

# Introduction

- British – American
- I work at the Delft University of Technology in the Department of Multi-Actor Systems
- Associate Professor in Systems Engineering and Policy Analysis
- Associate Editor of *Technological Forecasting and Social Change*

A black and white photograph of a microchip mounted on a circuit board. The chip is square with a grid of pins around its perimeter. On its surface, there is some faint, illegible text and a date stamp that appears to be '00-02-9014'. The circuit board has various traces and components visible.

 WILEY

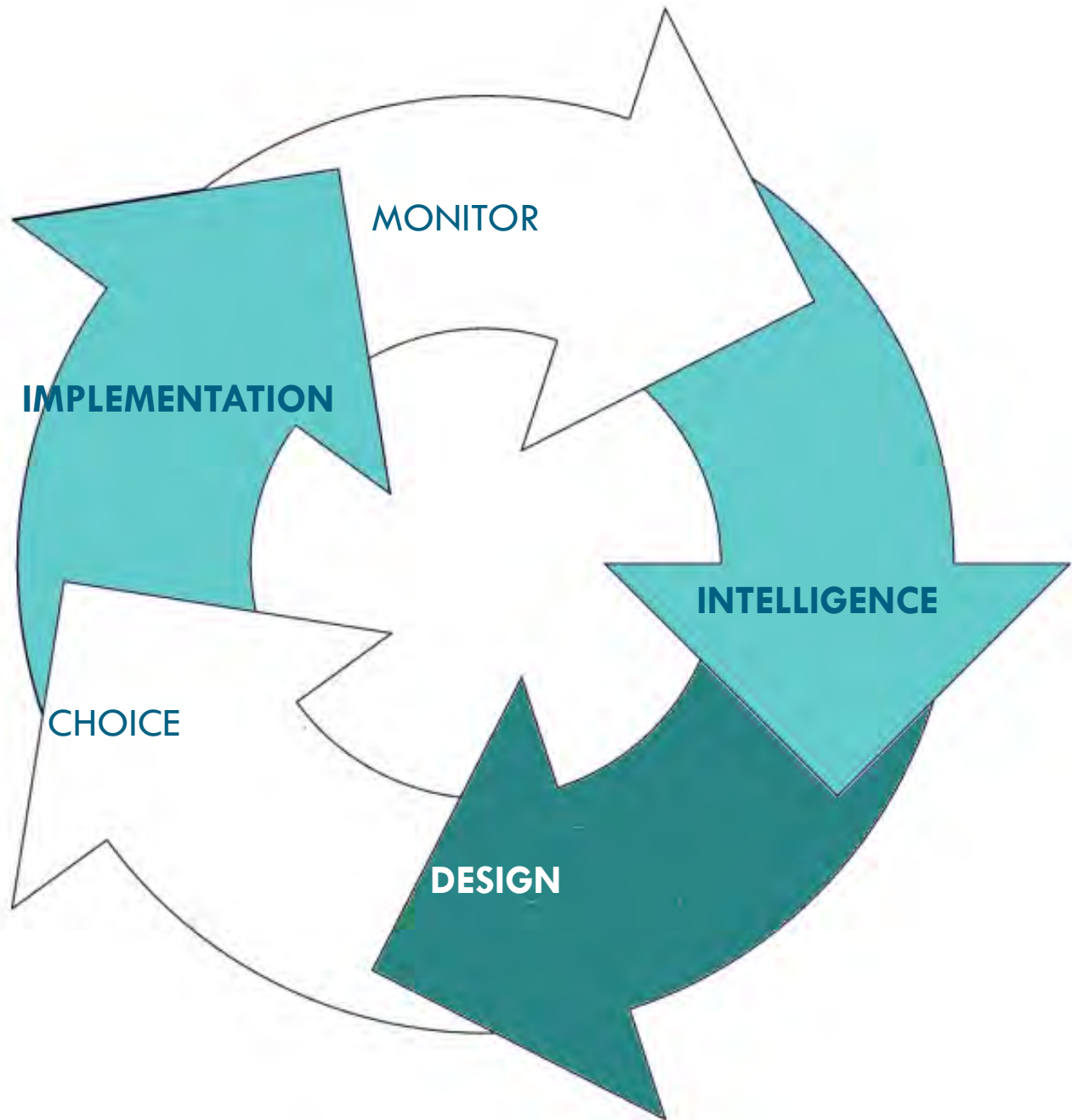
# TECH MINING

Exploiting New Technologies  
for Competitive Advantage



ALAN L. PORTER  
SCOTT W. CUNNINGHAM

*Wiley Series in System Engineering and Management*



**MONITOR**

# What is Tech Mining

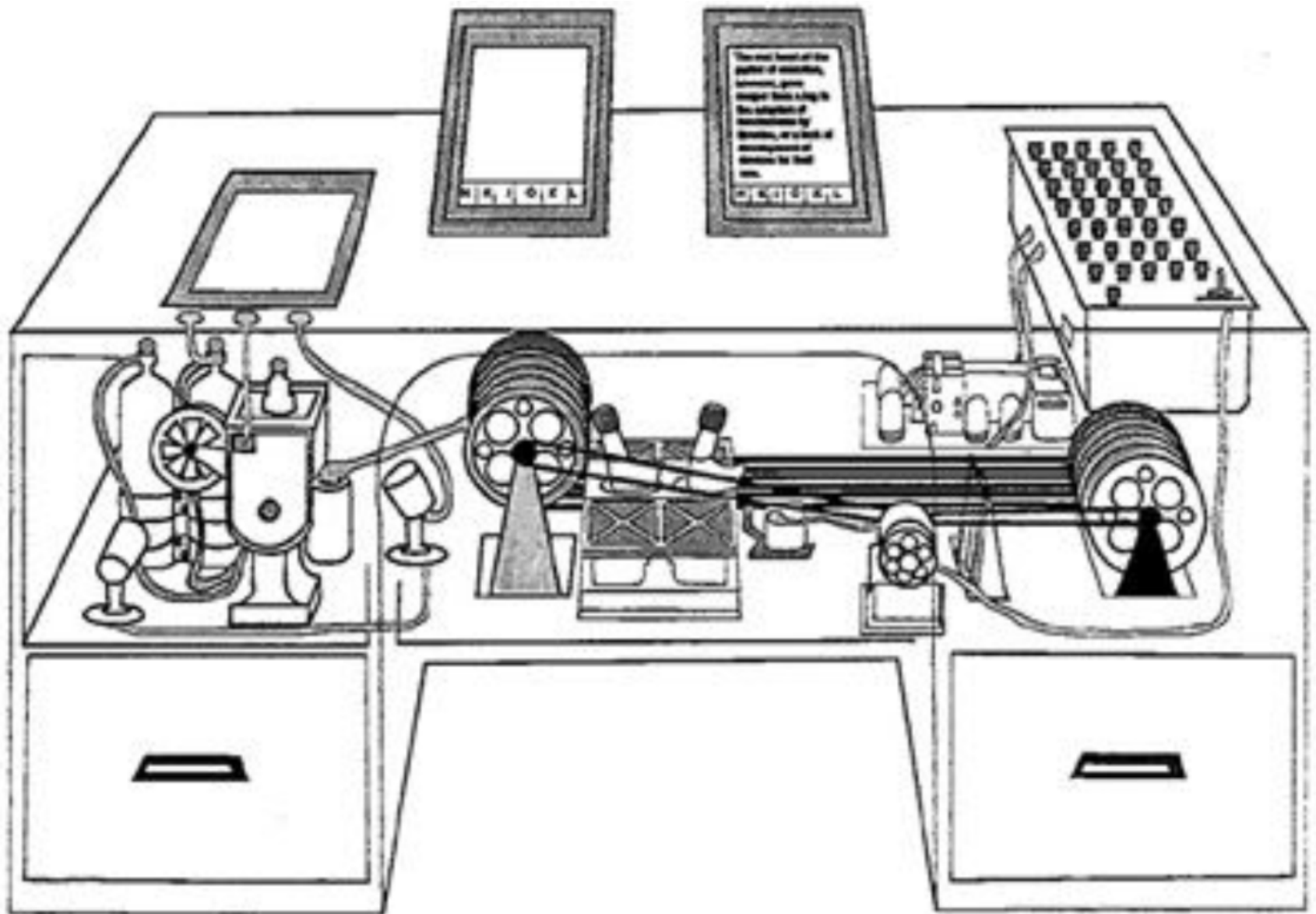
- Tech mining is the application of content analytic methods to the study of science, technology and innovation

# What is Content Analysis

- Content analysis is the analytical treatment of content
- Content meaning text, graphics, video and audio
- Analysis meaning detailed, objective, reproducible method

# Content Analysis in the Modern Era

- Prior to the internet, content analysis needed to be performed by hand
- Given the advent of vast amounts of machine readable text, we can automate our analytic procedures
- As time progresses, ever more text is being produced



Vannevar Bush's the Memex, 1945



# What Kind of Content for STI?

- To date, we've primarily been investigating content in the form of text
- There are two major sources of STI Content
  - Structured texts, and in particular scientific publications and technology patents
  - Unstructured texts, including social media and web pages

# What Kinds of Research Questions?

- Using publications and patents we can
  - Investigate dynamics in STI
  - Investigate content in STI
  - Investigate proximity in STI

# The Varied Nature of Proximity

- Proximity involves distances
- Proximity involves geography, regions and borders
- Proximity involves capabilities
- Proximity involves social, organizational and institutional factors as well

# The Research Approach

- Gets even more exciting when you combine these things
  - Dynamics + Content
  - Dynamics + Proximity
  - Content + Proximity

# Who Else is Interested in Tech Mining?

- Academic researchers and publishers, serving a professional development need
- Government agencies and non-governmental organizations, evaluation and indicators
- Private sector organizations, coordination, technological maturity, competitive intelligence

# What Are its Limitations?

- Output, not input indicators
- Partial, incomplete indicator.
- Early stages of tech maturity.
- Primarily retrospective



# What Are Your Alternatives?

- The clear alternative method is to seek expert opinion.
- Use tech mining to get the bigger contextual picture
- Talk to experts for the deeper, detailed, internally valid details
- The two approaches are quite complementary



**INTELLIGENCE**

# The Basis of Tech Mining

- What is a Scientific Paper?
- What is a Patent?

# A Scientific Paper Is

- A single, integral unit of scientific content
- Which is published in a scientific journal
- And is reviewed by an anonymous committee of scientific peers and deemed worth publishing

PHILOSOPHICAL  
TRANSACTIONS:  
GIVING SOME  
ACCOMPT  
OF THE PRESENT  
Undertakings, Studies, and Labours  
OF THE  
INGENIOUS  
IN MANY  
CONSIDERABLE PARTS  
OF THE  
WORLD.

*Vol I.*

For *Anno 1665*, and *1666*.

In the *SAVoy*,  
Printed by *T. N.* for *John Martyn* at the Bell, a little with-  
out *Temple-Bar*, and *James Allestry* in *Duck-Lane*,  
Printers to the *Royal Society*.

LE  
JOURNAL  
DES  
SCAVANS

*Du Lundy V. Janvier M. DC. LXV.*

Par le Sieur DE HEDOVILLE.



A PARIS.

Chez *JEAN CVSSON*, rue *S. Iacques*, à l'Ima-  
ge de *S. Iean Baptiste*.

M. DC. LXV.

AVEC PRIVILEGE DV ROY.

# Other Units of Scientific Content

- Conference Papers
- Book Reviews
- Editorials
- Books
- Book Chapters

# What Is a Scientific Journal?

- A group of papers published periodically
- Journals vary in their homogeneity of content
- Journals vary in their academic prestige
- Journals are mixed media, hard copy and electronic
- Many journals have a mixed mode of operation with both academic editors, as well as corporate publishers

# Why Publish at All?

- Scientific publishing is part of the open ethos of academic scientific research
- Citation is part of the gift economy of science
- Most papers are authored by academics, from a relatively select group of research universities
- Further, most scientific researchers publish very few articles

# Sociology of Science

- The publication and disciplinary life-cycle operates very differently by field
- This limits our ability to generalize across fields to disciplines as a whole
- Reasons are not very well-understood, but are loosely described as “sociology”



# Sociological Differences

- Nature of contributions
- Volume of contributions
- Aging of the literature
- Citation culture
- Publication outlets
- Cognitive differences
- Number of researchers
- Collaborative differences

# Databases of Science

- Large silos of knowledge produced and for whom access is sold by leading publishers
- These silos contain an fraction of the world's scientific literature; an elite group of flagship journals included
- Scientific articles are abstracted, with various value-added features included, as a business model

# Open Publication

- There is a move to more open publishing models
- Move from pay to access to pay to publish
- Academics say “we do the work”
- Publishers say “not so easy”
- Research funders say “why pay twice?”

# Anatomy of a Scientific Record

FN Thomson Reuters Web of Science™  
VR 1.0  
PT J  
AU Bouzas, V  
Haller, T  
Hobi, N  
Felder, E  
Pastoriza-Santos, I  
Perez-Gil, J  
AF Bouzas, Virginia  
Haller, Thomas  
Hobi, Nina  
Felder, Edward  
Pastoriza-Santos, Isabel  
Perez-Gil, Jesus  
TI Nontoxic impact of PEG-coated gold nanospheres on functional pulmonary  
surfactant-secreting alveolar type II cells  
SO NANOTOXICOLOGY  
LA English  
DT Article  
DE lung; alveolar type II; gold nanoparticles; toxicity; exocytosis  
ID DRUG-DELIVERY; INORGANIC NANOPARTICLES; CELLULAR TOXICITY; FLUORESCENT  
DYE; COLLOIDAL GOLD; ALAMAR BLUE; IN-VITRO; SIZE; CYTOTOXICITY; LUNG  
AB The outstanding properties of gold nanoparticles (NPs) make them very attractive for biomedical applications. In particular, the inhalation route has gained considerable interest as an innovative strategy for diagnosis and treatment of pulmonary diseases. It is, therefore, important to scrutinise the potentially deleterious or side effects of NPs on lung epithelium. The present study investigates, for the first time, the impact of polyethylene glycol (PEG)-coated NPs on freshly purified primary cultures of rat alveolar type II (ATII) cells. These cells play a central role in the respiratory function of the lungs. They are responsible for synthesizing and secreting pulmonary surfactant (PS), which is required to stabilise the respiratory surface during breathing dynamics. Cytotoxicity and cellular uptake of NPs was evaluated by analysing morphology, viability and exocytotic activity of ATII cells (PS secretion). The impact of ATII cells' exposure to NPs was studied in a wide range of gold concentration with particles sizes of 15 and 100 nm. The results show that PEG-coated NPs are very modestly internalised by ATII cells and it neither leads to detectable morphological changes nor to decreased cell viability nor to alterations in basic functional parameters such as PS secretion, even on exposure to high gold concentration (similar to 0.2 mM) during relatively long periods of time (24-48 h).  
C1 [Bouzas, Virginia; Perez-Gil, Jesus] Univ Complutense, Fac Biol, Dept Bioquim & Biol Mol 1, E-28040 Madrid, Spain.  
[Haller, Thomas; Hobi, Nina] Med Univ Innsbruck, Div Physiol, Dept Physiol & Med Phys, A-6020 Innsbruck, Austria.  
[Hobi, Nina; Felder, Edward] Univ Ulm, Inst Gen Physiol, D-89069 Ulm, Germany.  
[Pastoriza-Santos, Isabel] Univ Vigo, Fac Quim, Dept Quim Fis, Pontevedra, Spain.  
RP Perez-Gil, J (reprint author), Univ Complutense, Fac Biol, Dept Biochem, Madrid, Spain.  
EM thomas.haller@i-med.ac.at; jperezgil@bio.ucm.es  
FU Spanish Ministry of Economy and Competitiveness [BIO2012-30777-R]  
[MAT2010-15374]; Regional Government of Galicia [GRC2007-0010, MAT2010-15374]; Xunta de Galicia [INBIOMED-PA]

Journal

Content

Authors

Organizations

Regions

Nation

Funder

Citations

Times Cited

SC

Year

# What is a Patent?

- A patent is an exclusive but temporary right to develop and sell a technology
- Patents are awarded by individual nations which may vary in their criteria for awarding a patent

# Why Patent at All?

- Inventors often as a term of employment
- Inventors may also be entrepreneurs
- Assignees are granted the rights of the patent by the inventors
- Governments have a mixed interest in competition and also innovation
- Assignees (often companies) have an interest in protecting assets and appropriating the benefits of research

# Anatomy of a Patent Record

Inventors

Assignees

Regions

Family

Patent Office

Content

Classification

Citations

Filing Date

Priority Date

PT P  
PN WO2014057480-A2  
TI Analysis device used in automatic detection system for evaluating sample material e.g. blood or urine contained in sample volume has detector that detects light from sample volume in response to irradiation of sample volume by light source.  
AU BIAAF R I  
KENT N Y  
AE KONINK PHILIPS NV (PHIG-C)  
GA 201405744  
AB NOVELTY - The analysis device has a light source (307) that irradiates sample volume, and a detector that detects light from sample volume in response to irradiation of sample volume by the light source. An assessment unit determines the fill level of sample volume based on detected light.  
USE - Analysis device used in automatic detection system (claimed) for evaluating sample material e.g. blood or urine contained in sample volume to automatically detect filling states of sample volume used in field of home monitoring.  
ADVANTAGE - The potential blind spots at positions of air bubbles or other filling discrepancies are largely avoided due to detection of light from the entire sample volume. Improves image quality and reduces the maximum area of blind spots due to increased optical detection by using transparent material such as polycarbonate material. Allows detecting the presence of bodily fluids or substances which absorb light of the wavelength range of 475-575 nanometers or 260-380 nanometers.  
DETAILED DESCRIPTION - INDEPENDENT CLAIMS are included for the following:  
(1) a sample container configured for holding sample material in a sample volume; and  
(2) a method for assessing the fill level of a sample volume.  
DESCRIPTION OF DRAWING(S) - The drawing shows a schematic diagram of the system for automatic detection implemented in a stationary reader.  
Stationary reader (300)  
Aperture (302)  
Depth of field (305)  
Diffuser (306)  
Light source (307)  
DC A89 (Photographic, laboratory equipment, optical); B04 (Natural products and polymers, testing, compounds of unknown structure); J04 (Chemical/physical processes and apparatus including catalysis); F21 (Diagnosis, surgery); S03 (Scientific Instrumentation, photometry, calorimetry); S05 (Electrical Medical Equipment)  
MC A12-V03C2; B11-C06; B11-C07B2; B11-C07B7; B11-C11; B12-K04; J04-B01; S03-E15; S05-D01  
IP A61B-005/00; B01L-003/00; G01N-025/10  
PC WO2014057480-A2 17 Apr 2014 G01N-025/10 201428 Pages: 42 English  
AD WO2014057480-A2 WOIB059920 12 Oct 2013  
FI US713030P 12 Oct 2012  
DS WO2014057480-A2:  
(National): AE; AG; AL; AM; AO; AT; AU; AZ; BA; BB; BG; BH; BN; BR; BW; BY; BZ; CA; CH; CL; CN; CO; DE; DK; DM; DO; DS; EC; EE; EG; ES; FI; GB; GD; GE; GH; GM; GT; HN; HR; HU; ID; IL; IN; IR; IS; JP; KE; KG; KM; KN; KP; KR; KZ; LA; LB; LG; LI; LU; LV; LY; MA; MG; MK; MN; MW; MX; MY; MZ; NA; NG; NI; NO; NZ; OM; PA; PE; PG; PH; PL; PT; QA; RO; RU; RW; SA; SC; SD; SE; SG; SI; SK; SL; SM; ST; SV; SY; TH; TJ; TM; TN; TR; TT; TZ; UA; UG; US; UZ; VC; VN; ZA; ZM; ZW  
UT DIIDW:201405744  
ER

# Databases of Patents

- Derwent Patent Index
- Espacenet
- Global Patent Index, European Patent Office



# Fields to Questions

	Dynamics	Content	Proximity
Publications	<div>Year</div>	<div>Content</div> <div>Category</div>	<div>Author</div> <div>Organization</div> <div>Funder</div>
Patents	<div>Filing Date</div> <div>Priority Date</div>	<div>Content</div> <div>Classification</div>	<div>Inventor</div> <div>Assignee</div> <div>Patent Office</div>

# Querying the Database

- You need to select a subset of abstracts or records for analysis
- This involves designing and validating an appropriate Boolean query
- A good query has good recall (“all the good records”) as well as good precision (“and only the good records”)

# Tabulating

- Next steps
- Now that you have all this data at your finger tips
- Easy to automatically generate a variety of simple reports
  - Articles per year
  - Most published authors
  - Most patented assignees

# Pre-processing

- Remove stop words, common words like a, and, the, of
- Stemming and lemmatization – converting closely related words to a single common format
- Find and count elements of content – words, keywords, phrases

# Indexing Content



**DESIGN**

# Design, and its Challenges

- Tech mining can be a creative process
- Matching the data to the analysis to the real-world advisory needs of decision-makers a design problem
- However we face two analytic challenges

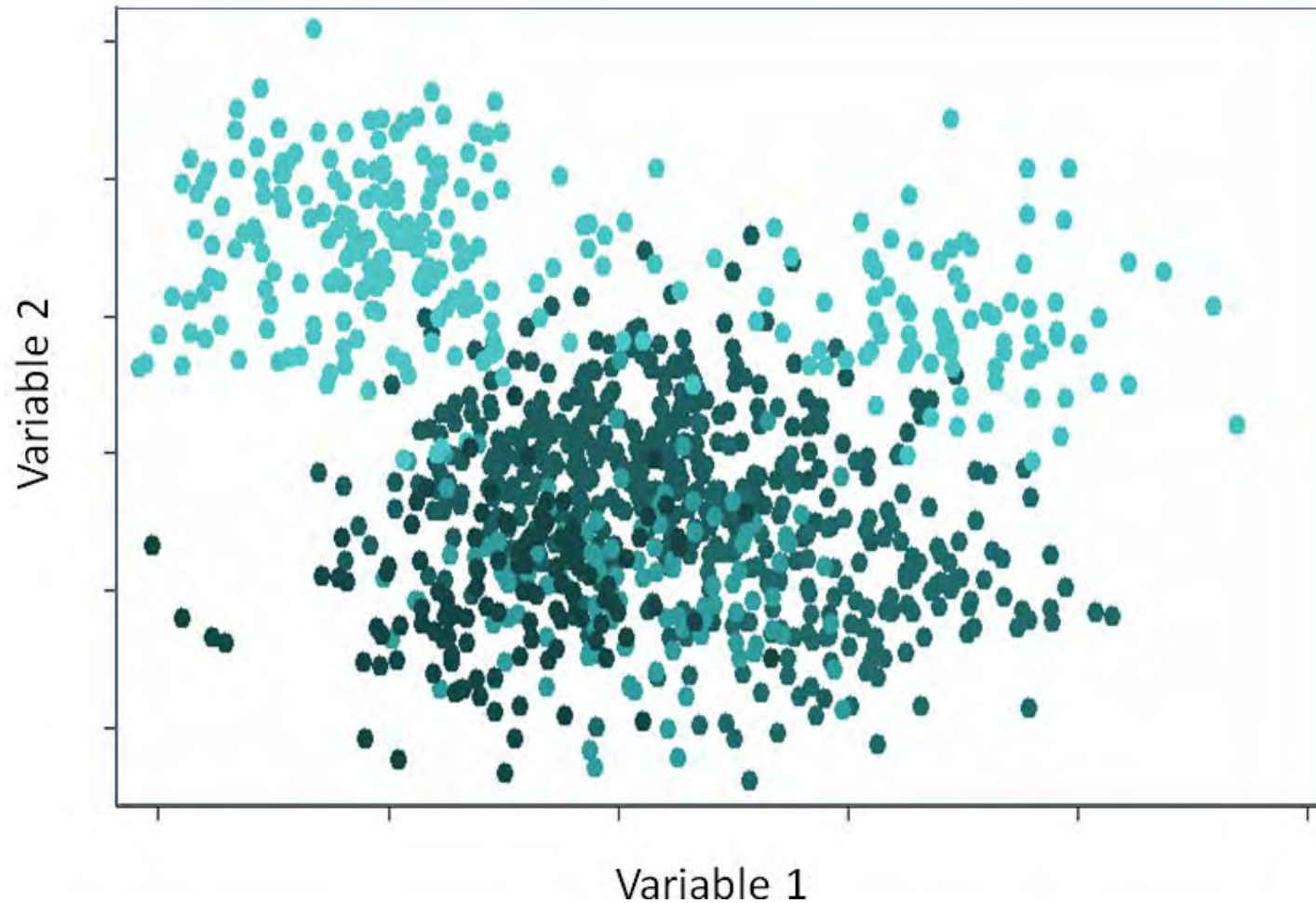
**STRUCTURE + NOISE**



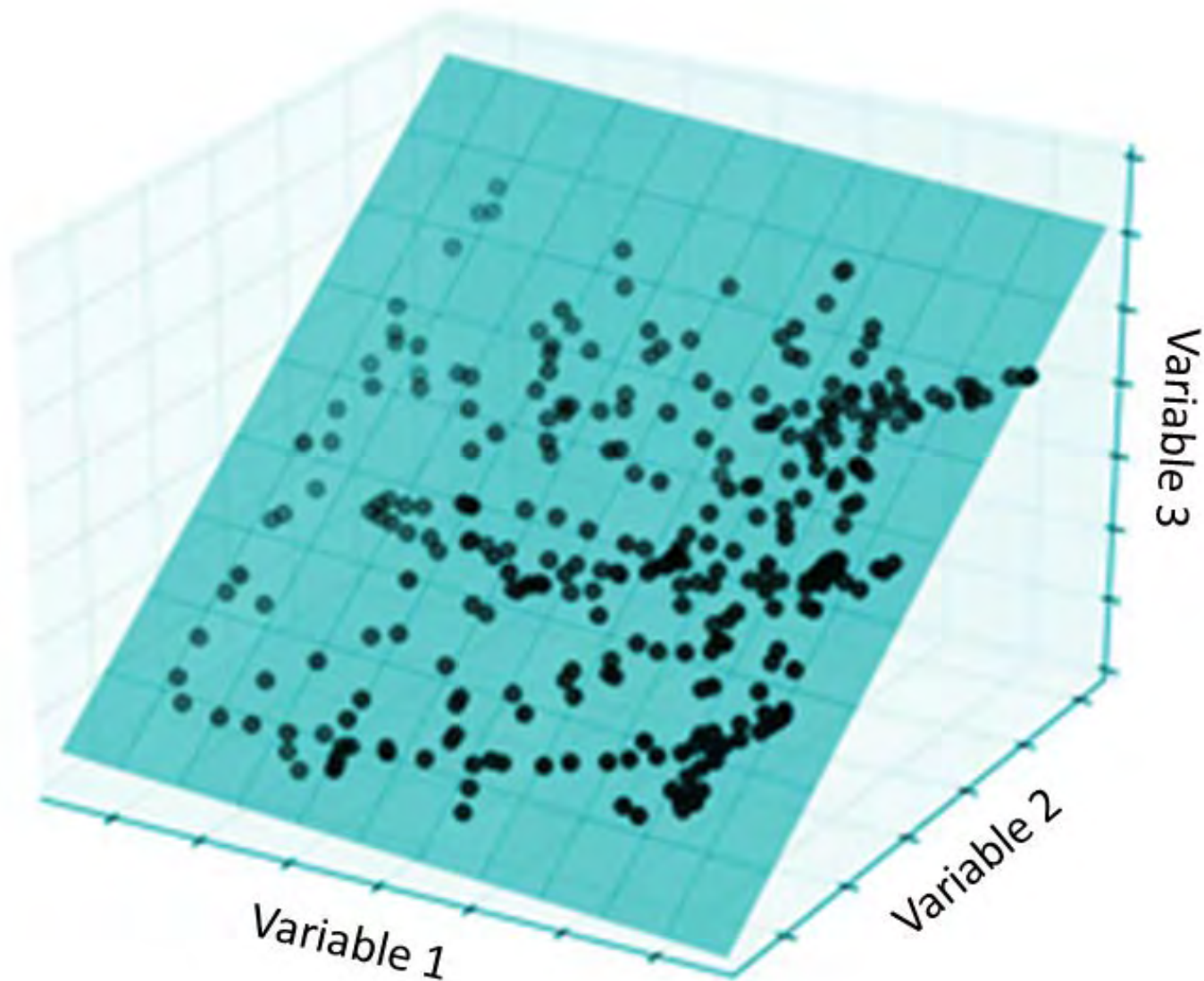
# Challenge #1: Analytical Structure

- The most useful statistical techniques for tech mining are less often taught in graduate schools
- This is because historically our toolboxes have been built for specific purposes – often biometry or manufacturing
- So I'll talk about five analytic techniques about which we need to know more

# Cluster Analysis



# Factor Analysis

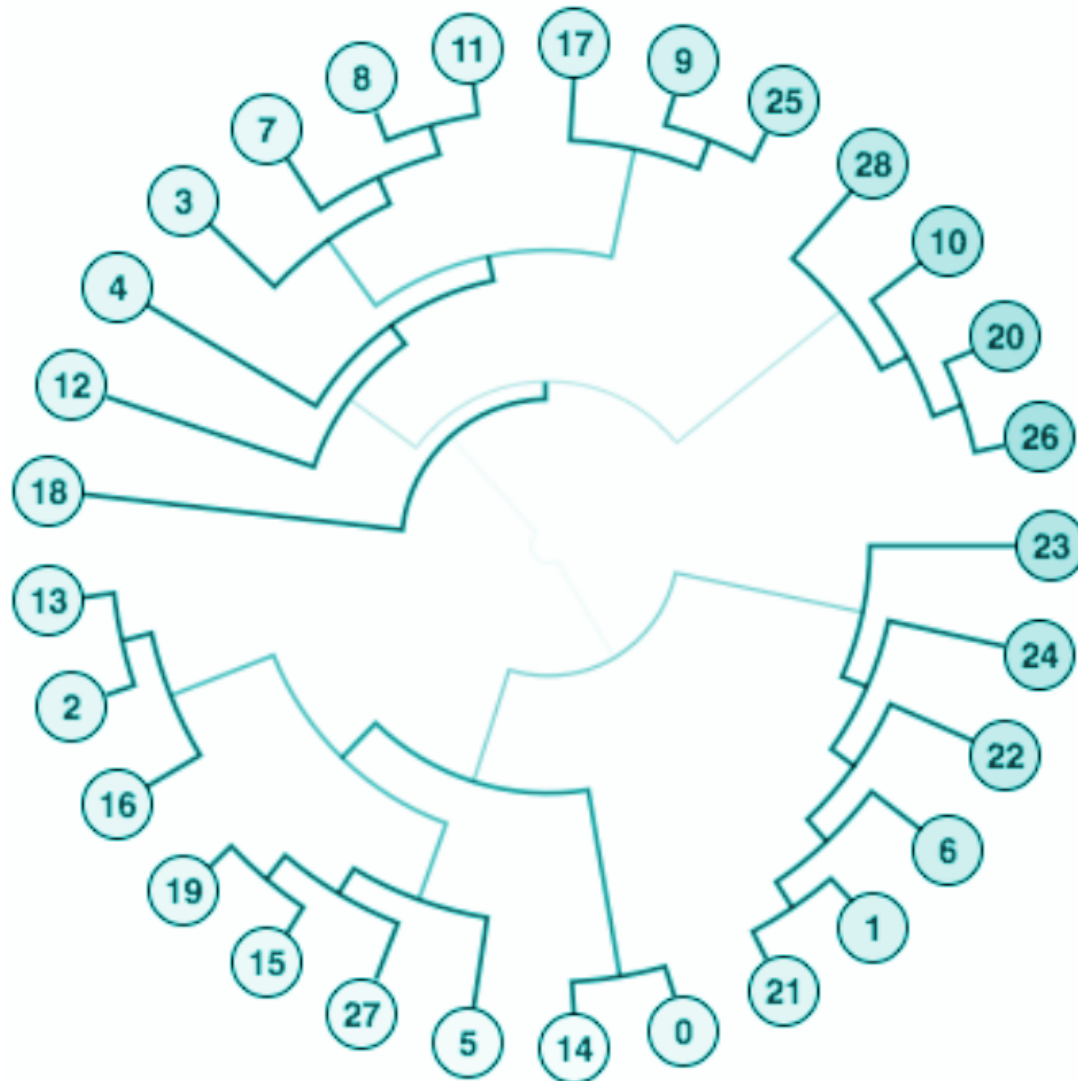




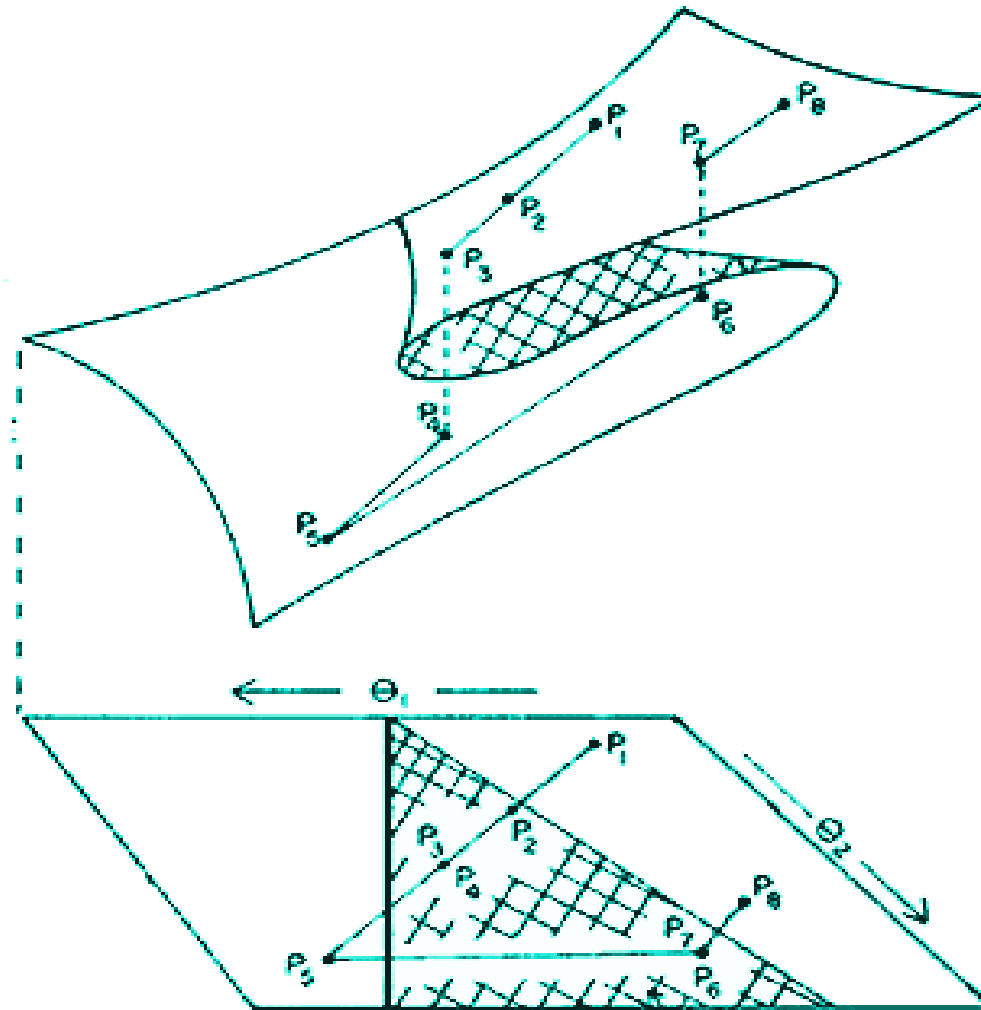
# Network Analysis



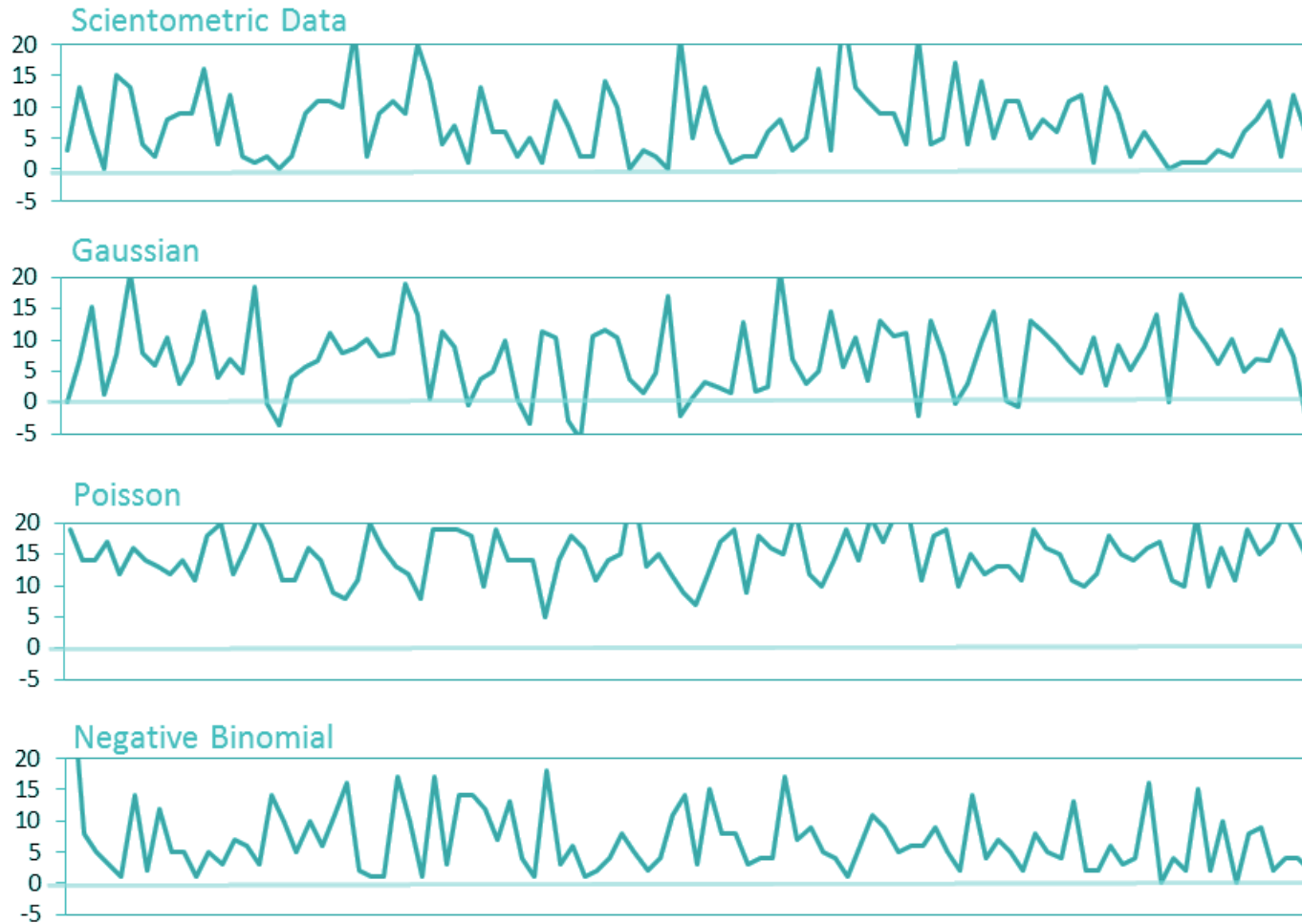
# Tree Analysis



# Trend Analyses



# Challenge #2: The Noise





	Cluster	Factor	Network	Tree	Trend
Content		<input type="checkbox"/>		<input type="checkbox"/>	
Dynamics					<input type="checkbox"/>
Proximity	<input type="checkbox"/>		<input type="checkbox"/>		

# Learning New Techniques

- Learning this quite a challenge
- State-of-the-art statistical techniques at your finger-tips
- Learn by doing!
- Prioritize according to need
- Really good MOOCs available

**CHOICE**

# Supporting Strategic Choices

- Medium-sized interdisciplinary research centre (105 FTE)
- Significant knowledge base (700+ WOS articles)
- Infrastructure Policy and Management
- Government-mandated self-assessment
- University-wide strategic reassessment

# Process Considerations

- Communicating with Dean and Rector
- Working with a short-list of international professors and peers in an interactive workshop
- Working with senior and junior academics
- Self-promoting, critical and analytic
- Research day activities for departments
- Round tables, breakouts for next steps

Journal Name	Quartile
COMPUTER AIDED CHEMICAL ENGINEERING	Q1
DEVELOPMENT AND CHANGE	Q1
ENERGY POLICY	Q1/Q2
ENVIRONMENT AND PLANNING B PLANNING DESIGN	Q3
EUROPEAN JOURNAL OF TRANSPORT AND INFRASTRUCTURE RESEARCH	Q3
EUROPEAN PLANNING STUDIES	Q4
FUTURES	Q2
GOVERNMENT INFORMATION QUARTERLY	Q1
GROUP DECISION AND NEGOTIATION	Q2/Q3
JOURNAL OF CLEANER PRODUCTION	Q1
JOURNAL OF TRANSPORT GEOGRAPHY	Q1
PUBLIC ADMINISTRATION	Q2
SAFETY SCIENCE	Q2
SCIENCE AND ENGINEERING ETHICS	Q1/Q2
TECHNOLOGICAL FORECASTING AND SOCIAL CHANGE	Q1
TELECOMMUNICATIONS POLICY	Q1
TRANSPORT REVIEWS	Q1
TRANSPORTATION PLANNING AND TECHNOLOGY	Q4
TRANSPORTATION RESEARCH PART A POLICY AND PRACTICE	Q1
TRANSPORTATION RESEARCH RECORD	Q3/Q4

# Domestic and International Context

## Domestic and International Peers

### Netherlands

Erasmus & Inst Social Studies	496.9
Delft Univ Technol	478.7
Univ Utrecht	158.3
Eindhoven Univ Technol	84.3
Univ Twente	73.3

### North America

Univ Calif Berkeley	185.2
MIT	161.7
Univ Illinois	139.1
Carnegie Mellon Univ	119.7
Univ Calif Davis	99.4

### Europe

UCL	314.2
Univ Manchester	221.7
Univ London Imperial Coll	218.7
Univ Leeds	207.0
Tech Univ Denmark	156.6

### Asia and Pacific

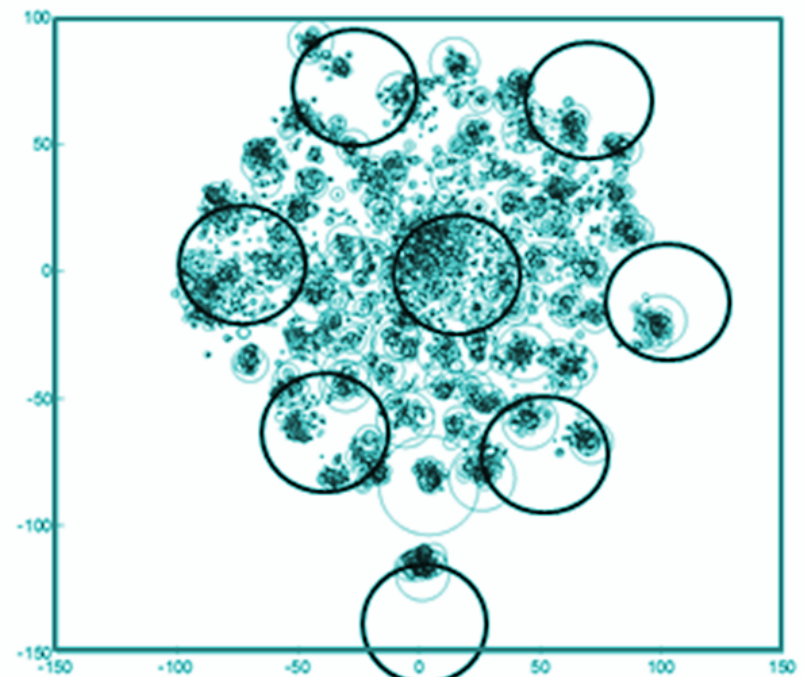
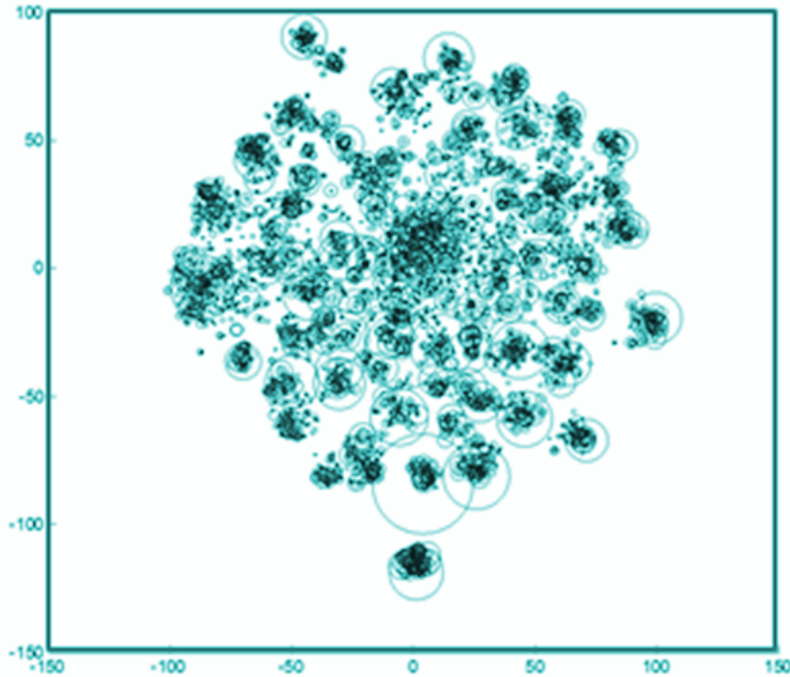
Natl Univ Singapore	157.0
Univ Sydney	105.1
Monash Univ	91.3
Indian Inst Technol	82.0
Tsinghua Univ	81.3

# Comparative Impact of Work

	Citations per Paper			Relative Impact	
	VSC	TPM	Peers	TPM Relative	Peer Relative
One Year Average	0.02	0.11	0.08	4.49	3.37
Two Year Average	0.40	0.50	0.60	1.23	1.49
Four Year Average	1.56	1.84	1.82	1.18	1.17
Eight Year Average	4.22	4.39	5.33	1.04	1.26
Sixteen Year Average	5.43	6.05	7.73	1.11	1.42
Single Peak	9.04	13.18	11.26	1.46	1.25

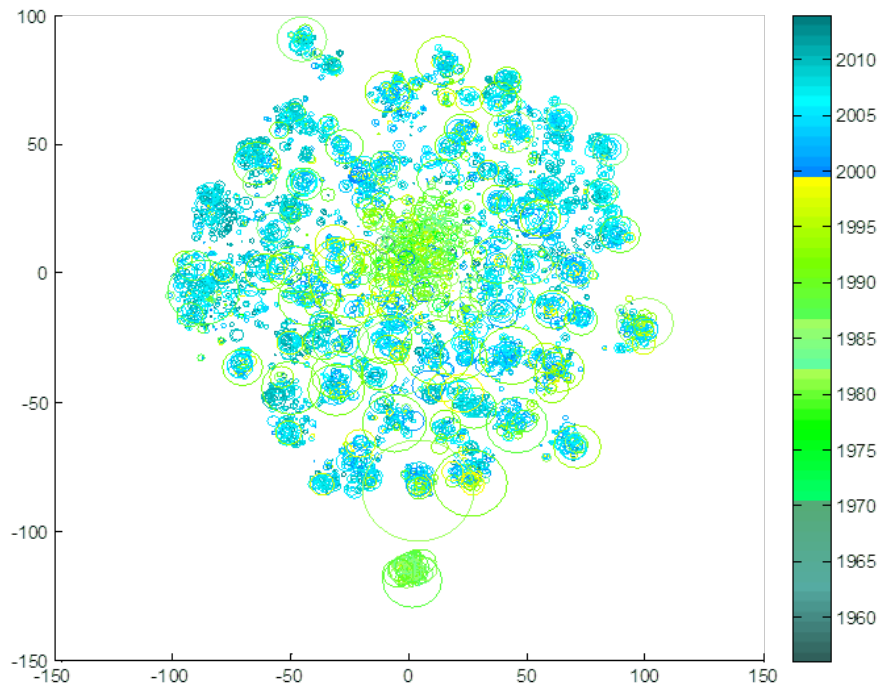


# Content Analysis and Segmentation

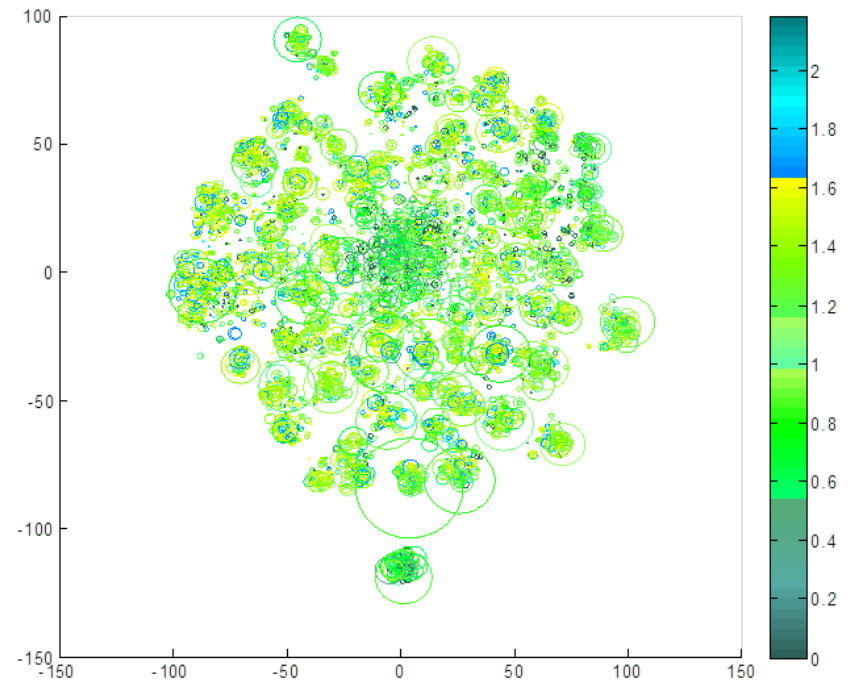


- 50k papers from leading TBM journals analysed
- Clustered using k-means algorithm into 5k clusters
- Further decomposed into 8 super-clusters for interpretation
- Visualized using the t-SNE algorithm (van der Maaten, 2004)

# Science Maps



F2. Average age of clusters  
F2. Average age of clusters



F3. Average citations per cluster. Scale modified to clarify differences in citation.

# Absolute and Relative Publication

Table 6. Absolute Publications by Peer Region and Cluster

	<b>EN</b>	<b>TR</b>	<b>IS</b>	<b>IN</b>	<b>RI</b>	<b>RE</b>	<b>VA</b>	<b>FU</b>
<u>Region/Cluster</u>	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>7</u>	<u>8</u>
Europe	194	357	169	75	127	38	354	22
North America	219	179	80	65	68	42	245	14
Asia and Pacific	184	140	46	33	69	36	143	7
Netherlands	91	105	45	164	98	23	446	37
Delft	97	108	106	60	106	6	142	10

Table 7. Relative Publications by Peer Region and Cluster

<u>Region/Cluster</u>	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>7</u>	<u>8</u>
Europe	82%	137%	140%	72%	91%	111%	87%	106%
North America	138%	103%	95%	87%	70%	159%	93%	87%
Asia and Pacific	172%	120%	78%	63%	102%	177%	84%	58%
Netherlands	58%	61%	54%	218%	101%	89%	169%	219%
Delft	94%	96%	182%	116%	160%	34%	86%	81%

**IMPLEMENTATION**

# Implementation Choices

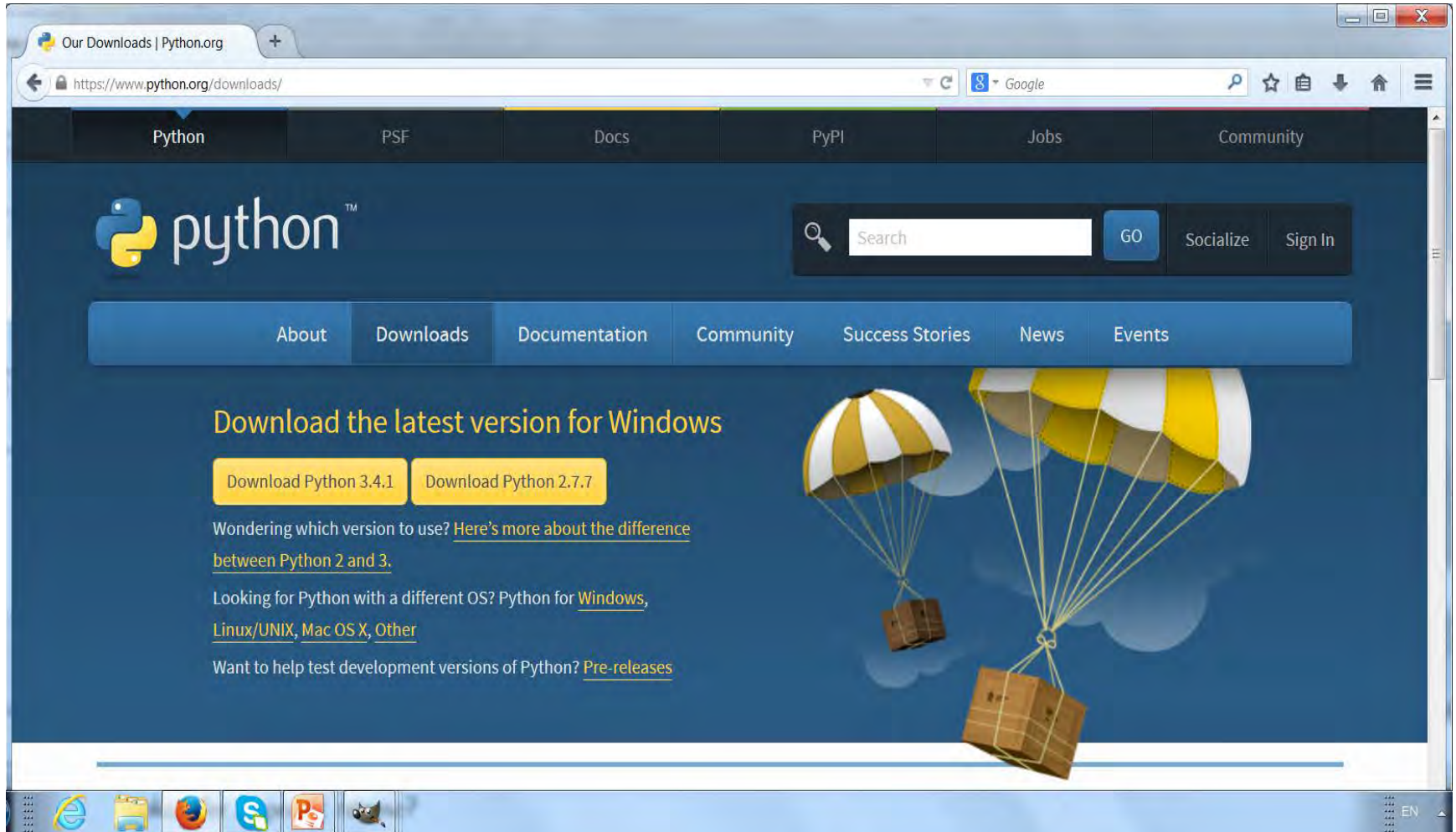
- Proprietary Packages
- Scientific Computing and Statistical Packages
- Open Source Software and Scripting
- Tech Mining is a pragmatic field!

	Proprietary	Scientific Computing	Open Source
Advantages	<ul style="list-style-type: none"> <li>• Standardized</li> <li>• Supports range of skill</li> <li>• Easier to maintain</li> </ul>	<ul style="list-style-type: none"> <li>• Innovative</li> </ul>	<ul style="list-style-type: none"> <li>• Affordable</li> <li>• Innovative</li> <li>• Easier to customize</li> <li>• Internet native</li> </ul>
Disadvantages	<ul style="list-style-type: none"> <li>• Costly</li> <li>• Slower</li> <li>• Lags behind in features</li> </ul>	<ul style="list-style-type: none"> <li>• No text handling</li> <li>• Skill intensive</li> <li>• Memory intensive</li> </ul>	<ul style="list-style-type: none"> <li>• Time intensive</li> <li>• Skill intensive</li> </ul>

# Python

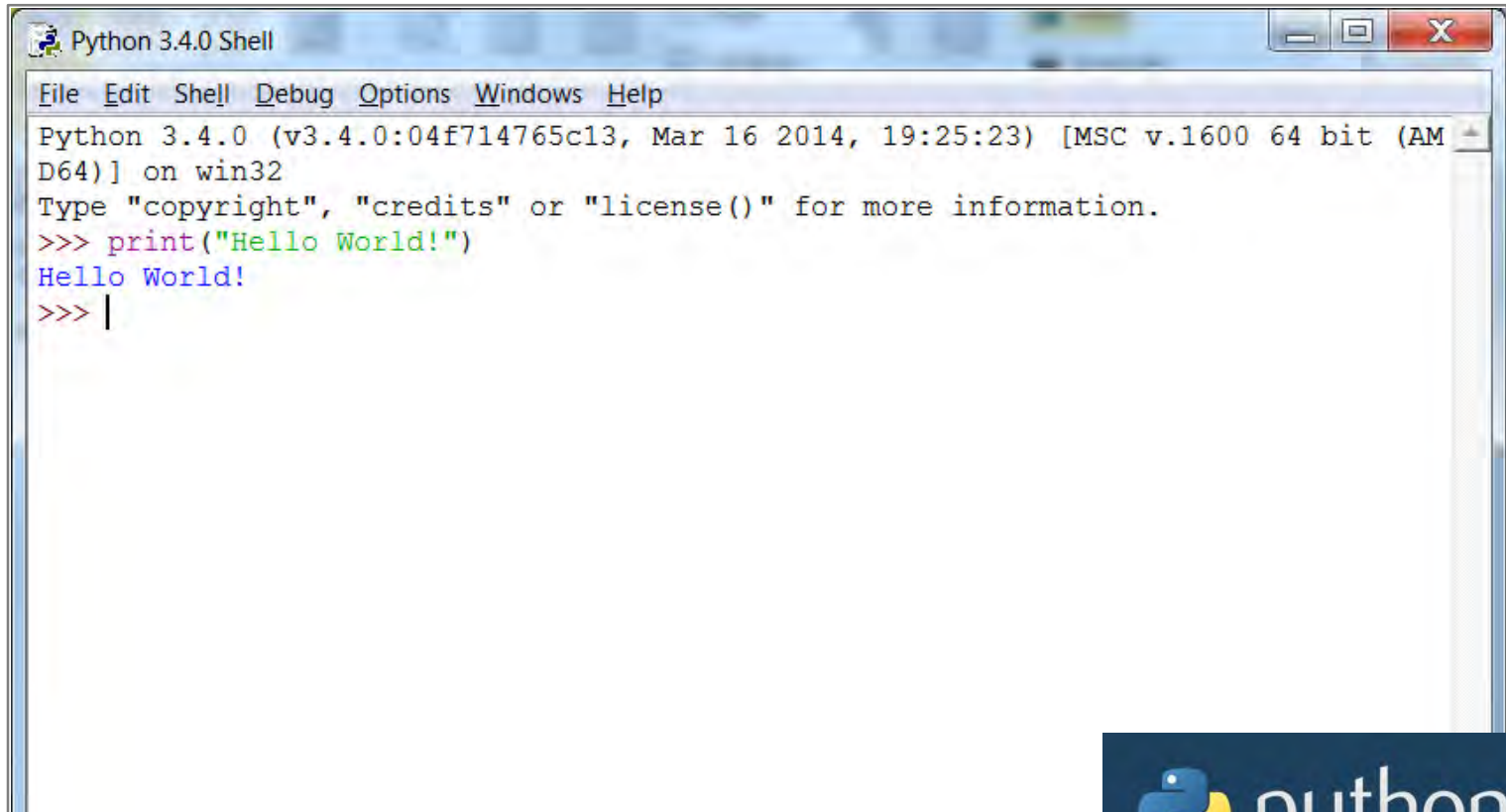
- Widely used
- High-level
- General purpose
- Scripting
- First appeared in 1991; 23 years ago

# Downloading Python





# IDLE: Programming in Python



The screenshot shows a window titled "Python 3.4.0 Shell". The menu bar includes "File", "Edit", "Shell", "Debug", "Options", "Windows", and "Help". The main text area displays the following content:

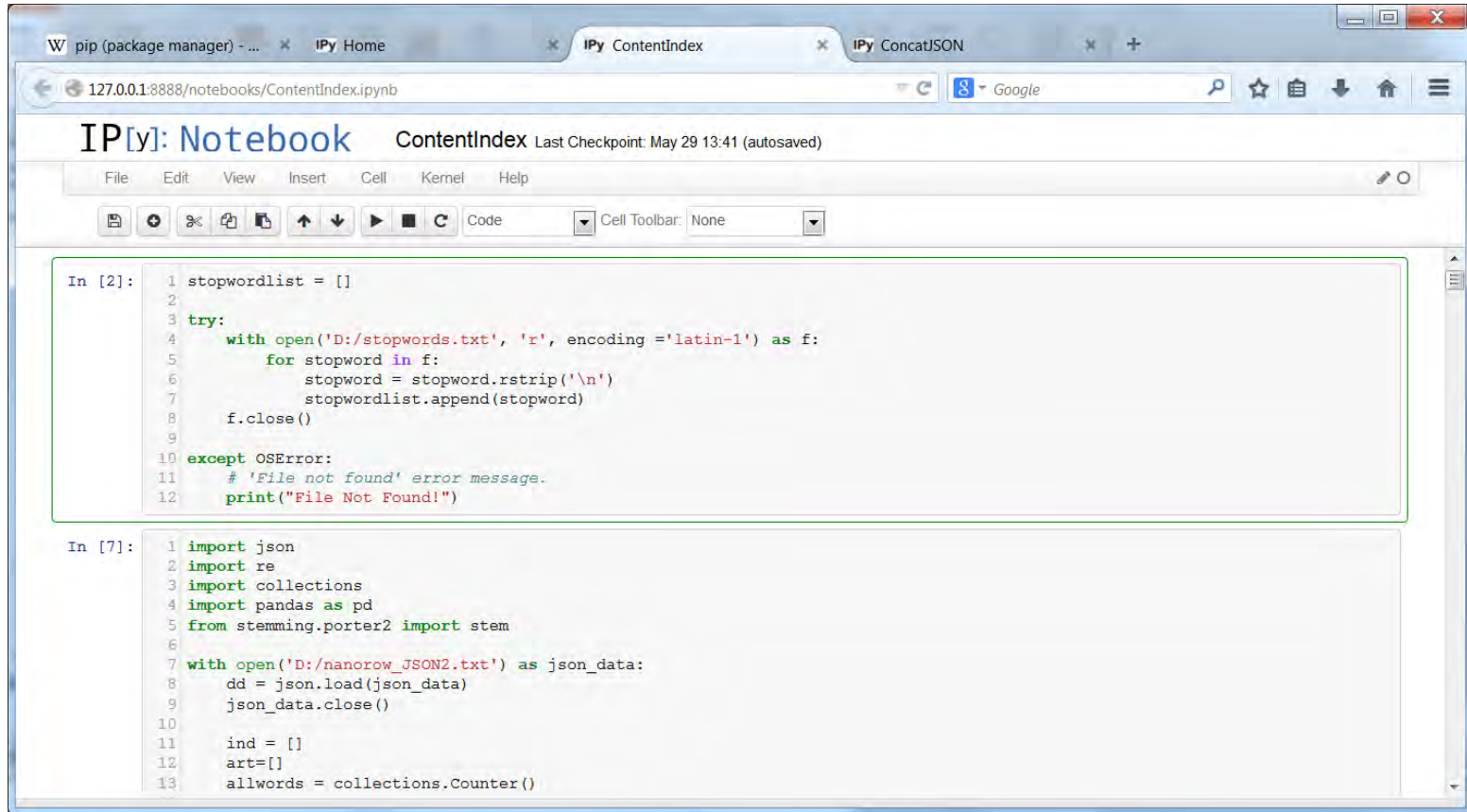
```
Python 3.4.0 (v3.4.0:04f714765c13, Mar 16 2014, 19:25:23) [MSC v.1600 64 bit (AMD64)] on win32
Type "copyright", "credits" or "license()" for more information.
>>> print("Hello World!")
Hello World!
>>> |
```



# Expanded Python Capabilities

- Many extended capabilities added to Python by third party developers
- Text Processing, Reporting, Scientific Computing, Visualization
- iPython notebook

# iPython Notebook



The screenshot displays the iPython Notebook web interface in a browser. The address bar shows the URL `127.0.0.1:8888/notebooks/ContentIndex.ipynb`. The notebook title is `ContentIndex`, and it indicates the last checkpoint was on May 29 at 13:41 (autosaved). The interface includes a menu bar (File, Edit, View, Insert, Cell, Kernel, Help) and a toolbar with icons for file operations and execution. Two code cells are visible:

```
In [2]: 1 stopwordlist = []
2
3 try:
4     with open('D:/stopwords.txt', 'r', encoding='latin-1') as f:
5         for stopword in f:
6             stopword = stopword.rstrip('\n')
7             stopwordlist.append(stopword)
8         f.close()
9
10 except OSError:
11     # 'File not found' error message.
12     print("File Not Found!")
```

```
In [7]: 1 import json
2 import re
3 import collections
4 import pandas as pd
5 from stemming.porter2 import stem
6
7 with open('D:/nanorow_JSON2.txt') as json_data:
8     dd = json.load(json_data)
9     json_data.close()
10
11 ind = []
12 art=[]
13 allwords = collections.Counter()
```





# New Book, Momentum Press

## TABLE OF CONTENTS

- Chapter 1 – Tech Mining Using Open Source Tools
- Chapter 2 – Downloading and Processing the Scientific Article
- Chapter 3 – Processing and Storing the Scientific Article
- Chapter 4 – Parsing the Patent
- Chapter 5 – Parsing the Web and Social Media
- Chapter 6 – Creating Bots, Spiders and Crawlers
- Chapter 7 – Tabulating and Filtering Records
- Chapter 8 - Indexing the Content
- Chapter 9 – Creating Clusters of Related Content
- Chapter 10 – Creating Factors of Related Content
- Chapter 11 – Creating Science and Technology Maps
- Chapter 12 – Creating Networks of Content
- Chapter 13 – Analyzing Dynamic Content
- Chapter 14 - Analyzing Geographic and Institutional Data

# Example Scripts

## PARTIAL LIST OF FIGURES, EXAMPLES, AND TABLES

### Examples

Example 3.1: Example for Loading Column Structured Text

Example 3.2: Printing the Data Structure

Example 3.3: Example of Parsing Plain Text Articles

Example 3.4: Example of Growing an Existing Corpus of Text

Example 3.5: Read-in Supplementary Information

Example 4.1: Inspecting Espacenet Patent Records

Example 4.2: Cleaning and Parsing the Espacenet Format

Example 4.3: Unpacking Global Patent Index Records

Example 4.4: Function for Printing XML Elements

Example 8.1: Extracting Content from a Record

Example 8.2: Indexing Content from a Record

```
1 import json
2
3 linecount = 0
4 try:
5     # read using with to close filehandle when complete
6     # use 'latin-1' encoding as a safe bet
7     with open('D:/nano_column.txt', 'r', encoding='latin-1') as f:
8         d = {} #This is the dictionary for a single record
9         dd = {} #This is the dictionary for all the articles
10
11         #read the line
12         for line in f:
13             line=line.rstrip('\n')
14             fields=line.split('\t')
15
16             if (linecount==0):
17                 headers = fields
18             else:
19                 line = line.rstrip('\t')
20                 fields=line.split('\t')
21                 article_id=fields[-1]
22                 pair = zip(headers,fields)
23                 d=dict(pair)
24                 dd[article_id]=[d]
25
26             linecount += 1
27
28 except OSError:
29     # 'File not found' error message.
30     print("File Not Found!")
31
32 with open('D:/nanocolumn_JSON.txt', 'w') as g:
33     json.dump(dd,g)
```

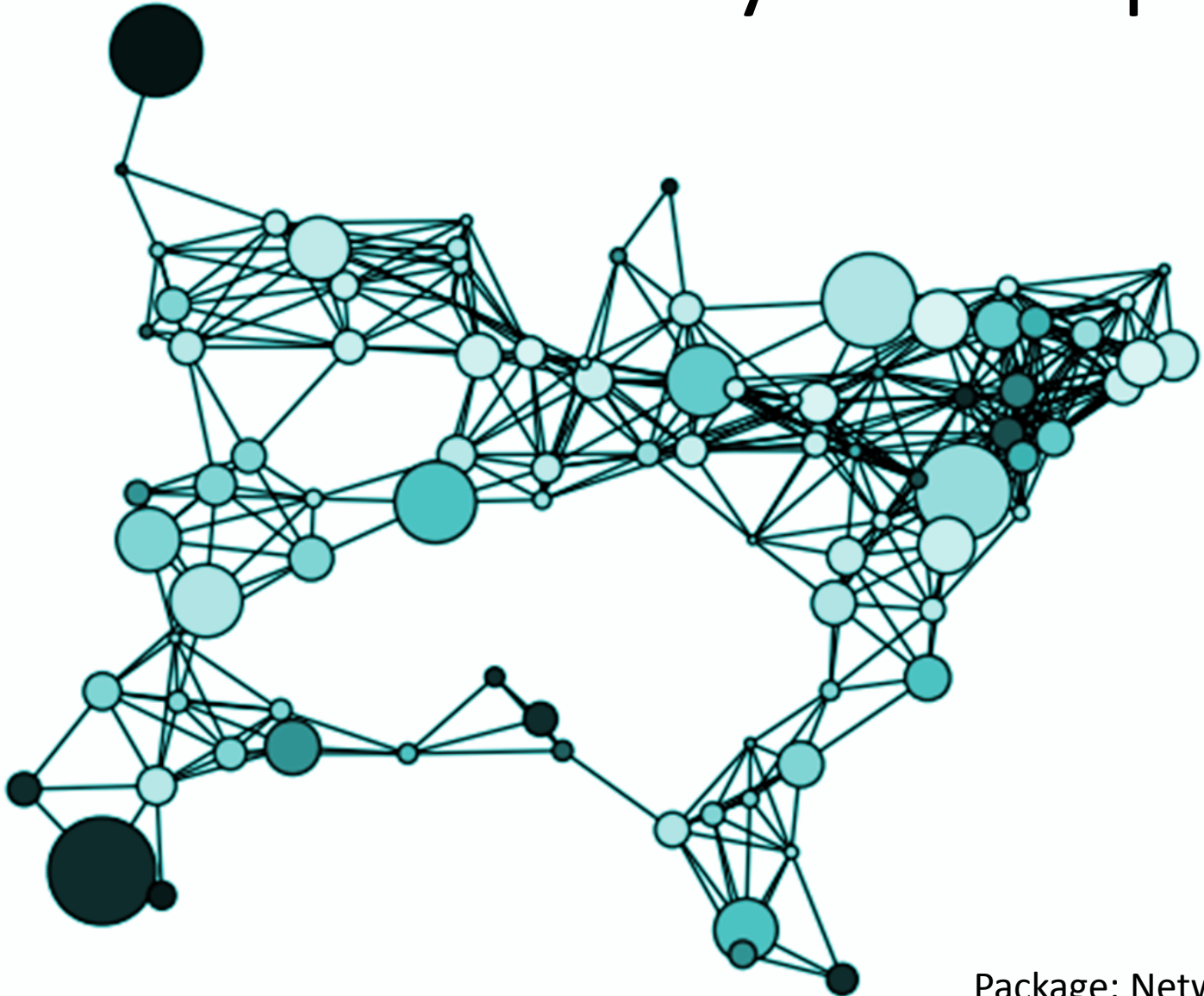
# Python Word Cloud Example



Package: wordcloud  
<https://github.com/amueller>

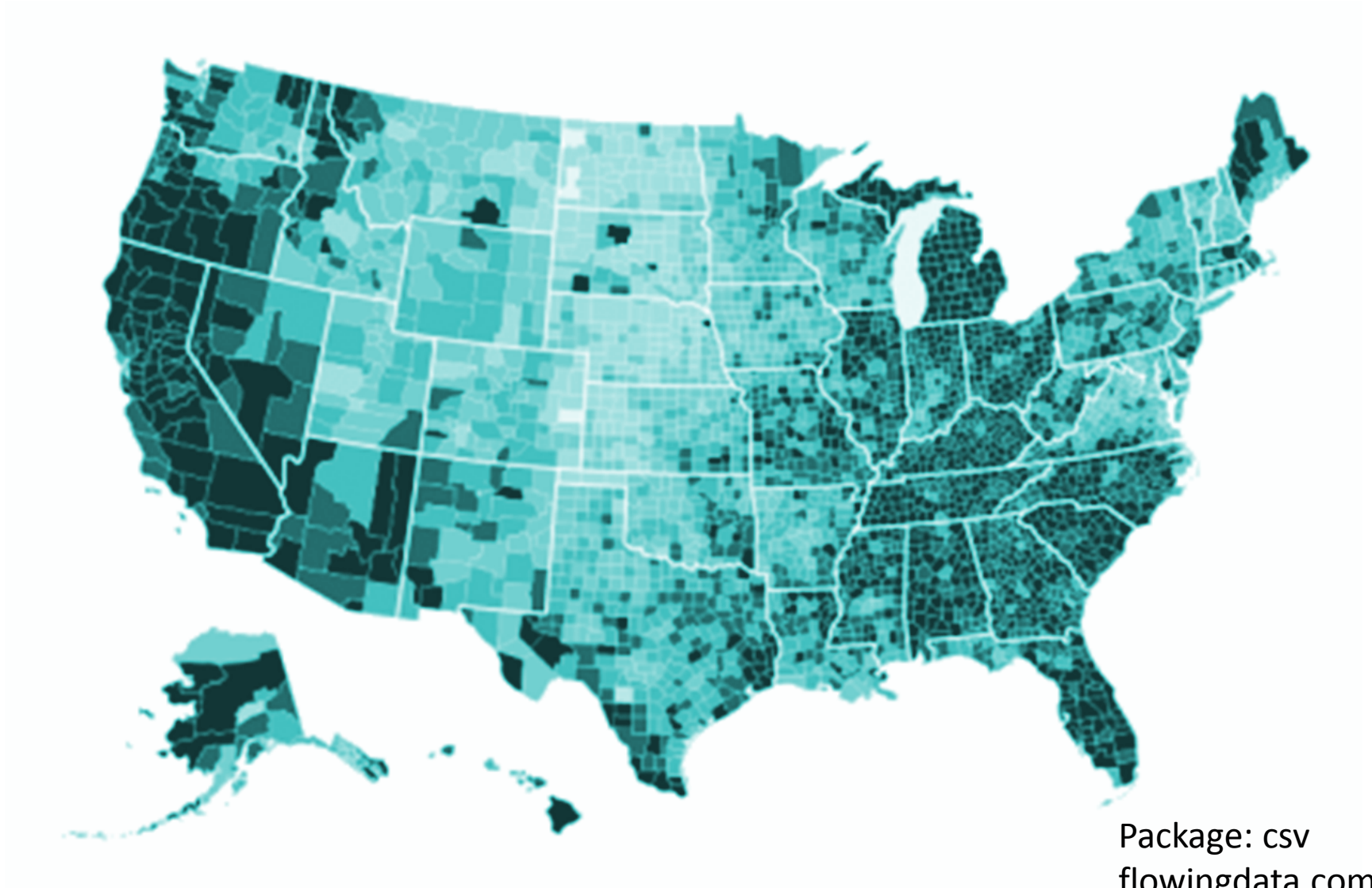


# Python Network Analysis Example



Package: NetworkX

# Python Choropleth Example



# Python Arcs and Globes



Package: d3py  
[github.com/dwtkns](https://github.com/dwtkns)

# Conclusions

Email: [s.cunningham@tudelft.nl](mailto:s.cunningham@tudelft.nl)

Twitter: @DelftScott