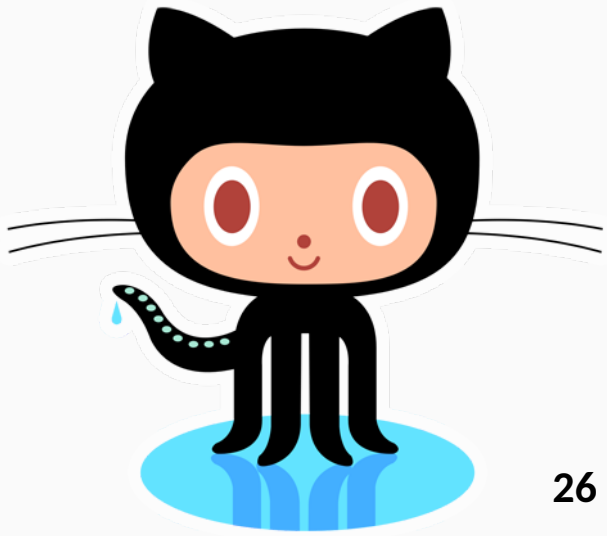
Version control

Version control

- `thesis_v1`, `thesis_v2`, etc., is not adequate version control.
- You need to track changes over time, have a *remote* repository, and be able to roll back as required.
- Your repository should contain *everything* required to produce your document including computer code, references, writing.
- Your repository should have an obvious structure and be fully documented.
- **Github** solves these problems
- Read “Happy git with R”: `happygitwithr.com`

Version control with git

- RStudio integrates with github, so version control is built in.
- But github can be used with *any* text-based language including Matlab, Stata, Python, LaTeX, R, Rmarkdown, markdown, etc.
- Git allows you to:
  - track changes
  - experiment in branches
  - undo
- Github provides:
  - backup and restore
  - synchronisation



Outline

Names and titles

- Your name is your brand. Choose one and stick to it.

Names and titles

- Your name is your brand. Choose one and stick to it.
- Titles should be informative, short and catchy (in that order).

Names and titles

- Your name is your brand. Choose one and stick to it.
- Titles should be informative, short and catchy (in that order).
  - Optimal forecast reconciliation for hierarchical and grouped time series through trace minimization
  - Exploring the sources of uncertainty: why does bagging for time series forecasting work?
  - Do levels of airborne grass pollen influence asthma hospital admissions?
  - Unmasking the Theta method
  - I am not an econometrician

Names and titles

- Your name is your brand. Choose one and stick to it.
- Titles should be informative, short and catchy (in that order).
  - Optimal forecast reconciliation for hierarchical and grouped time series through trace minimization
  - Exploring the sources of uncertainty: why does bagging for time series forecasting work?
  - Do levels of airborne grass pollen influence asthma hospital admissions?
  - Unmasking the Theta method
  - I am not an econometrician
- Some journals don't like imaginative titles.

Who should be listed as an author?

Monash authorship policy

. . . in all cases authorship must be based on making a substantial intellectual contribution to the work described and taking sole or joint responsibility for that contribution or, where appropriate, the work as a whole. Accordingly, authorship must be based upon a substantial contribution and responsibility for at least one, and usually more than one, of the following activities:

- Conception and design of the project;
- Analysis and interpretation of research data;
- Drafting significant parts of the work or critically revising it so as to contribute to the interpretation.

Who should be listed as an author?

Unacceptable inclusions of authorship

- Being head of department, holding other positions of authority, or personal friendship with the authors;
- Providing a routine technical contribution;
- Providing routine assistance in some aspects of the project;
- Acquisition of funding;
- General supervision of the research team;
- Providing data that has already been published or materials obtained from third parties (including the routine collation and provision of research source material).

Who should be listed as an author?

Unacceptable inclusions of authorship

- Being head of department, holding other positions of authority, or personal friendship with the authors;
- Providing a routine technical contribution;
- Providing routine assistance in some aspects of the project;
- Acquisition of funding;
- General supervision of the research team;
- Providing data that has already been published or materials obtained from third parties (including the routine collation and provision of research source material).

Acknowledge everyone who helped but is not an author.

Writing an abstract

1. What did you do?
2. Why did you do it? What question were you trying to answer?
3. How did you do it? State your methods.
4. What did you learn? State your major results.
5. Why does it matter? Point out at least one significant implication.

Writing an abstract

Original Article

Visualizing statistical models: Removing the blindfold

Abstract

Visualization can help in model building, diagnosis, and in developing an understanding about how a model summarizes data. This paper proposes three strategies for visualizing statistical models: (i) display the model in the data space, (ii) look at all members of a collection, and (iii) explore the process of model fitting, not just the end result. Each strategy is accompanied by examples, including MANOVA, classification algorithms, hierarchical clustering, ensembles of linear models, projection pursuit, self-organizing maps, and neural networks.

Writing an abstract

1. What did you do?
2. Why did you do it? What question were you trying to answer?
3. How did you do it? State your methods.
4. What did you learn? State your major results.
5. Why does it matter? Point out at least one

ELSEVIER

International Journal of Forecasting

Volume 23, Issue 2, April–June 2007, Pages 189-203

Bias in macroeconomic forecasts

Roy Batchelor

Show more

https://doi.org/10.1016/j.ijforecast.2007.01.004

Abstract

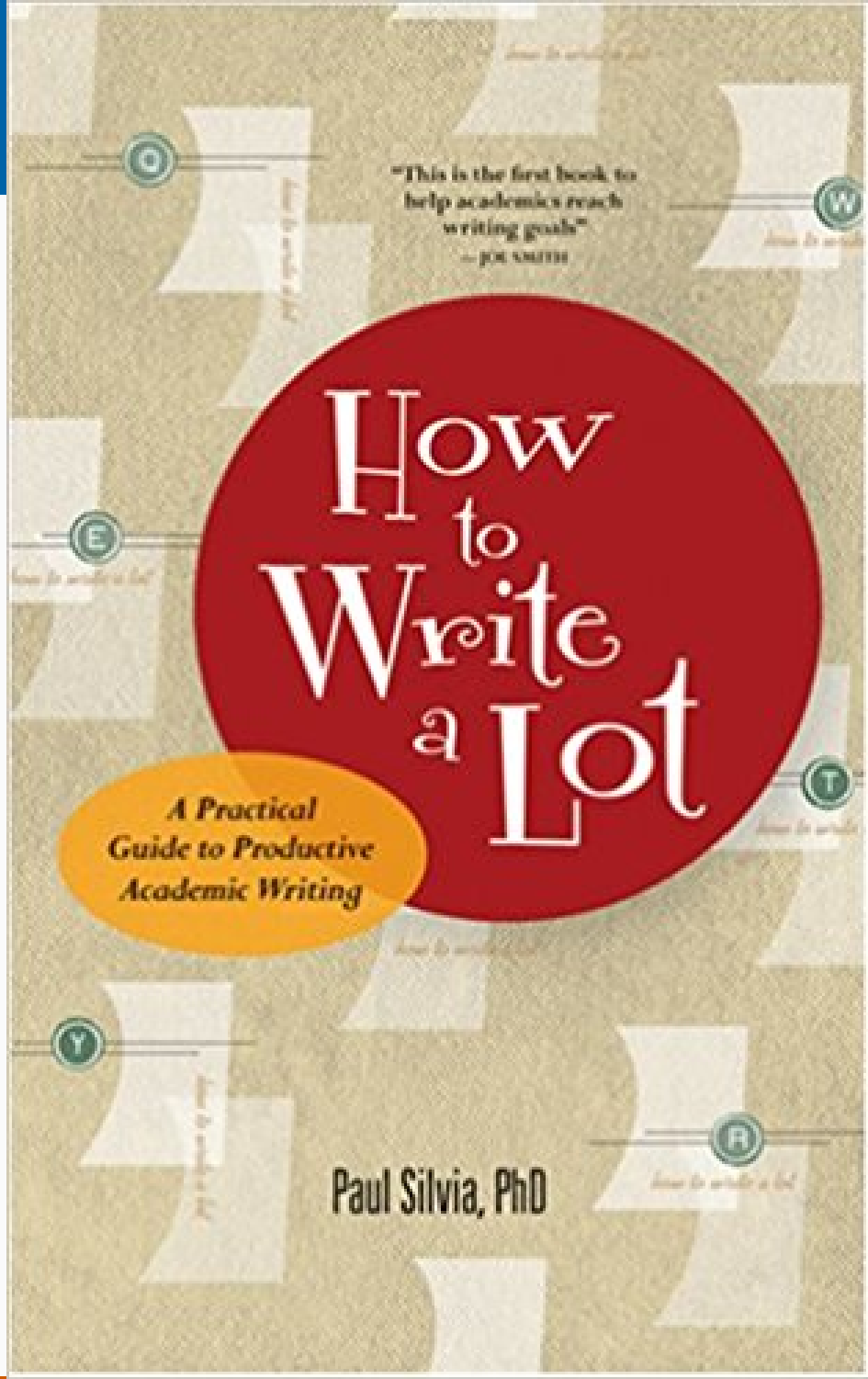
This paper documents the presence of systematic bias in the real GDP and inflation forecasts of private sector forecasters in the G7 economies in the years 1990–2005. The data come from the monthly Consensus Economics forecasting service, and bias is measured and tested for significance using parametric fixed effect panel regressions and nonparametric tests on accuracy ranks. We examine patterns across countries and forecasters to establish whether the bias reflects the inefficient use of information, or whether it reflects a rational response to financial, reputational and other incentives operating for forecasters. In several G7 countries – Japan, Italy, Germany and France – there is evidence of a change in the trend growth rate. In these circumstances, standard tests for rationality are inappropriate, and a bias towards optimism in the consensus forecast is inevitable as rational forecasters learn about the new trend. In all countries there is evidence that individual forecasters converge on the consensus forecast too slowly. However, the persistent optimism of some forecasters, and the persistent pessimism of others, is not consistent with the predictions of models of “rational bias” that have become popular in the finance and economics literature.

Writing an abstract

- Should be a stand-alone summary. It is the only thing most people will read.
- No references or citations.
- Write in the past tense.
- Be explicit, precise and concise.
- Stick to a single paragraph.
- Restrict background information to a sentence or two at most.
- Make sure that your abstract is consistent with what you reported in the paper.
- Write the abstract last

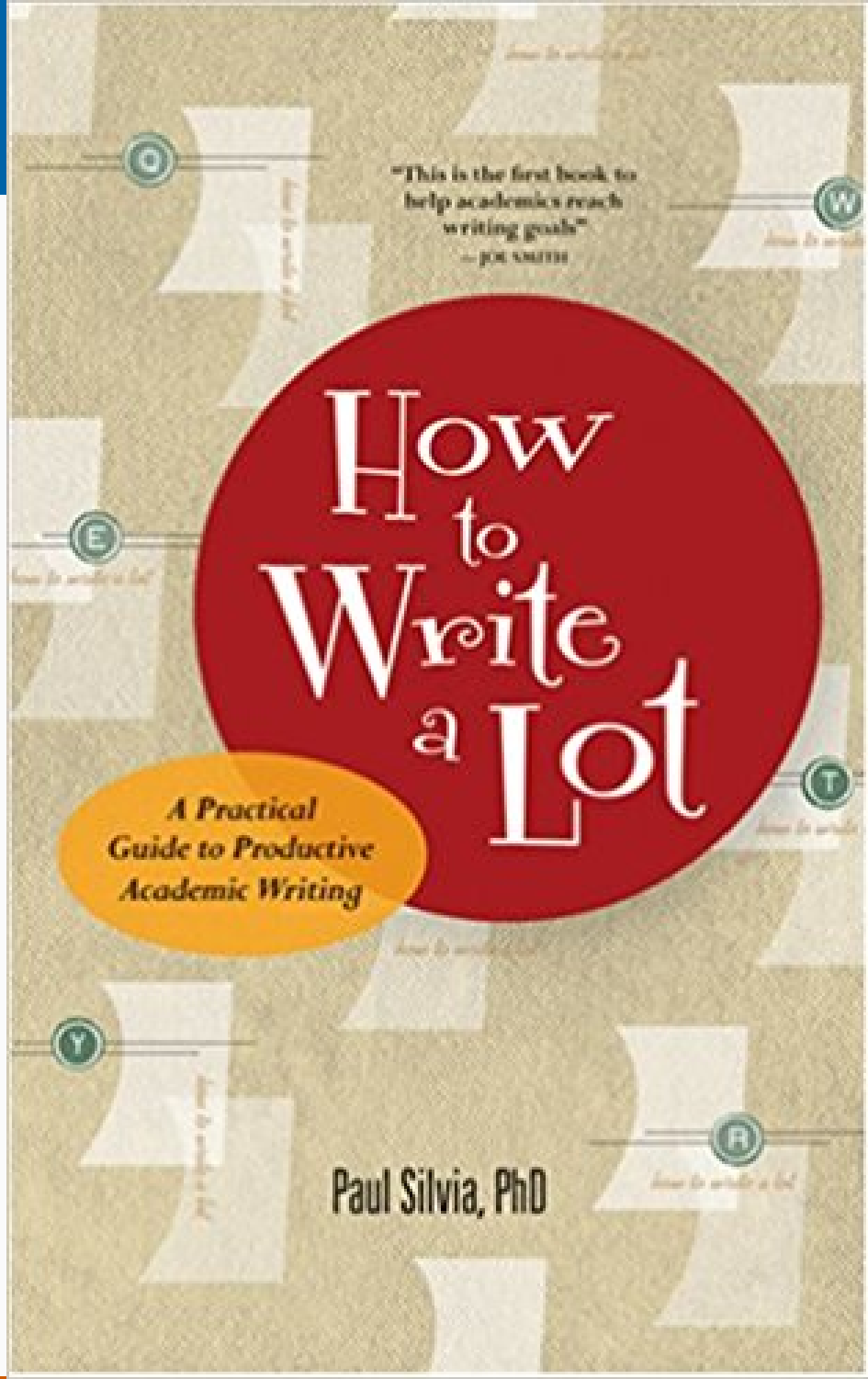
How to write a lot

- Block out a regular writing time (e.g., 2 hours every morning).
- Write something every day.
- Set a word or paragraph goal for each session.
- Practice makes ~~perfect~~ better.
- Writing clarifies thinking.
- Build on a scaffold.



Excuses

- There's not enough time
- I need to read more articles before I begin
- I need a new computer, a better printer, etc.
- I'm waiting for inspiration

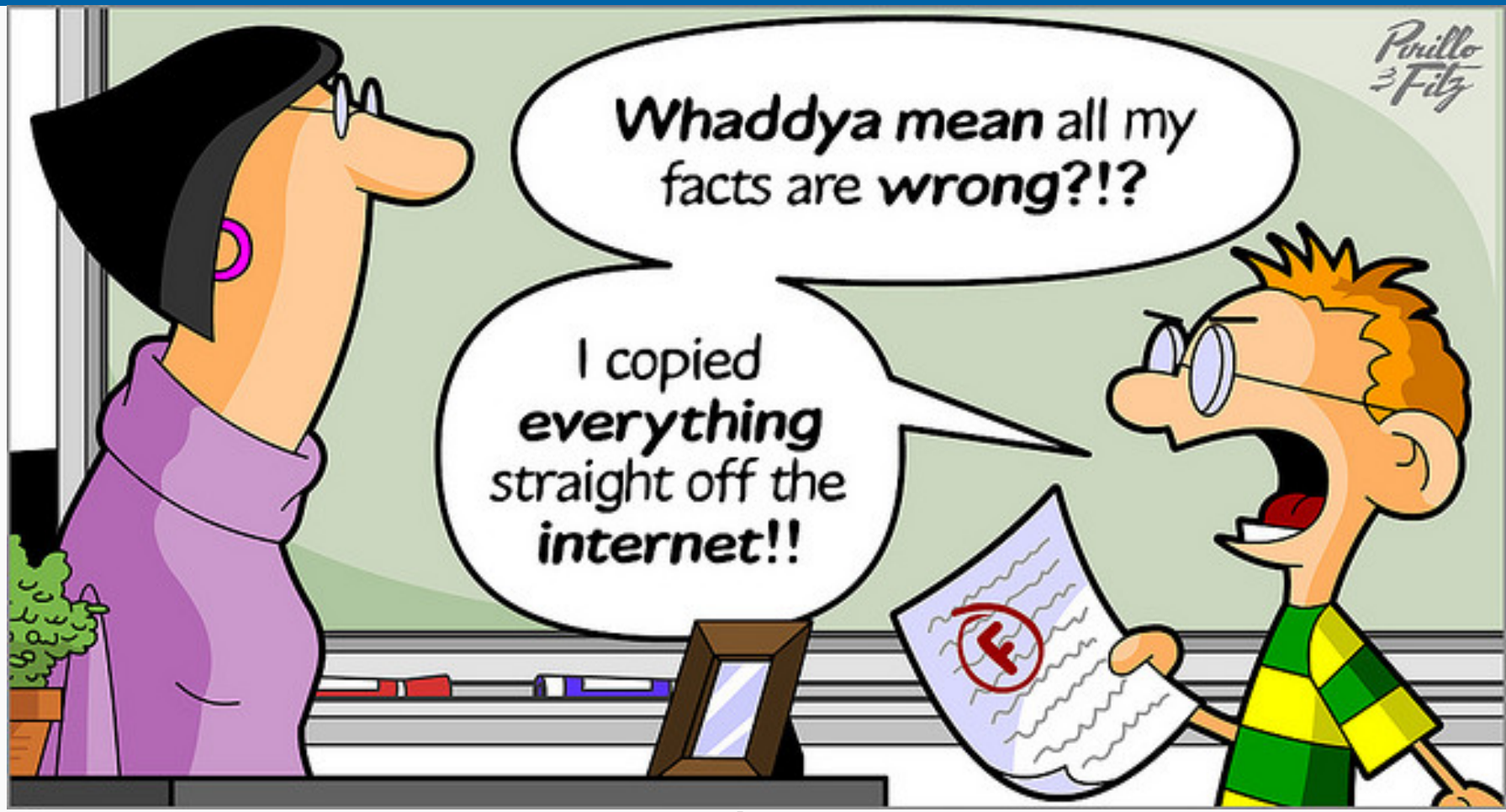


Words to avoid

According to Andrew Gelman

- Note that
- Interestingly
- It is interesting to note that
- Obviously
- It is clear that
- very
- quite
- of course
- Notice that

Plagiarism



- Never plagiarise from other papers — not even sentence fragments. Use your own words.
- Don’t plagiarise from your own papers either.

Outline

1. References
2. Languages and editors
3. Reproducibility & version control
4. Writing a paper or thesis
5. **Getting published**
6. Reviewer reports
7. Writing grant applications
8. Writing industry and media reports

Finding the right journal

- Beware of academic phishing!
- Consider journals of papers you cite.
- Check impact factors and journal rankings.
- Aim as high as possible, but be realistic.

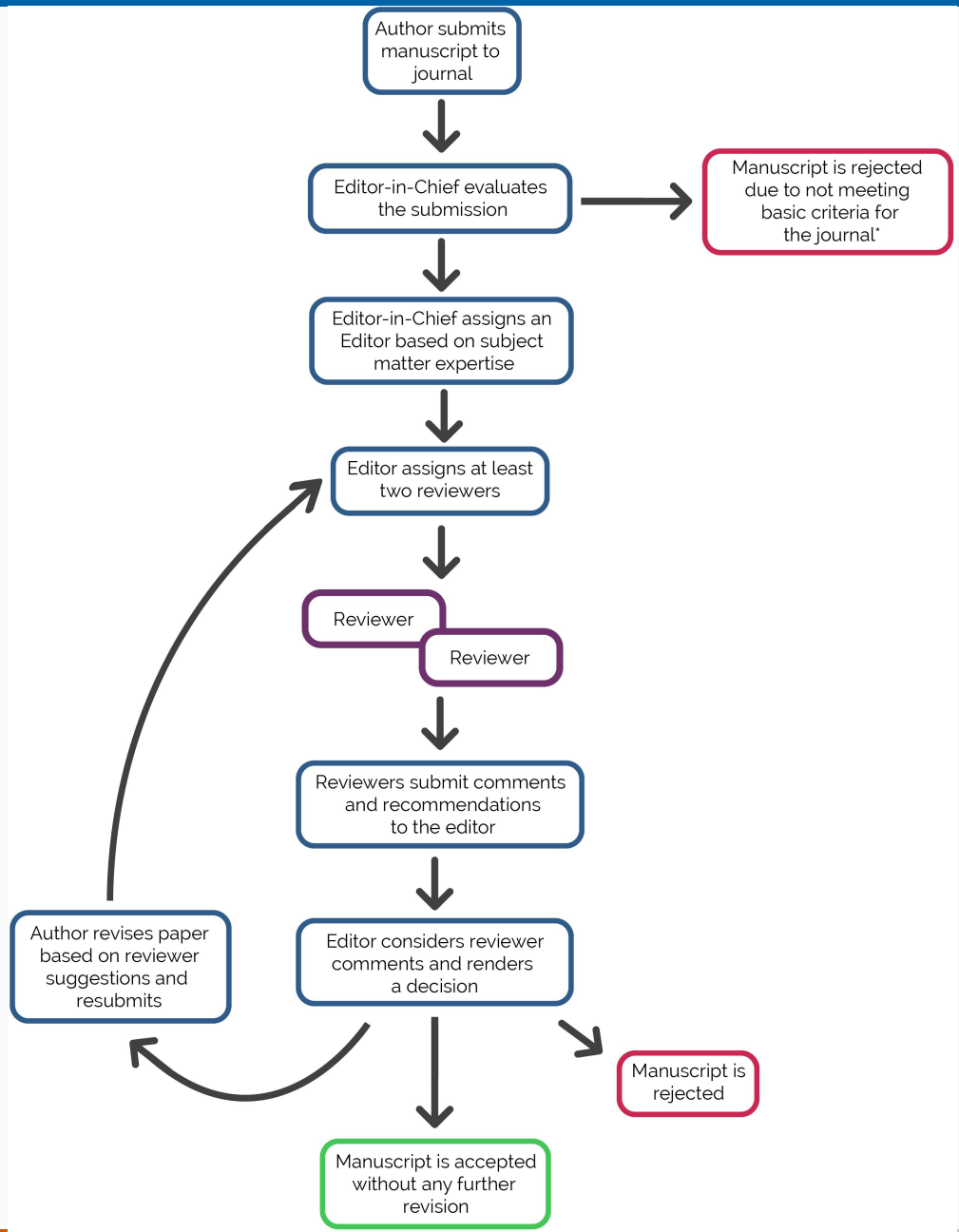
How do journals work?

Editor-in-Chief

Editors

Associate Editors

Reviewers



How do journals work?

Editor-in-Chief

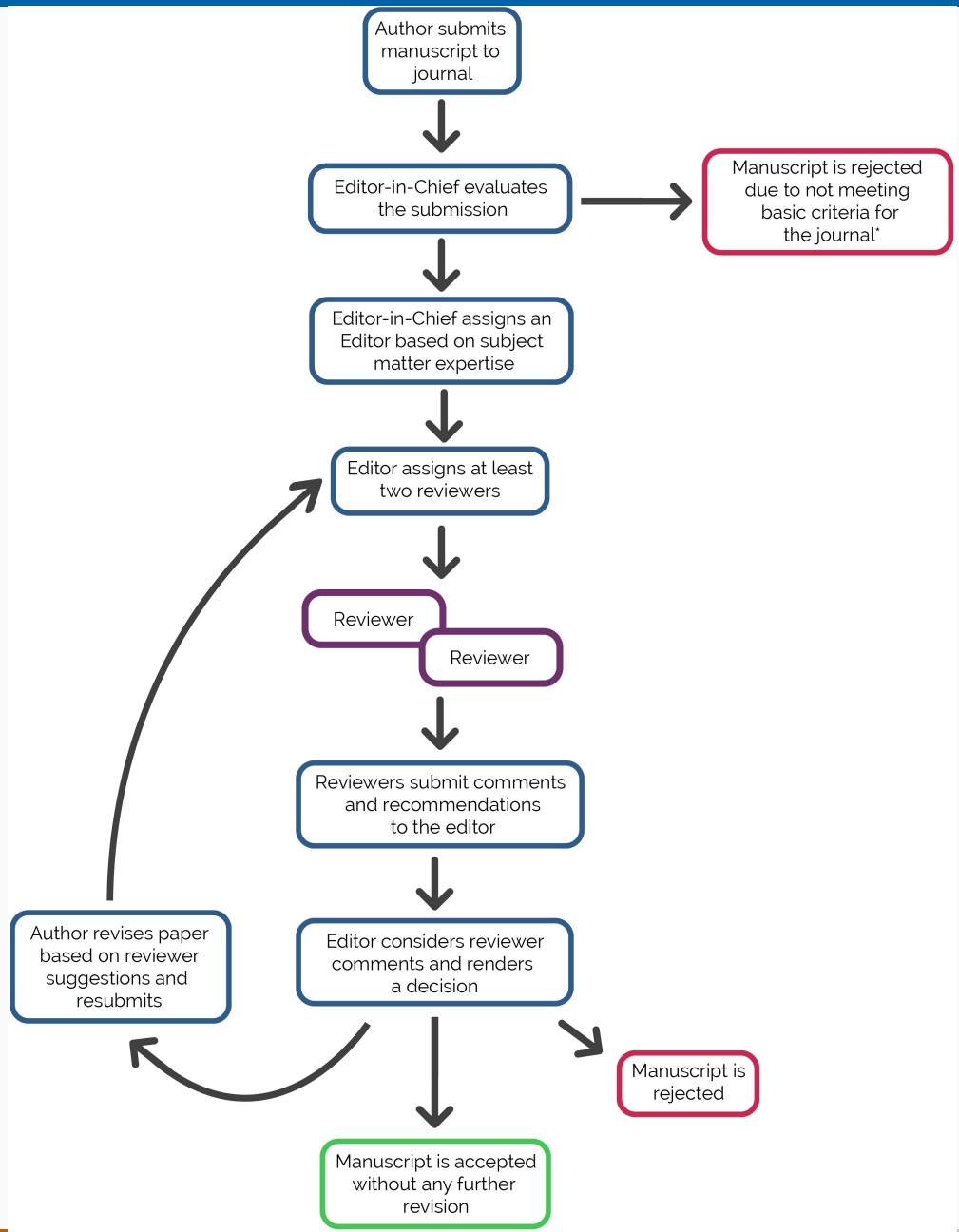
Editors

Associate Editors

Reviewers

Possible outcomes

1. Accept
2. Revise and resubmit
3. Reject and resubmit
4. Reject



Submitting to journals

- Don’t be too fussed about journal styles. Most journals are much more lenient than the guide to authors suggests.
- Use biblatex which makes it easy to change bibliographic styles if necessary.
- Don’t bother with long cover letters.
- Don’t grovel.
- Check the submission when requested.

An IJF rejection letter

Thank you for this submission, but as it consists entirely of the IJF author guidelines, it is not suitable for publication in the IJF. We publish original research, not author guidelines. Perhaps the *Journal for Guidelines* would be an appropriate outlet.

In future, when you are asked to check the pdf of your paper, you might find it useful to actually do so, rather than just claim to have done so. That way, you will avoid this kind of mistake.

Common reasons for rejection at the IJF

- Sending it to the wrong journal.
- Poor literature review
- No new ideas
- Limited empirical evaluation
- Outrageous claims

Outline

Dealing with reviewer reports

- Put the reviews aside for a couple of days until you calm down.
- Poor reviews indicate poor editors.
- The best journals have the best reviewers.
- If the reviewers misunderstood your paper, then it is not explained clearly enough.
- Unless you strongly disagree, do what the reviewers have requested.
- Make the changes, even if the paper has been rejected and you are sending it to a new journal.

Writing responses to reviewer reports

If the journal allows a resubmission, you need to write a response to the reviewers.

Author responses to Associate Editor comments

1. In Section 2.1: the notion of a reconciliation matrix $\boldsymbol{P}$ is introduced. This will not be clear to a majority of the readership of JASA. I suggest that a specific example of $\boldsymbol{P}$. I would also like to see an expanded discussion of the remark that "$\boldsymbol{SPS} = \boldsymbol{S}$ is required for unbiased forecasts". This would be helpful to the general readership. Its not obvious (at least, not to me). It also seems to be important since it was used in page 5 to explain why $\boldsymbol{\Sigma}_h$ is not identifiable.

   We have now included on page 7 examples for two choices of $\boldsymbol{P}$ which lead to the commonly used bottom-up and top-down approaches and a detailed explanation as to why $\boldsymbol{SPS} = \boldsymbol{S}$ is required for unbiased reconciled forecasts assuming that the base forecasts are unbiased.

2. On Equation (2) [now (3)]: What conditions do we need about dependence between $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_T$ and $\boldsymbol{\varepsilon}_h$? Is $\boldsymbol{\Sigma}_h$ the unconditional covariance matrix of $\boldsymbol{\varepsilon}_h$ or is it the conditional covariance (given $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_T$)? Of course both will be equivalent under independence between $\boldsymbol{\varepsilon}_h$ and $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_T$.

   It is assumed that $\boldsymbol{\varepsilon}_h$ is independent of observations $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_T$. We explicitly state this now right after equation (3).

3. From the definition of $\boldsymbol{W}_h$ in Lemma 1, the errors in Equations (5) and (6) have mean 0. How is this expectation computed? Is this conditional on $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_T$?

   As you have correctly anticipated, in Equations (5) and (6) errors have mean zero as the

Writing responses to reviewer reports

If the journal allows a resubmission, you need to write a response to the reviewers.

- No grovelling
- Cut and paste reviewer comments into response, then add your own comments beneath in a different colour/font.
- Give page/paragraph numbers for all changes.
- Respond to *all* the points with a simple but specific explanation of what you have done.
- If you strongly disagree, you need to persuade the editor (not the reviewer) of your perspective.
- Exception: bad editors sometimes act as rubber stamps for reviewers.
- Keep your response as short as possible. Respect the editor's time.

Becoming a reviewer

1. Write good articles
2. Get them published

Becoming a reviewer

1. Write good articles
2. Get them published

Why review?

- You learn a lot.
- You get better known by the research leaders in your area.
- You get to see the latest research before everyone else.
- The scholarly publishing system depends on it.

Writing a good review

1. What is the paper about?
2. What is the gap that it is trying to solve?
3. How does it address the gap? Do the methods/theory work, check what is promised
4. What sort of application is discussed? Is it contemporary, and interesting data problem, or data pulled from another paper, and a bit tired?
5. How well does the title/abstract describe the main contributions of the paper?
6. Is the introduction readable? If you have trouble understanding the problem from the intro there will be many other readers in the same situation
7. Is the solution original? Are there other published papers on the same problem? Have they been cited appropriately? Are they missing major existing work?

Writing a good review

1. Provide a general summary of the paper and its contribution.
2. Describe the major problems that need addressing.
3. List minor corrections required.

- Do not include a recommendation about whether to publish in the report itself.
- Be the reviewer you would like to have.

Outline

Writing grant applications

- Read the guidelines and rules and follow them.
- Read successful grant applications in area.
- Grab attention in the first paragraph
- Clearly specify aims
- Sell the project — why should someone spend money on you?
- Be clear and succinct
- Write for a non-expert audience. Decision-makers probably not in your research field.
- Explain project management and timelines
- Demonstrate credibility
- Explain role of every person on the grant and why they are needed.
- Build partnerships
- Ensure budget is realistic and justifiable.
- Allow time for many drafts.
- Find peers to review it.

Outline

Industry reports

- Reproducibility is crucial. The data will likely change several times during the project.
- Fully document what you did.
- Write in layers:
  1. Executive summary
  2. Introduction
  3. Body of report
  4. Appendices

  Each layer provides more details.
- Most people will only read the executive summary.
- Highlight recommendations in the first few pages.
- Remember the readers may not be experts.
- Include contact information somewhere.

Media reports

- Find a headline hook
- Write in layers
  1. Headline
  2. First paragraph provides a summary
  3. Next paragraphs provide more detail
  4. Finish with contact information
- Inverted pyramid style: details in descending order of importance.
- Avoid jargon, cliches and slang.
- Don’t over-simplify
- Don’t claim more than is true.
- Ask your media office for help.

Slides and resources

`robjhyndman.com/seminars/Writing`