WHAT THE LIBRARY CAN DO FOR YOU?

* Acquire for you pretty much any book, journal article, conference paper or thesis from any library in the world via interlibrary loans.

* Talk with you one on one about your project and assist you in developing a search strategy, finding resources and using them efficiently.

* Assist you with referencing the things that you find or using software to help you reference.

* Give you another pair of hands to seek out solutions to technical problems you may have.
MEET THE RESEARCH CYCLE…

Note the circularity – Searching and reading for a big assignment often happens in cycles over time. All along the way you can document your search and archive literature to refer to when you’re writing.
A BIT ON SEARCHING GENERALLY

* The size of the available literature will shape the way you search

* Generally you can only load 3 or 4 separate ideas into a single search before results get small.

* You can use synonymous terms to increase your reach and save you time when searching (though you can over do this too)

* It’s a good rule of thumb to search in more than one place with more than one variation of search terms.
PLACES TO LOOK — LIBRARY CATALOGUE

THE LIBRARY CATALOGUE CAN FIND BOOKS AND JOURNAL ARTICLES (AND MORE) WITH THE ONE SEARCH

Library Catalogue

data science

Sign in for all available resources, to save this search, or to request

0 selected  PAGE 1  42,513 Results

1. BOOK
Data science
Pierson, Lillian, author; Porway, Jake, writer of introductory text. 2017
Available at UTS Library. Open Collection (004 PIER)

* Remember, we can buy books and add them to our collection or borrow books from other libraries on your behalf. Just email me - david.litting@uts.edu.au
Databases vs the catalogue

* In addition to the catalogue, the library subscribes to specialized search engines called databases.

* Databases can either multidisciplinary or specialised.

* Databases also often allow you to control your search in more delicate ways than Google or Google Scholar.

Searching a database is a bit different to catalogue searching because of the matching that goes on.

* A catalogue record has only a few dozen to a couple of hundred words on it to match your search terms.

* Databases often search across entire documents so you get far more matches.

* If you are searching a database for a journal article, be aware that the titles of the papers are not as opaque as books. A book about management accounting will probably be called ‘management accounting’. This is often not true of journal articles...
CHOOSING A JOURNAL ARTICLE DATABASE...
(CLICKING ON ANY OF THE ICONS IN THE PDF VERSION WILL TAKE YOU TO THE DATABASE)
THINGS YOU CAN DO IN A DATABASE

Refining

Metrics

Combining sets
A DEMO SEARCH

Search

terroris* OR "counter terroris*" OR "protective security"

E.g., "Cognitive architectures" AND robots

AND

Search

design OR planning OR "built environment" OR urban OR property
This search was then limited by **Time** and by **Subject Heading**

(Patron was trying to filter out in Natural Disasters, Biological and IT based Terrorism)
RESULTS WERE HANDED BACK TO THE RESEARCHER IN **EXCEL** AND **ENDNOTE** FORM.
Citation tracking can also be useful...

ProQuest, Scopus, and Google Scholar are all good at this.

Take this article:

Reference lists look backwards in time

‘Cited by’ lists move forward in time
AFTER COLLECTING LITERATURE
IT’S A GOOD IDEA TO...

* Document the places you looked and the words you searched with – this will help you remember where you’ve been and prevent you from doing the same work twice.

* Keep a list of the things you’ve found (e.g. saving URLs of searches, or exporting results to a spreadsheet or reference manager). At the very least, save your PDF’s into a single folder on your computer, within a comprehensible and consistent file naming approach.

* Annotating the things you read may help you digest information, draw relationships and also save you having to re-read things in their entirety later on.

* Backing up to a cloud based server will save you if your computer dies.
IBISWorld Industry Report H4512
Fast Food and Takeaway Food Services in Australia

The healthy option: Operators are changing menus as consumer health consciousness rises
Marketline Advantage is like IBISWorld but tends to concentrate on consumer goods, rather than the big production and service industries. It also does data
Passport Euromonitor is like a cross between Marketline and IBISWorld, but has more of an international flavour than IBISWorld.
PLACES TO LOOK - RAW DATA

* Data can be found anywhere – including academic data that can be pulled from an article and business data that can be pulled from a business database. This is often ‘cooked’ data however – in the sense that it has already been interpreted and packaged...

* For raw (or lightly seared) data (and in addition to things you can find with a regular Google search) there are things like…
Tools like Aurin, ABS Tablebuilder and Data.gov.au allow you to call up data and visualisations mostly just by clicking a mouse.

The data you call up from these platforms can usually be then exported to more sophisticated tools if needed.
Google Datasets aggregates a search across various platforms – and represents a pretty interesting move forward in this still nascent field.

Academics also publish their raw data in places like Kaggle, Figshare and Research Data Australia. Kaggle even has a list of datasets they make sure are up to date and accurate. (Kaggle requires you to create a login before downloading)

There are many more places to get data (too many!) – if you want a piece of data and can’t find it, email david.litting@uts.edu.au.
DATA VISUALIZATION!

- There are a few different kinds of software that you can get for free as a UTS student that can do visualization, including:
  - Excel 2016
  - Tableau Academic
  - Power BI (scroll down in the list of apps to locate)

- There are also fancier (free) tools beyond that you can use, including Natural Language Processing Toolkits, JSON, Python and R
**DESKTOP DATA VIZ SOLUTIONS**

**Excel 2016**
* Free to enrolled staff and students
* Is familiar to many
* Creates charts without hassle
* 2016 can also do mapping and more advanced ‘Power Queries’

**Power BI**
* Free to enrolled staff and students (PC only though)
* Can import data from a variety of web sources and file types
* Easier than Tableau in some ways, but not always as configurable

**Tableau Academic**
*Free to enrolled staff and students
* Lets you control the look of a visualisation really well
* Works better if you understand pivot tables
This Guardian article is an interactive web-based data visualisation created by journalist Nick Evershed. You can actually build this visualisation yourself (it requires no coding) using freely available tools and data. Only a basic Excel type knowledge is required.
In this visualisation myself and a colleague found the heights of Spanish cities and their rainfall and tried to see if the rain falls mainly on the plain

(It turns out it doesn’t really - it tends to fall on the coast 😊)
EASY TO MIDDING DATA VIZ - A POWER BI DASHBOARD
This technique is used to analyse huge textual corpuses that a human would never have the time to parse manually.

The analysis is run by code-based programs (some like Stanford’s are free) that you have to train for your particular situation.

Whilst you don’t have to write code to use it you will need to be able to understand a bit about what it’s supposed to do and how to manipulate it.

Lynda has courses on how to do this.

* Never done it so just a guess 😊
MEDIUM TO HARD DATA VIZ - R PROGRAMMING LANGUAGE

R is a free open source programming language

Requires the use of programming language to call functions and analyse data.

So, rather than click a mouse to make something happen you write something like this:

```r
ggplot(df, aes(listicle_size)) +
  geom_histogram(binwidth=1, fill="#c0392b", alpha=0.75) +
  theme() +
  labs(title="Distribution of Listicle Sizes for BuzzFeed Listicles", x="# of Entries in Listicle", y="# of Listicles") +
  scale_x_continuous(breaks=seq(0,50, by=5)) +
  scale_y_continuous(labels=comma) +
  geom_hline(yintercept=0, size=0.4, color="black")
```
I’d like to do some of this stuff, where do I start?

• A lot of this stuff ends up being self-taught
• LinkedIn Learning has courses on all the software and techniques discussed today, including Power BI, Excel, Natural Language Processing and R
• Formal and informal support can be found at places like and Hacky Hour
• The Library can help too, even if it’s only pointing you in the direction of appropriate help
Interlibrary loan (used to order books journals etc from other libraries)

Proquest Ebooks – A powerful ebook platform.

Company and industry databases

Microsoft Office for students / Microsoft Power BI for students (Microsoft only, scroll down until you see Power BI in the list) / Tableau Academic for students

Any questions can always be directed to david.litting@uts.edu.au / 95143390