

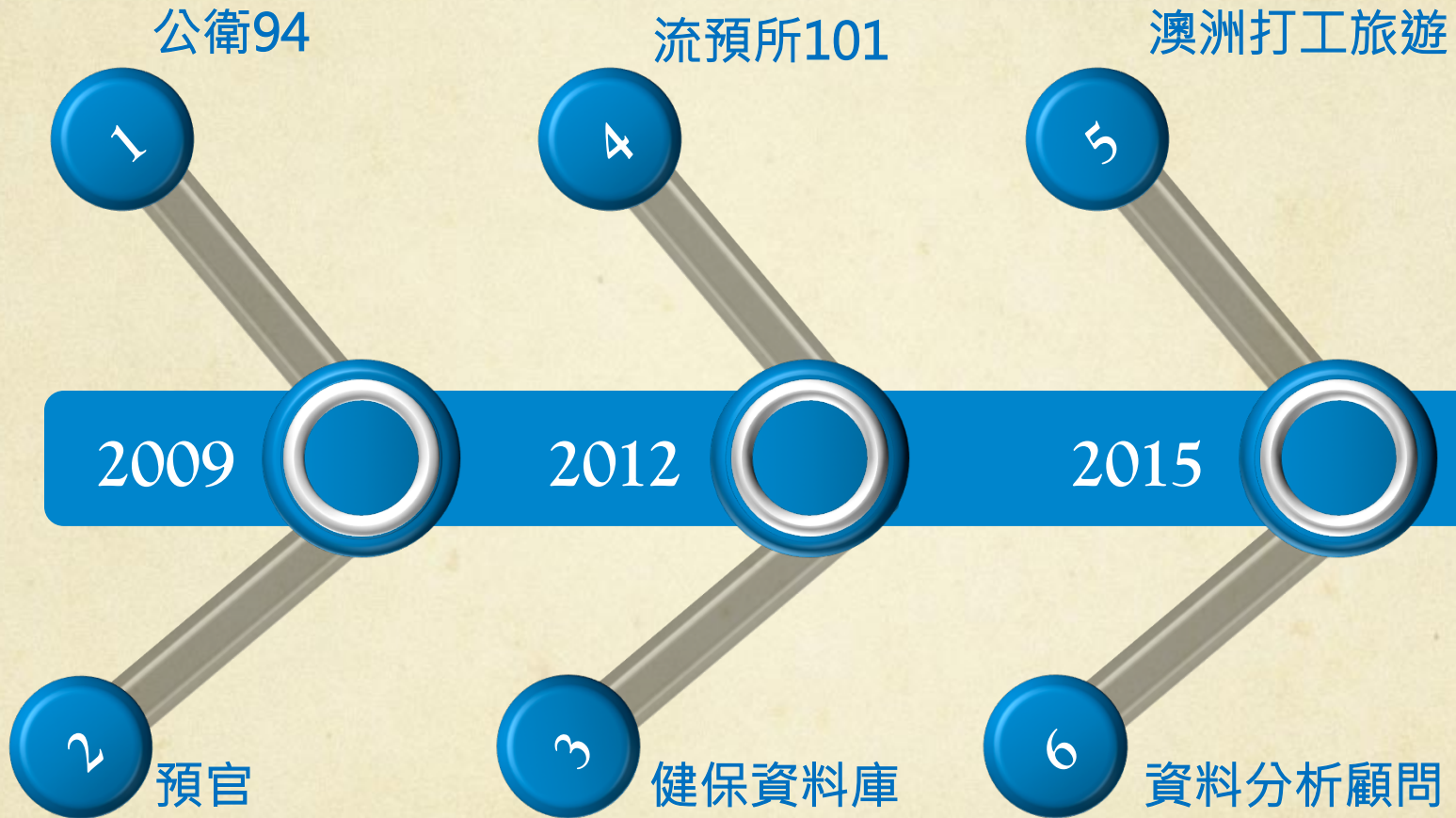


職涯經驗分享

陳建瑋

Data Analyst, Omniwaresoft

學經歷



澳洲打工旅遊

○目的：賺錢？練習英文？看看這世界！



我在做什麼？

DATA SCIENCE WORKFLOW



Storage and management

Novel tools such as **NoSQL** and **MapReduce** are bolstered by growth of global data, expected to reach 40 zettabytes by 2020.



Visualization

Flexible visualization tools such as **D3.js** and **Processing** extract insight from data and easily integrate with existing frameworks.

Visualization

1

Data acquisition and cleanup



Many **Python** libraries and specialized tools like **OpenRefine** and **Wrangler** aim to lower costs of data cleanup, which can claim up to 80% of development time.

2

3

Analysis



Analysis often involves revisiting raw data

Data scientists who use open-source tools such as statistical packages in **R** and **Python** report higher salaries than those who use commercial software.

4

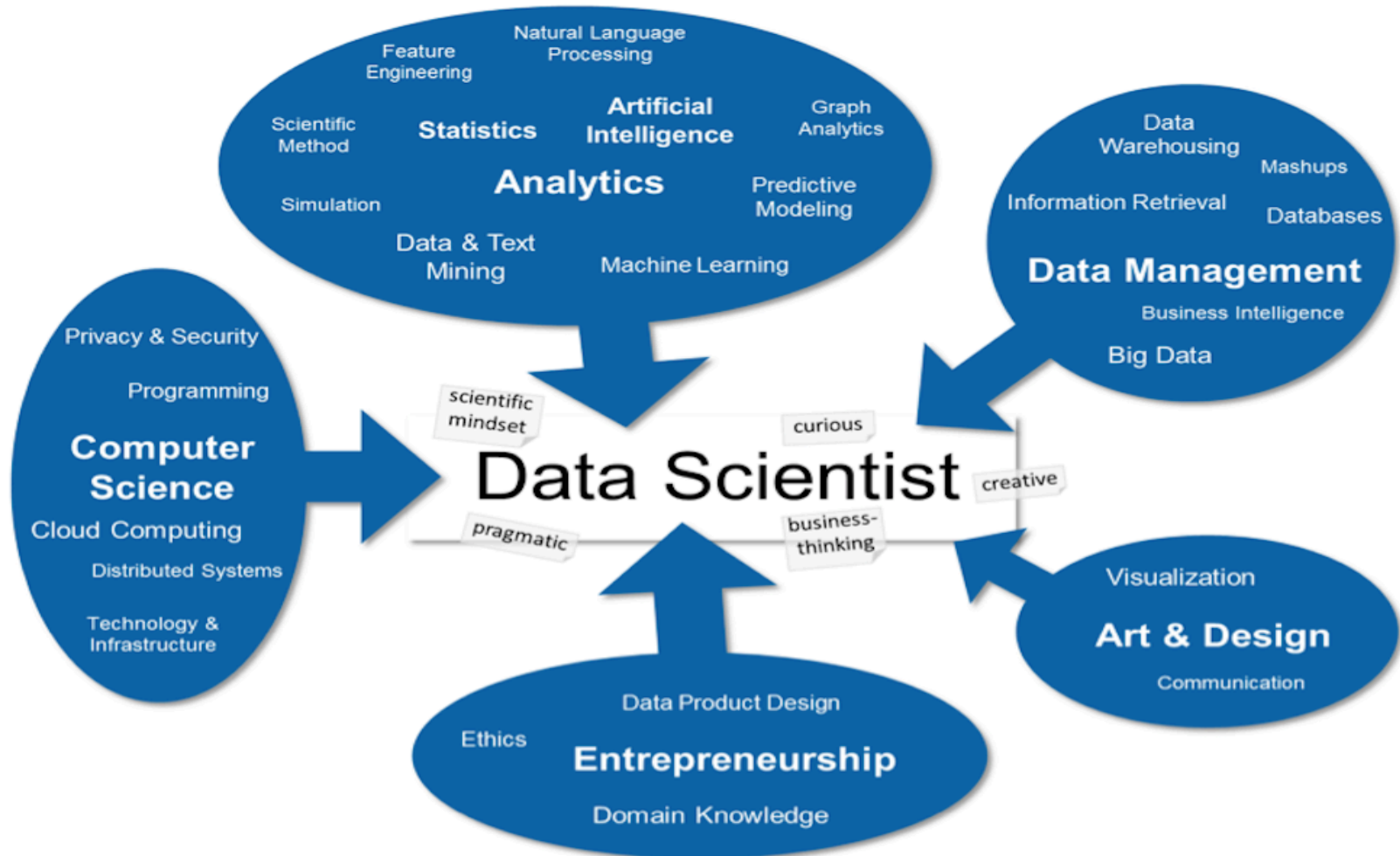
5

Communication



Collaborative services such as **GitHub** and **Bitbucket** simplify sharing code and distributing results, which in turn increases reproducibility.

Data Analyst → Data Scientist



找工作

- 只會SAS、R、SQL
- 打開104，這殺毀？
 - Database: MySQL、Hadoop、MongoDB...
 - Programming: Python
 - Visualization: SAS VA、Tableau...
 - ETL: Pentaho Data Integration...
- Coursera、Google...

建議

- 大一、大二：不同學院課程，打工
- 大三、大四：專精某項課程，打工
 - 統計：線性代數、多變量分析、迴歸分析...
- 畢業後
 - 當兵？(外交替代役)
 - 研究所？工作？打工旅遊？
- 碩一、碩二：解決問題 & 推銷產品能力，助教
- 共同：作品檔案，提早投履歷
- 2015年NTUPH線上就業博覽會

你比你所想的
更棒！