The data wars: moving from management information to data driven intelligence

UKSG: April 2019

Ken Chad

Ken Chad Consulting Ltd

Twitter @kenchad

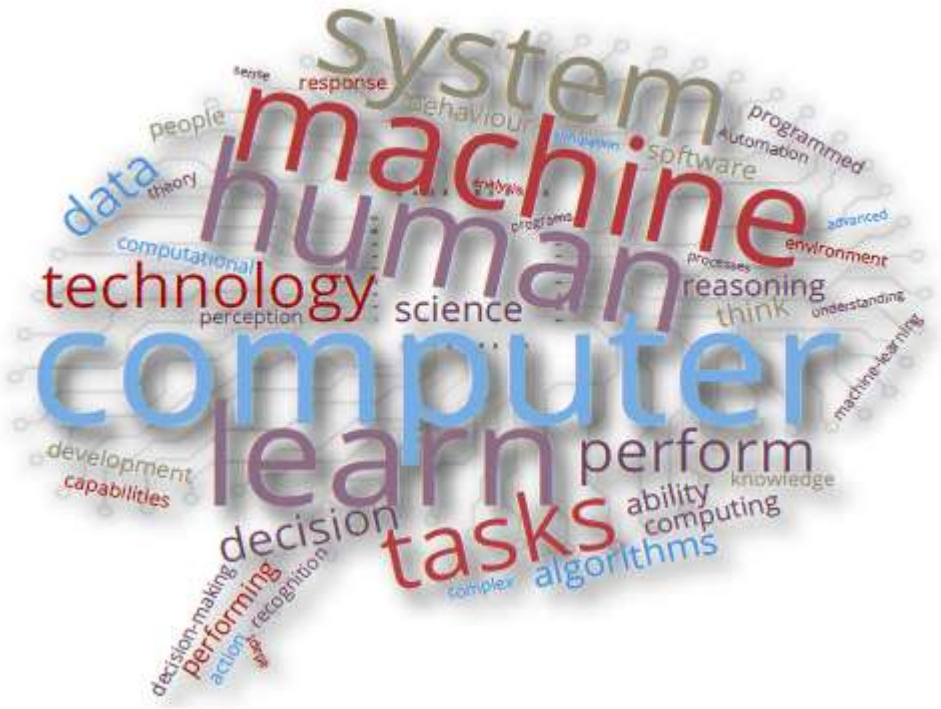
ken@kenchadconsulting.com

Tel: +44 (0)7788 727 845

www.kenchadconsulting.com

Setting your expectations.....

My contribution today should be treated as "food-for-thought"



Data

"The world's most valuable resource is no longer oil, but data. Smartphones and the internet have made data abundant, ubiquitous and far more valuable"

The world's most valuable resource is no longer oil, but data. The Economist. 6 May 2017
http://www.economist.com/news/leaders/21721656-data-economy-demands-new-approach-antitrust-rules-worlds-most-valuable-resource

Libraries and data

Why are libraries interested? -Trends

"the library has an integral role in the university's mission of research and teaching, not just a supporting one."

UK Scholarly Reading and the Value of Library Resources: library may have a correlation to the overall strength Summary Results of the Study Conducted Spring 2011. Carol Tenopir and Rachel Volentine. Jisc Collections. February 2012

Why are libraries interested? -Trends


JOURNAL OF LIBRARY INNOVATION

Volume 6, Issue 1, 2015

Article

Academic Libraries and Innovation: A Literature Review

Curtis Brundy

Academic libraries are facing times of unprecedented challenge and unparalleled change. **Innovation has moved from a consideration to a necessity**.

http://www.libraryinnovation.org/article/view/420/625

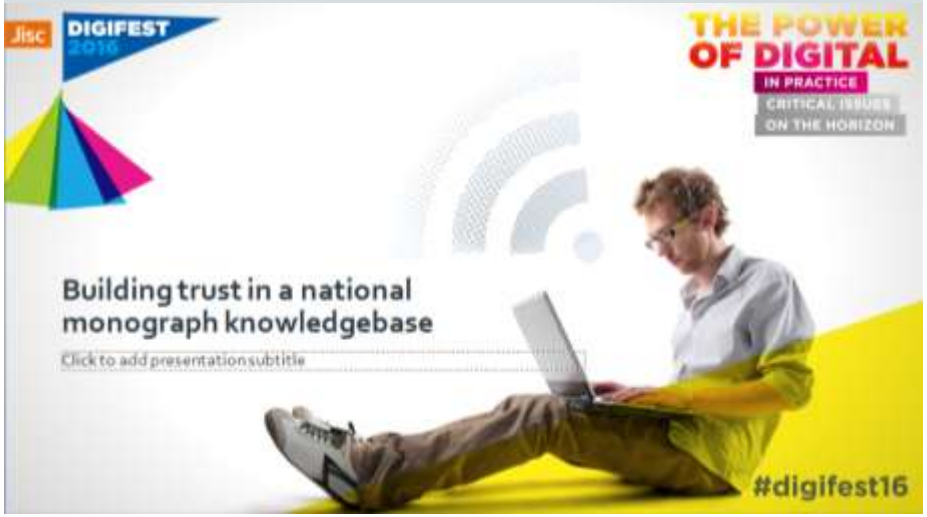
Why are libraries interested? -Trends

In challenging times, the requirement to make decisions (particularly in relation to resources spend) that are **data-driven and evidence-based** have become increasingly significant.

Back to basics: the importance of library collections. By Steve Rose. Sconul. Focus 70 [Editorial] 2017
https://www.sconul.ac.uk/sites/default/files/documents/3.Editorial.pdf

Why are libraries interested? -Trends

What problems are we trying to solve?



"Libraries want to make **data-driven decisions about the management of their digital and print book collections** but the data that is currently available does not allow them to do this with confidence"

http://www.slideshare.net/JISC/building-trust-in-a-njiational-monograph-knowledgebase-sc-digifest-2016

Why are libraries interested? -Trends

Trend 1 :Rising interest in big data, data science and AI in general

Trend 2: Library systems are becoming more open and more capable at analytics

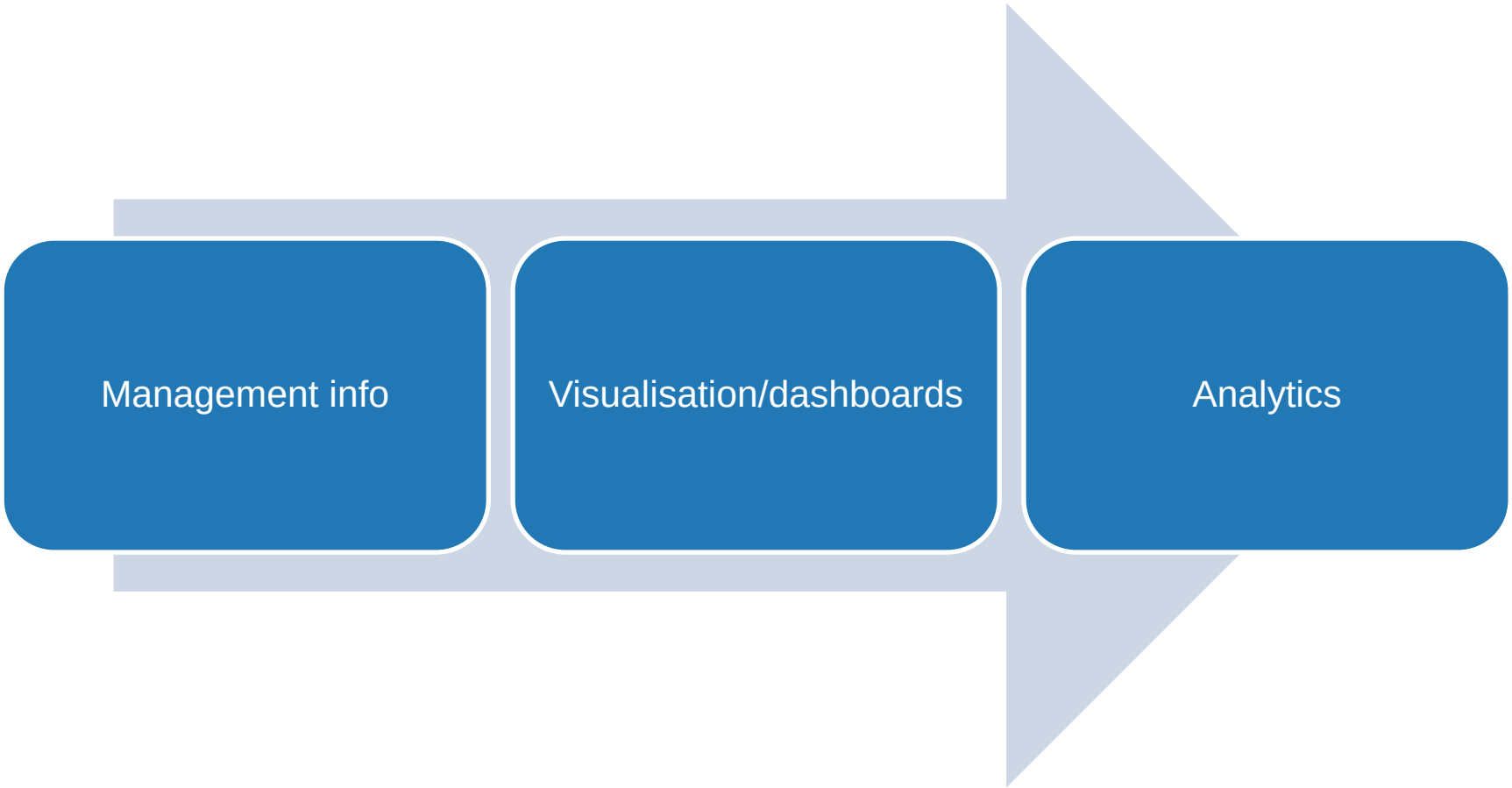
Trend 3 : Assessment and increasing demand to show value are hot trends

Trend 4 : Rising interest in learning analytics

Trend 5 : Increasing academic focus on managing research data provides synergy

5 reasons why library analytics is on the rise. By Aaron Tay. Musings about Librarianship [blog] 18th November 2016 http://musingsaboutlibrarianship.blogspot.com/2016/11/5-reasons-why-library-analytics-is-on.html

PART 1 : From management information to analytics



Management information –reports from the ILS

Out-of-the-Box Reports

Report	Description
Acquisitions (based on Funds Expenditures SA)	
Detailed Expenditures	Displays fund, title, and price information for items.
Expenditure Per Acquisition Method	Displays annual expenditures per acquisition method.
Expenditure Per Acquisition Method – Annual Trend	Displays a line chart of annual expenditures per acquisition method.
Expenditure Per Acquisition Method - Top 10	Displays a bar chart of the top 10 expenditures per acquisition method.
Expenditure Per Classifications	Displays annual expenditure per classification.
Expenditure Per Classifications - Top 10	Displays a bar chart of the top 10 annual expenditures per classification.
Expenditure Per Fund Ledger Code	Displays annual expenditures per fund ledger code.
Expenditure Per Fund Ledger Code - Top 10	Displays a bar chart of the top 10 expenditures per fund ledger code.

https://knowledge.exlibrisgroup.com/Alma/Product_Documentation/010Alma_Online_Help_(English)/080Analytics/060Out-of-the-Box-Reports

Bringing library related data together –dashboards/visualisation

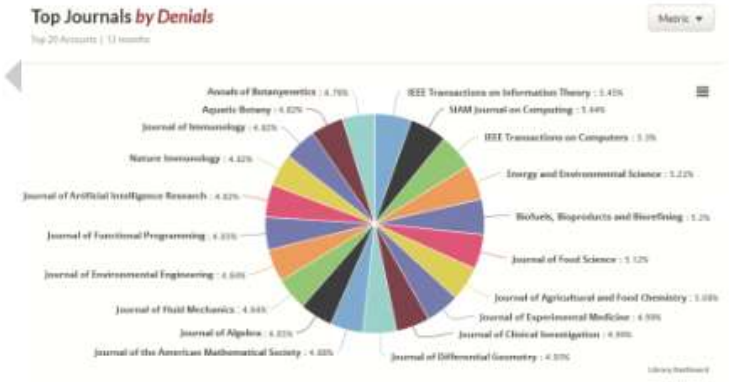
RedLink

FOR LIBRARIES FOR PUBLISHERS FOR BOTH WHY REDLINK? COMPANY

MAKE MORE DATA-DRIVEN DECISIONS

Integrate and organize data from dozens of sources covering journals, databases, eBooks, and bundles, all in one place so you and your staff can make more informed spending and collection decisions.

". . . helps us to efficiently allocate resources to meet user needs and demands." – **Scott Anderson, Millersville University**



https://redlink.com/library-dashboard/

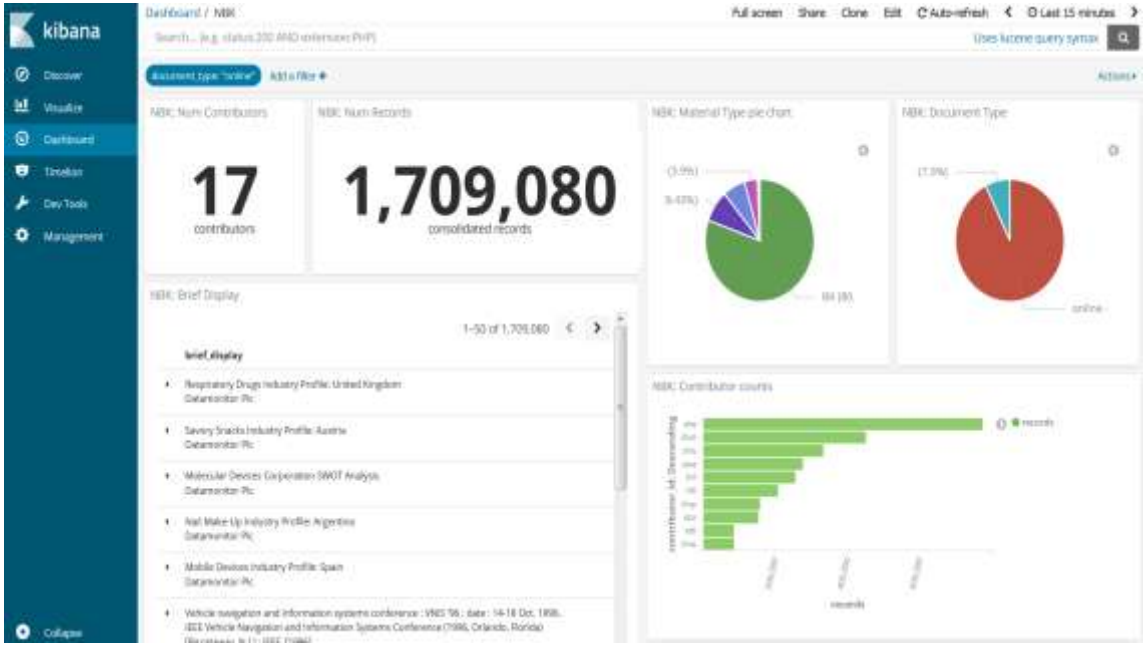
Shared data initiatives

https://jusp.jisc.ac.uk/

Shared initiatives: The NBK

During 2019 NBK will work to deliver improved collection analysis tools using Elastic Search:

- Faceted search options, such as format, date, material type: enabling better collection analysis
- More data visualisation and benchmarking tools
- Data quality indicators
- Institutional data analysis tools:



Trying to get more useful results…..

Usage data on their own…give libraries and publishers very little insight into how content is being used or how much it is being looked at.

In spite of the huge amount of data that are now available to libraries, it feels as if **little progress has been made in developing metrics that may give an indication of how resources are being used** and the extent to which library users value the resources provided. These perceived shortcomings in conventional usage data led Nottingham Trent University and Alexander Street to partner in piloting an in-depth view of **analytics, demonstrating user engagement and impact of use**.

Project COUNTER …..provides consistency in reporting, but it does not reveal which individual titles or subjects are used or for what purpose.

Adey, H., & Eastman-Mullins, A. (2017). User engagement analytics case study: how customer behaviour can drive intelligent library decision making. Insights, 30(3), 138–147. DOI: http://doi.org/10.1629/uksg.387 https://insights.uksg.org/articles/10.1629/uksg.387/

Hierarchy of analytics use in libraries

Level 1 - Any analysis done is library function specific. Typically ad-hoc analytics but there might be dashboard systems created for only one specific area (e.g. collection dashboard for Alma or web dashboard for Google analytics)

Level 2 - A centralised library wide dashboard is created covering most functional areas in the library

Level 3 - Library "shows value" runs correlation studies etc

Level 4 - Library ventures into predictive analytics or learning analytics

By the time you reach level 4, it would be almost impossible fo library to go it alone.

5 reasons why library analytics is on the rise. Aaron Tay. Musings about librarianship [blog] http://musingsaboutlibrarianship.blogspot.com/2016/11/5-reasons-why-library-analytics-is-on.html

Library related data only gets you so far

"Many firms are not able to exploit their data in the most effective manner to drive the best insights when the data is stored across different silos." (PwC Global CEO Survey)

Data driven businesses race ahead but change doesn't come easily. By Nick Bouch. In Future of Data. Raconteur. 26 March 2019

Beyond the library - Learning analytics

"Every time a student interacts with their university – be that going to the library, logging into their virtual learning environment or submitting assessments online – they leave behind a digital footprint. **Learning analytics is the process of using this data to improve learning and teaching"**

Learning analytics in higher education: A review of UK and international practice. Read our updated briefing on learning analytics and student success from January 2017. By Niall Sclater, Alice Peasgood & Joel Mullen. Jisc. 2016. https://www.jisc.ac.uk/reports/learning-analytics-in-higher-education

Northumbria University's approach to the **utilisation of Educational Analytics is directly linked to the University Strategy**

In the future, the use of Educational Analytics may be extended to personalised earning paths, adaptive learning, personalised feedback, visualisations of study journey, intelligent e-tutoring, intelligent peer support, etc.

Furthermore, new technological innovations might allow for more targeted, measured approaches.

https://www.northumbria.ac.uk/-/media/corporate-website/new-sitecore-gallery/services/academic-registry/documents/qte/student-engagement/ethical-use-of-student-data-for-educational-analytics.pdf?la=en&hash=EEB8CF87D03669F66A935ECEA17D084F05947832

Northumbria University NEWCASTLE

Using Student Data for Educational Analytics

The following data, which is currently captured by the University, is initially in scope
for Educational Analytics:

- personal information provided by the student at registration
- student level study records held by the University including assessment marks
- details of a student's assigned Personal Tutor system-generated data from Blackboard, such as the date and frequency of accessing pages
- student attendance data
- library borrowing logs
- smart card activity log on Campus
- Northumbria gym membership

This data will be used in line with the University's Student and Applicant Privacy Notice.

https://www.northumbria.ac.uk/-/media/corporate-website/new-sitecore-gallery/services/academic-registry/documents/qte/student-engagement/ethical-use-of-student-data-for-educational-analytics.pdf?la=en&hash=EEB8CF87D03669F66A935ECEA17D084F05947832

Data could save lives

The Guardian

The bereaved father who says data could save lives

Murray is proposing that universities introduce a system that would automatically pull together disparate data – from schools, the student, attendance, assessment and library access. Multiple red flags would trigger an early warning alert and a meeting with a student support professional.

Student suicides: the bereaved father who says data could save lives
Guardian Education 8 October 2018
https://www.theguardian.com/education/2018/oct/08/student-suicides-the-bereaved-father-who-says-data-could-save-live

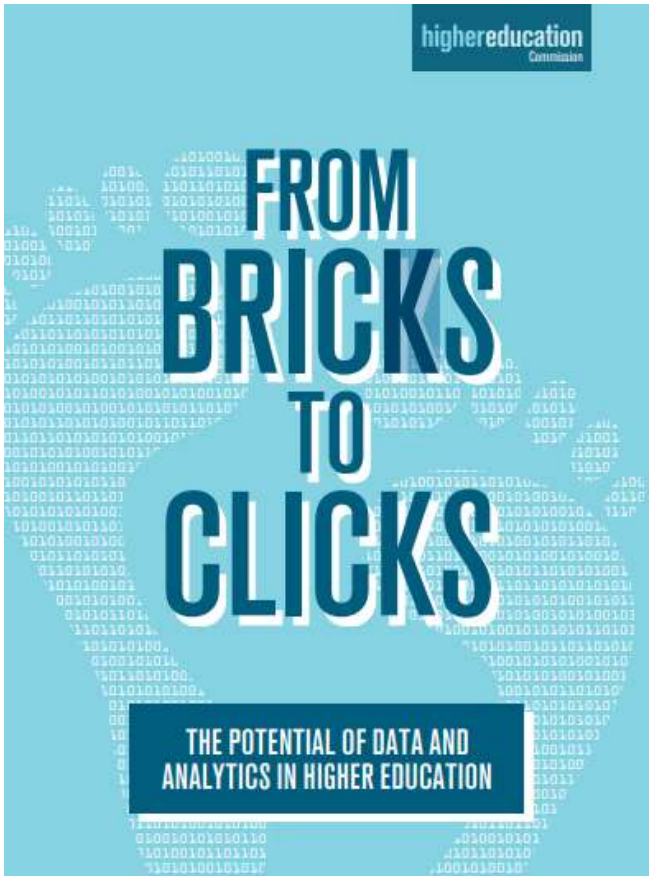
We're working in collaboration to build a learning analytics service for the sector. There are over 50 universities and colleges signed up to the initial phases of the implementation.

What we are making

1. A basic learning analytics solution

This will include **everything you require to track student learning activity** so that you **can improve retention and attainment**. It will include an app for students to allow them to maximise their learning potential by tracking their learning activity.

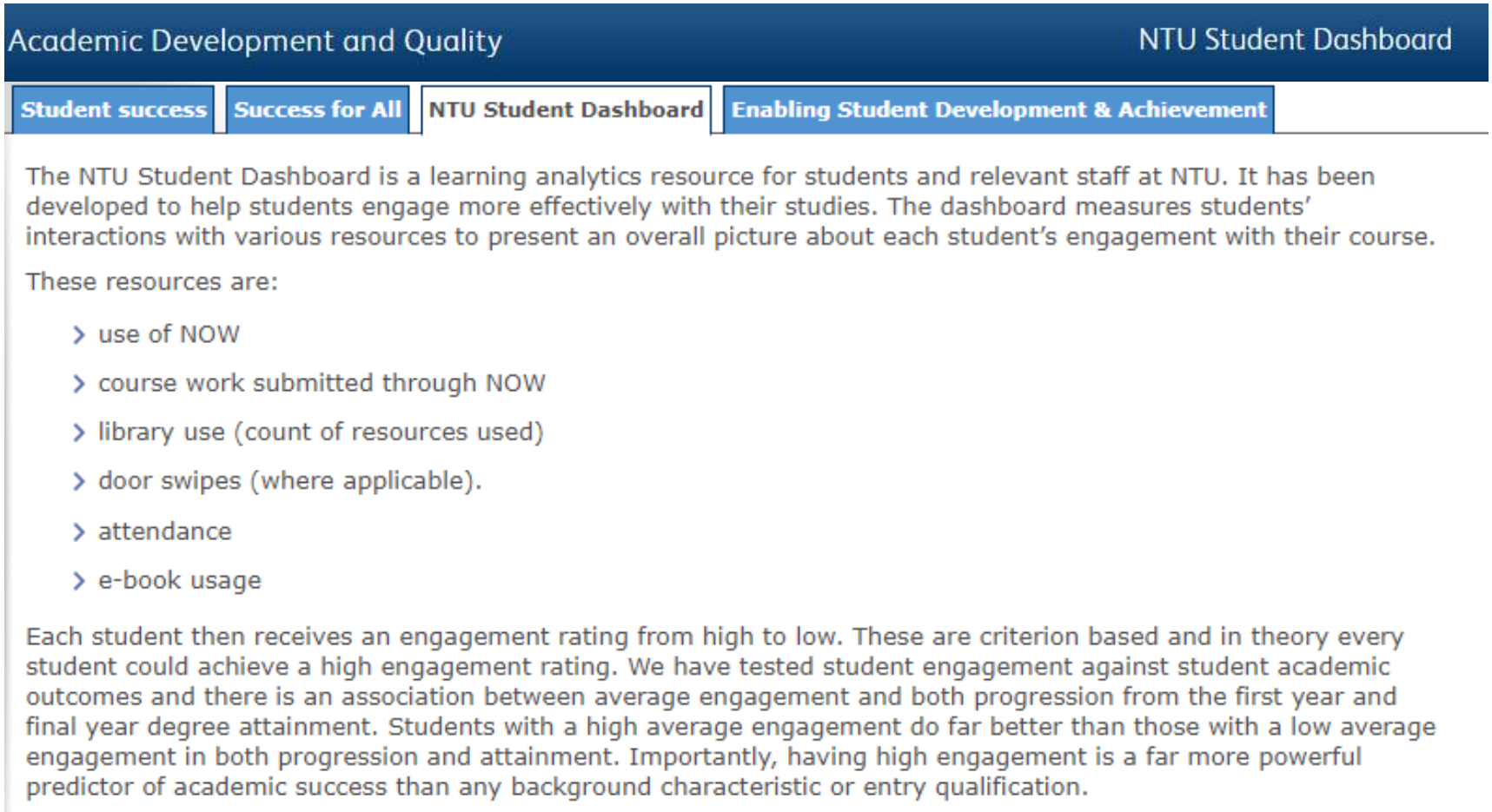
https://www.jisc.ac.uk/rd/projects/effective-learning-analytics



We **recommend that all institutions should consider introducing an appropriate learning analytics system** to improve student support and performance.

https://www.policyconnect.org.uk/hec/research/report-bricks-clicks-potential-data-and-analytics-higher-education

Learning analytics – also ***for*** students



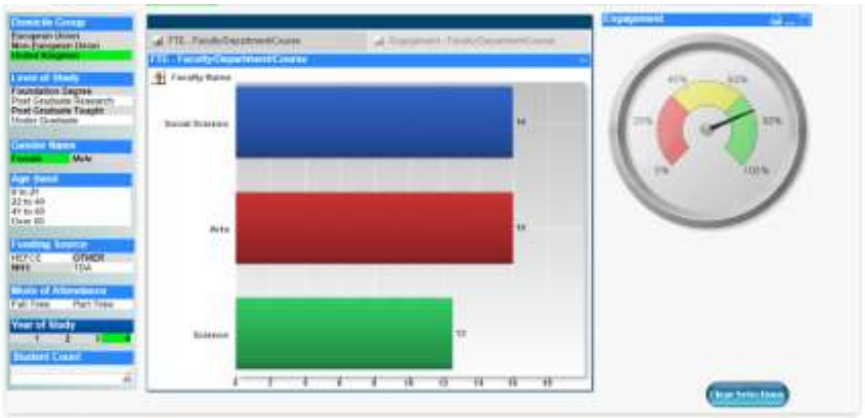
Using generic BI tools

Learning analytics using business intelligence systems

By Niall Sclater on October 22, 2014

Systems

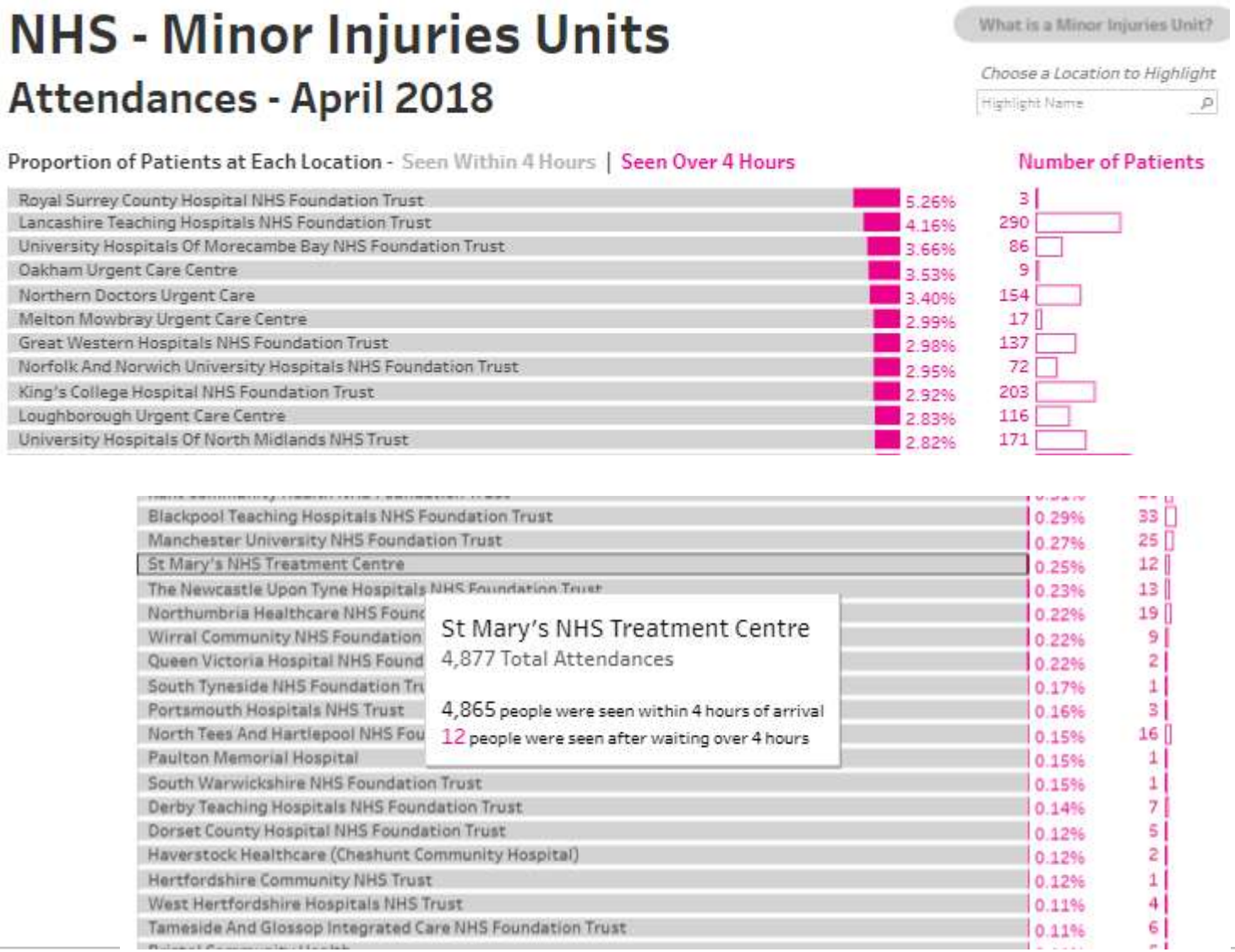
IBM Cognos, QlikView, Tableau



"Many universities and colleges are using **generic business intelligence (BI) systems** to obtain insight on various aspects of their operations – and are beginning to use them for analysing the student learning experience too. The most widely used systems IBM *Cognos*, *QlikView* and *Tableau*.

https://analytics.jiscinvolve.org/wp/2014/10/22/learning-analytics-using-business-intelligence-systems/

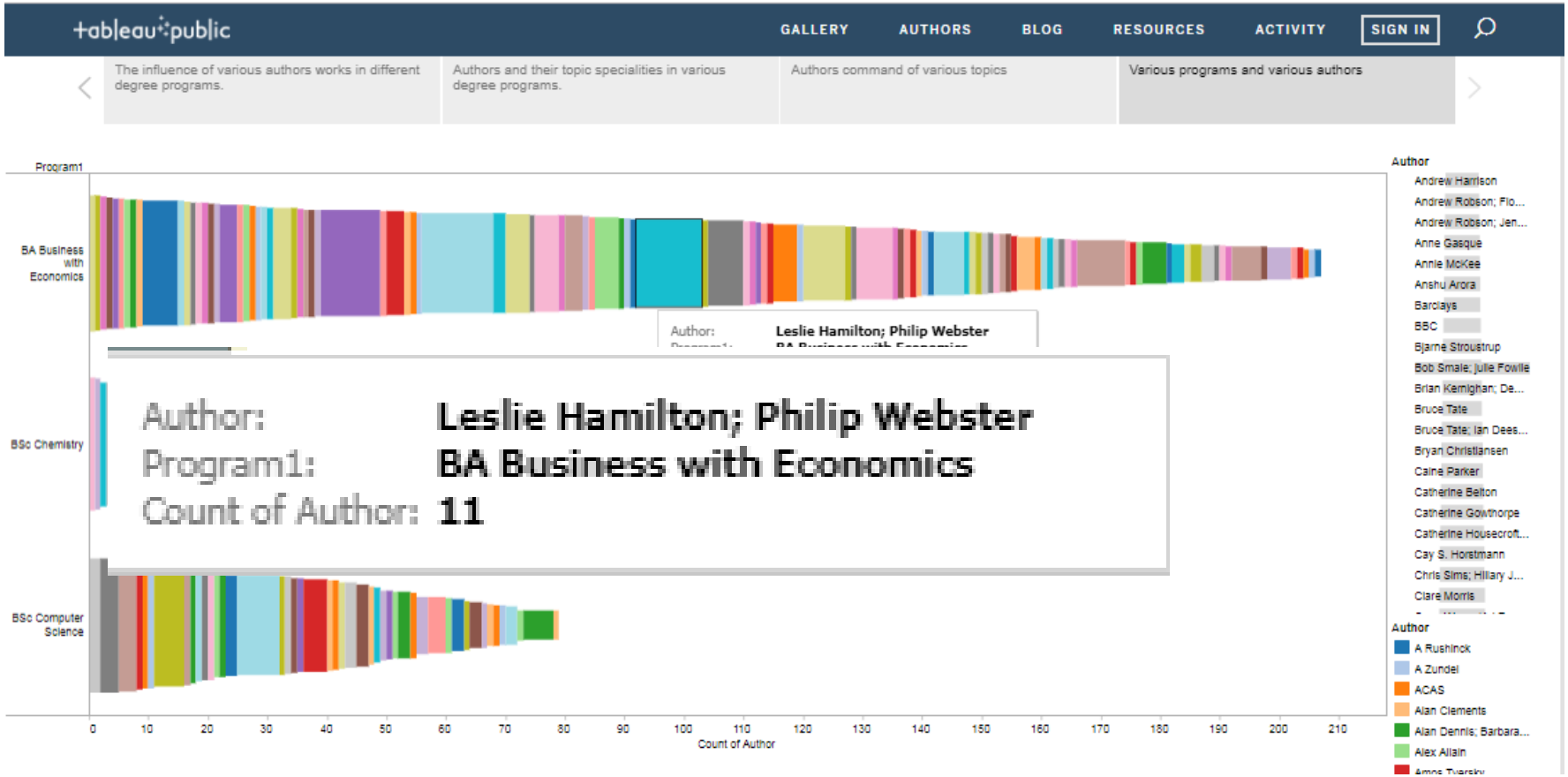
Visualise and compare



Kenchadconsulting helping create more effective libraries…..

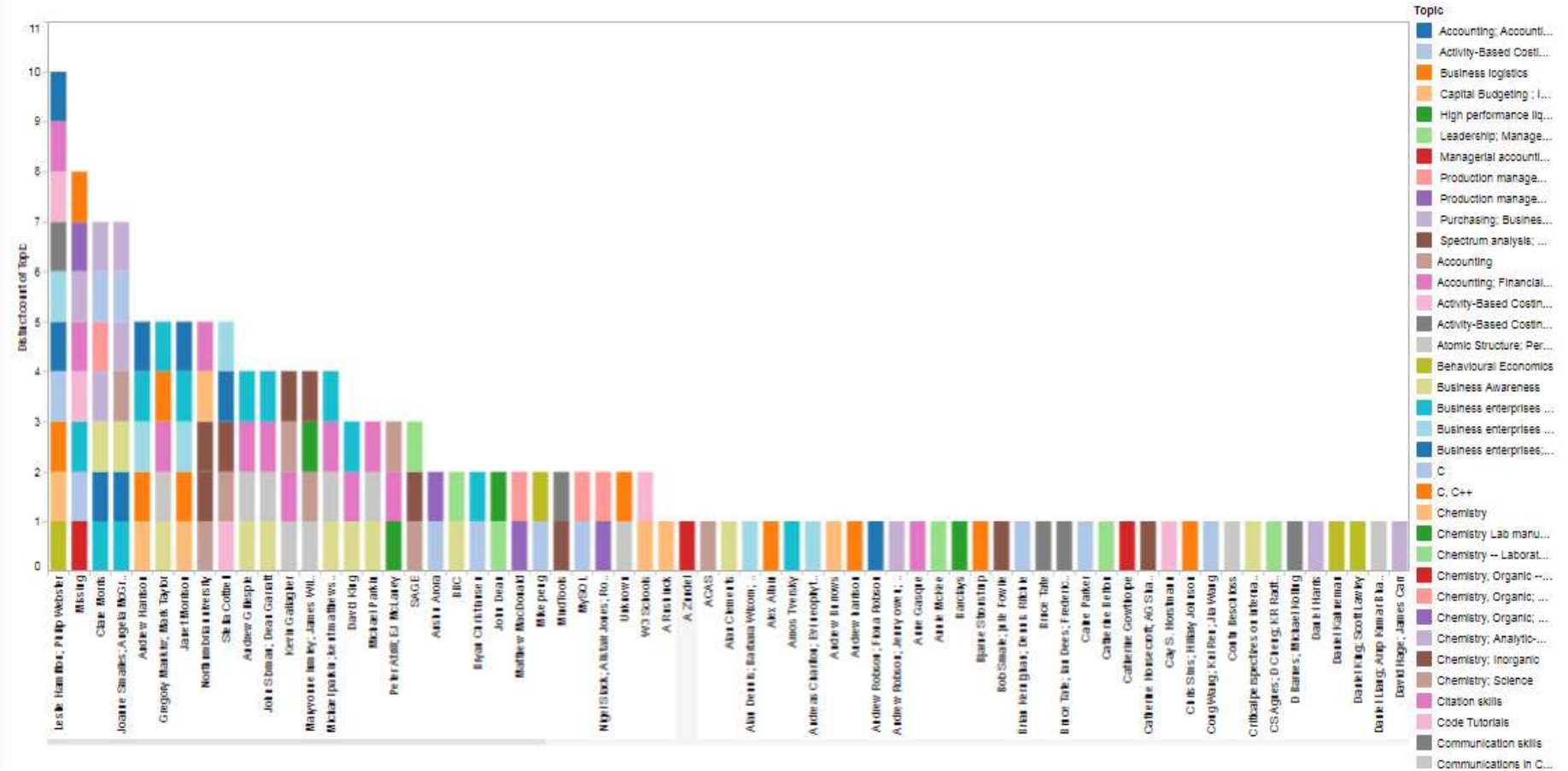
Visualise and compare

Various programmes and various authors



https://public.tableau.com/profile/kushwanth#!/vizhome/JISCdata1/Story1

The same data visualised differently



So what to do?

Put your data to work

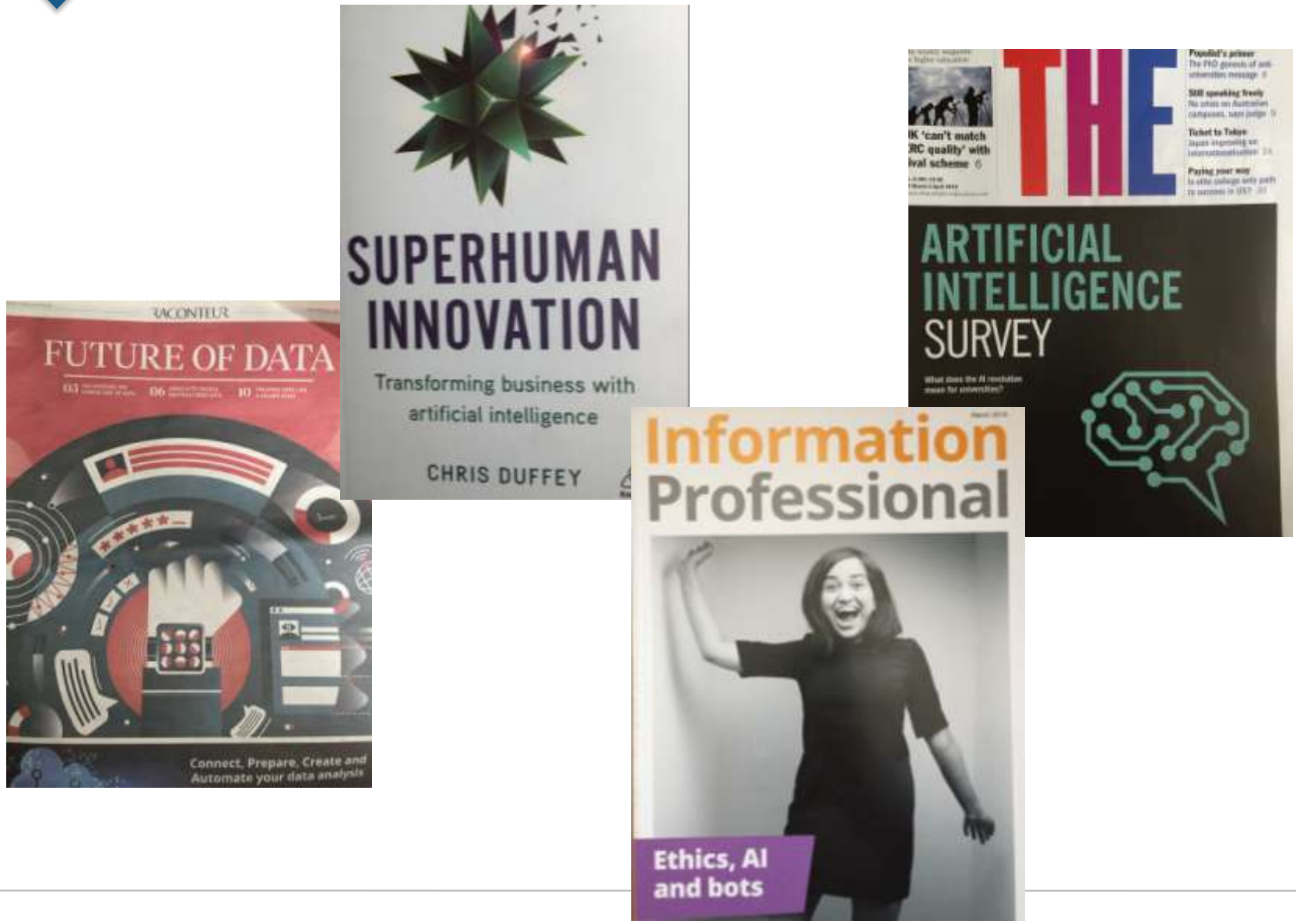
Your university or college already collects a variety of data about students and you can use it to:

- **Transform their learning experience**
- **Support their wellbeing**
- **Help them to achieve more**

Our learning analytics service helps you put your data to work to tackle some of the big strategic challenges. It is the world's first national learning analytics service developed to address the **key strategic goals of HE and FE** organisations.

https://www.jisc.ac.uk/learning-analytics

Part 2: Artificial Intelligence



Kenchadconsulting helping create more effective libraries…..

Adopting AI ...is a journey, not a silver bullet that will solve problems in an instant. It begins with gathering data into simple visualizations and statistical processes that allow you to better understand your data and get your processes under control. From there, you'll progress through increasingly advanced analytical capabilities, until you achieve that utopian goal .....

Data Is The Foundation For Artificial Intelligence And Machine Learning, By Willem Sundblad Forbes [magazine] 18 October 2018
https://www.forbes.com/sites/willemsundbladeurope/2018/10/18/data-is-the-foundation-for-artificial-intelligence-and-machine-learning/#6750b0a151b4

So what is AI?

To understand where AI should be used and will be most successful, one must understand what AI really is. AI, or machine learning, refers to **a broad set of algorithms that can solve a specific set of problems**, if trained properly.

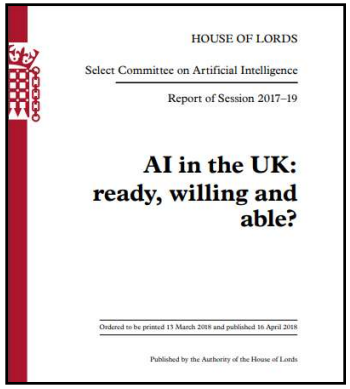
The success of artificial intelligence depends on data, Nick Ismail Information Age [blog] 23 April 2018 https://www.information-age.com/success-artificial-intelligence-data-123471607/

The AI bucket consists of:

- Big data
- Analytics
- Machine learning
- Natural language processing
- Data visualisation
- Decision logic

Cox, A.M. Pinfield, S. and Rutter, S. (2018) The intelligent library: Thought leaders' views on the likely impact of artificial intelligence on academic libraries. Library Hi Tech. ISSN 0737-8831 https://doi.org/10.1108/LHT-08-2018-0105

Artificial Intelligence What is AI?


HOUSE OF LORDS

Select Committee on Artificial Intelligence

Report of Session 2017–19

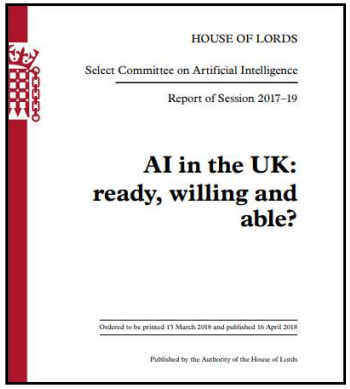
AI in the UK: ready, willing and able?

Ordered to be printed 15 March 2018 and published 16 April 2018

Published by the Authority of the House of Lords

AI in the UK: ready, willing and able? HOUSE OF LORDS Select Committee on Artificial Intelligence. Report of Session 2017–19 HL Paper 100 16 April 2018

Common terms used in artificial intelligence


HOUSE OF LORDS

Select Committee on Artificial Intelligence

Report of Session 2017–19

AI in the UK: ready, willing and able?

Ordered to be printed 15 March 2018 and published 16 April 2018

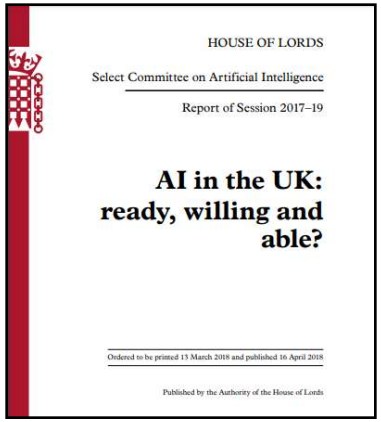
Published by the Authority of the House of Lords

Algorithm

A series of instructions for performing a calculation or solving a problem, especially with a computer. They form the basis for everything a computer can do, and are therefore a fundamental aspect of all AI systems.

Expert system

A computer system that mimics the decision-making ability of a human expert by following pre-programmed rules, such as ‘if this occurs, then do that’. These systems fuelled much of the earlier excitement surrounding AI in the 1980s, but have since become less fashionable, particularly with the rise of neural networks.

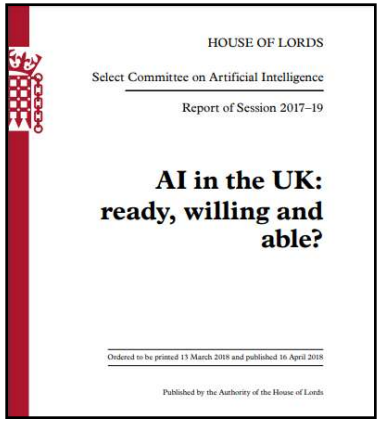


Machine learning

One particular form of AI, which gives computers the ability to learn from and improve with experience, without being explicitly programmed. When provided with sufficient data, a machine learning algorithm can learn to make predictions or solve problems, such as identifying objects in pictures or winning at particular games, for example.

Neural network

Also known as an artificial neural network, this is a type of machine learning loosely inspired by the structure of the human brain. A neural network is composed of simple processing nodes, or 'artificial neurons', which are connected to one another in layers. Each node will receive data from several nodes 'above 'it, and give data to several nodes 'below' it. Nodes attach a 'weight' to the data they receive, and attribute a value to that data. If the data does not pass a certain threshold, it is not passed on to another node. The weights and thresholds of the nodes are adjusted when the algorithm is trained until similar data input results
in consistent outputs.


HOUSE OF LORDS

Select Committee on Artificial Intelligence

Report of Session 2017–19

AI in the UK: ready, willing and able?

Ordered to be printed 13 March 2018 and published 16 April 2018

Published by the Authority of the House of Lords

Deep learning

A more recent variation of neural networks, which uses many layers of artificial neurons to solve more difficult problems. Its popularity as a technique increased significantly from the mid-2000s onwards, as it is behind much of the wider interest in AI today. It is often used to classify information from images, text or sound

AI : What is the opportunity?

Our inquiry has concluded that **the UK is in a strong position to be among the world leaders in the development of artificial intelligence**

AI presents a significant opportunity to solve complex problems and potentially improve productivity, which the UK is right to embrace.

AI : What is the opportunity?

"Human beings and artificial intelligence will work together to create a **brave new world**...This will be a world where **people are freed** to use their natural creative abilities and their amplified intelligence with concern for the drudgery of mundane, repetitive and quite frankly boring tasks. **This is the true mission of artificial intelligence"**

Is AI just about a packet of biscuits?

"Imagine walking into a grocery store, toward the produce aisle, but then getting a ping on your phone that the cookies you bought last week are on sale. You put them to your cart, put your phone away and

keep shopping"

AI : What is the opportunity?

Powerful as it is, one of the problems is that **artificial intelligence is not aimed at the worthiest of the worlds problems**. Today if you look at the very successful AI applications at scale they are in the field of making people click more ads

AI and bots can create fairer world. Rob Mackinley -interview with Kriti Sharma. Information Professional; March 2019

AI in Higher Education – a diversity of views

Nearly all agree that AI will be a very big issue for higher education. ..**most of the university leaders among the respondents express an intention to develop strategies** where they do not already have one.

universities will use AI to select the best candidates for degree courses,

not one of *THE*'s respondents finds it easy to recruit and retain academic staff able to teach and research AI,

"There is a lot of work going on in AI by way of tech ..but what a lot of people are trying to look into is the ethical tone of it all," ..it's really difficult to do any humanities- and social sciences-based work on it because grants are not tailored towards it."

Universities today cannot compete against the Googles of the world because they do not have that data.

respondents rank research as the area of university management and practice likely to be most affected by AI

https://www.timeshighereducation.com/features/microsoft-survey-ai

AI in Higher Education – a diversity of views

Most agree that AI will have the cognitive capacity to participate in scientific advancement, at least to some extent. Exactly half believe that AI will be able to direct the testing of scientific hypotheses at least as well as humans can, and 52 per cent think machines will be able to generate new scientific hypotheses as well as humans can.

respondents are keen on the idea that not only science students but also humanities and social science students will need to be taught specific technical skills to help them programme and interact with artificial intelligence productively

Respondents are reasonably confident that AI will be able to provide student feedback at least as well as humans can, with student assessment another area where AI could play a big role.

AI in Higher Education – a diversity of views

"AI can help flip the classroom, personalise education and tackle the increasingly and distressingly prohibitive cost of delivering that education. Some of the skills that universities help people learn will change. But the skills that will be most in demand will tend to be old-fashioned ones, that universities used to deliver, such as analytical and creative thinking."

"The key problem, as ever, is that a small pool of academics have managed to push politicians to think that investing in AI research is going to change the world. I don't think that is right,"

"AI machine learning can never replace that until you make a totally new, self-replicating machine or life form with artificial consciousness…And that will remain firmly in the realm of science fiction for many hundreds of years."

Gartner®

Artificial intelligence has opened up strategic opportunities for tertiary institutions to fundamentally redesign their businesses to deliver premium personalized services. Education CIOs can learn how virtual personal assistants provide an important building block toward that objective.

Use AI to Take Student Success to the Next Level of Personalization in Higher Education. Nick Ingelbrecht & Jan-Martin Lowendahl. Gartner [report]. 14 February 2018
https://www.gartner.com/doc/3857266/use-ai-student-success-level

Student success

AI based Teaching Assistants

INSIDE HIGHER ED

At the Georgia Institute of Technology, Ashok Goel, professor of computer and cognitive science, has been working with virtual teaching assistants for several years….In 2016, Goel made headlines after revealing that some of his **students (in a master's-level online course in AI) were unable to distinguish between AI and human** TAs answering questions in a discussion forum

The **AI TAs can't answer deep questions about conte**nt -- only human teaching assistants can do that -- said Goel, but the AI TAs are useful "because students tend to ask the same questions again and again." Questions about assessment or deadlines are easily handled by AI, he said.

Pushing the Boundaries of Learning With AI. By Lindsay McKenzie. Inside Higher Ed 26 September 2018 https://www.insidehighered.com/digital-learning/article/2018/09/26/academics-push-expand-use-ai-higher-ed-teaching-and-learning

Identify “at risk” students-predictive analytics

As your command center for student success, Illume® uses your institution’s data to develop **personalized predictions** that are timely, **accurate, and actionable**.

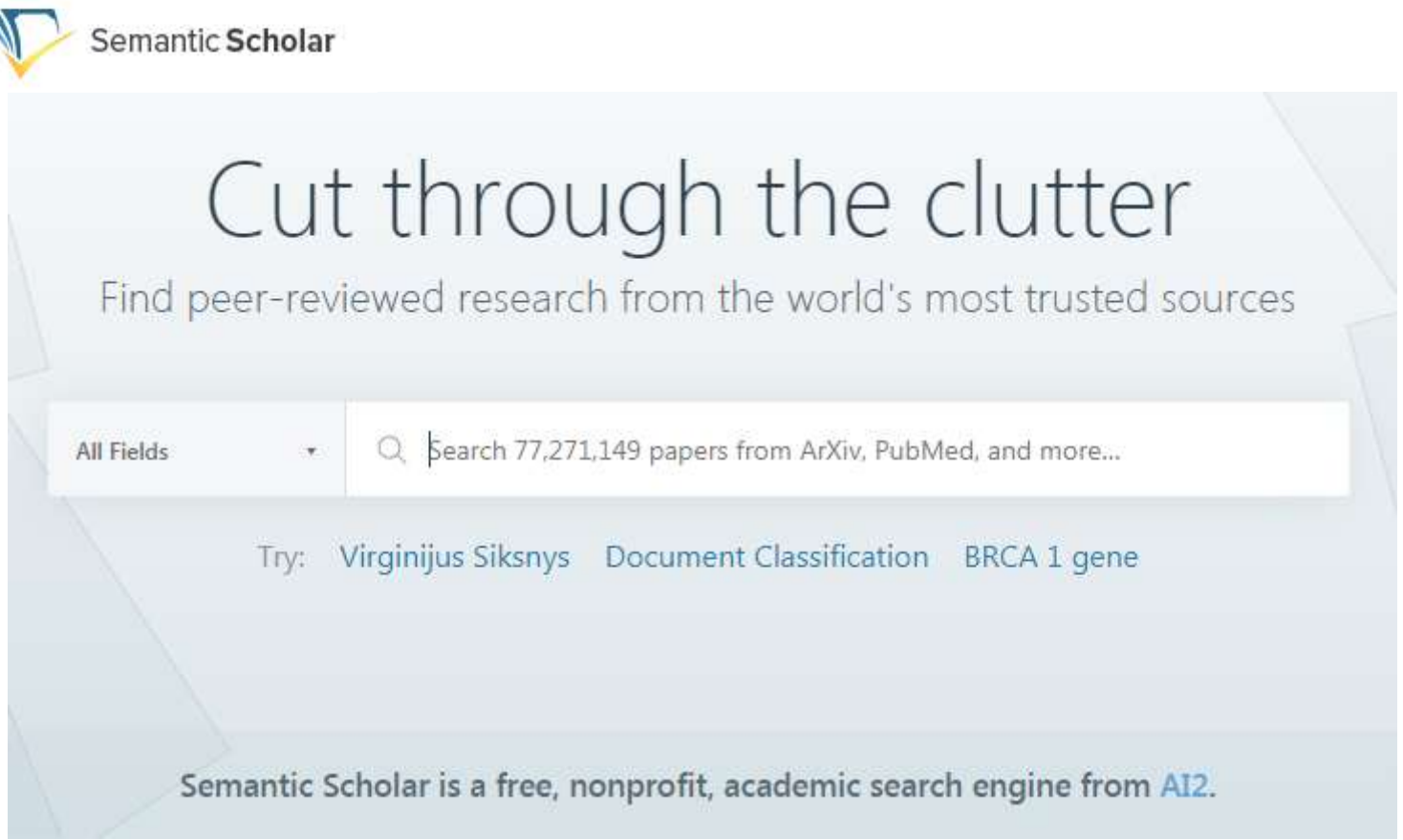
Pinpoint at-risk students hiding in plain sight, develop coordinated student success initiatives, and deploy targeted email interventions, all from one place.

Each student’s data tells a story. Are you listening?

https://www.civitaslearning.com/

Helping research

Helping research



Helping research

Semantic Scholar

How it works

We've pulled over 77 million scientific papers from sources like PubMed, Nature, and ArXiv.

Our AI analyzes research papers and pulls out authors, references, figures, and topics.

We link all of this information together into a comprehensive picture of cutting-edge research.

Yewno Discover: *Case Study*

Kamran Naim,
Stanford University
Doctoral Candidate & Instructor

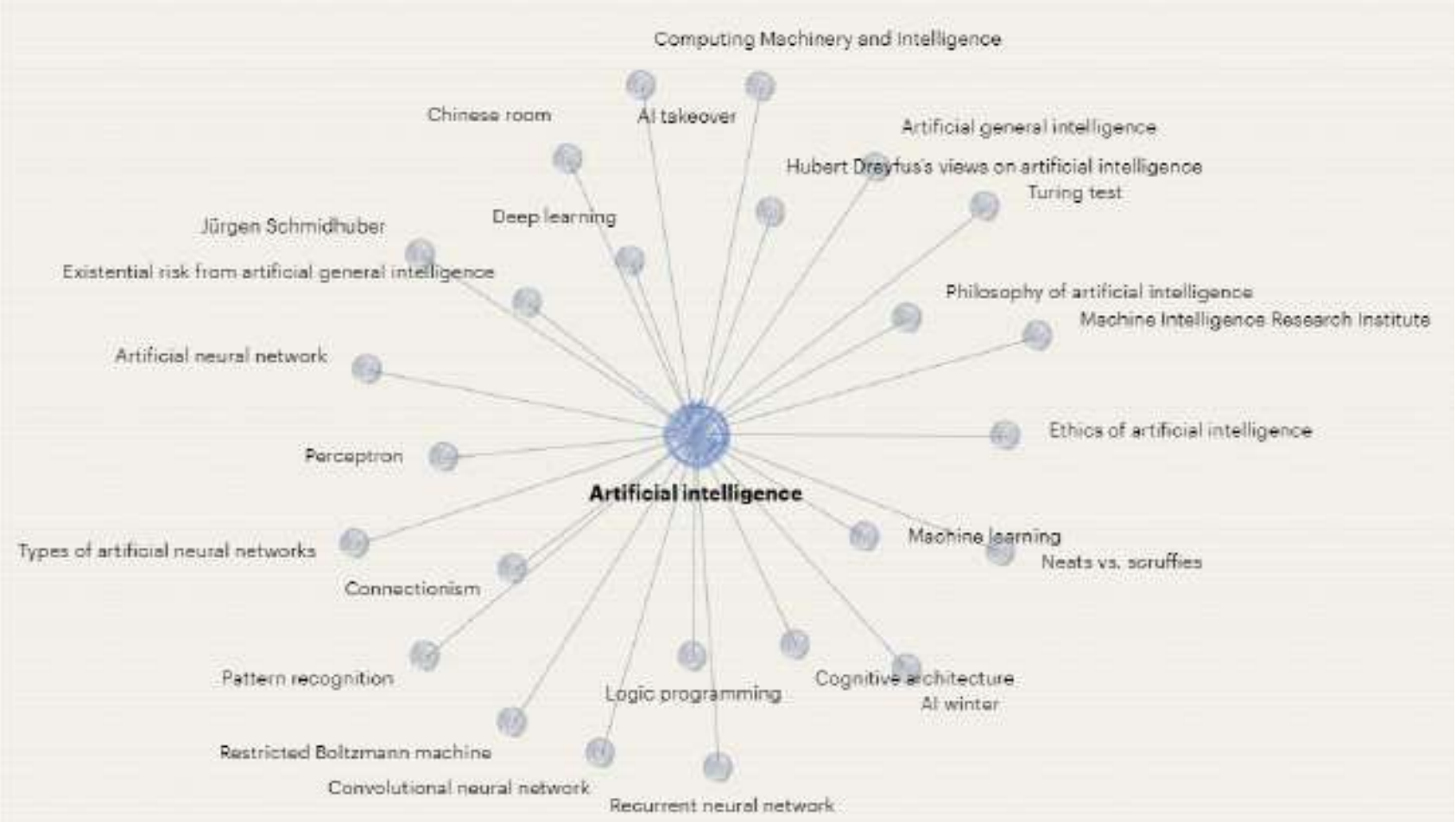
As doctoral candidate and part-time instructor in Stanford's graduate school of Education, Kamran Naim is keenly aware of the importance of developing digital literacy skills to survive today's deluge of information.

"The volume of scholarly content that is produced far exceeds anyone's ability to read it."

He sees information tools powered by neural networks and artificial intelligence (AI) -- where computers replicate human processes like decision making or perception -- as a cutting-edge method to navigate this growing body of literature.

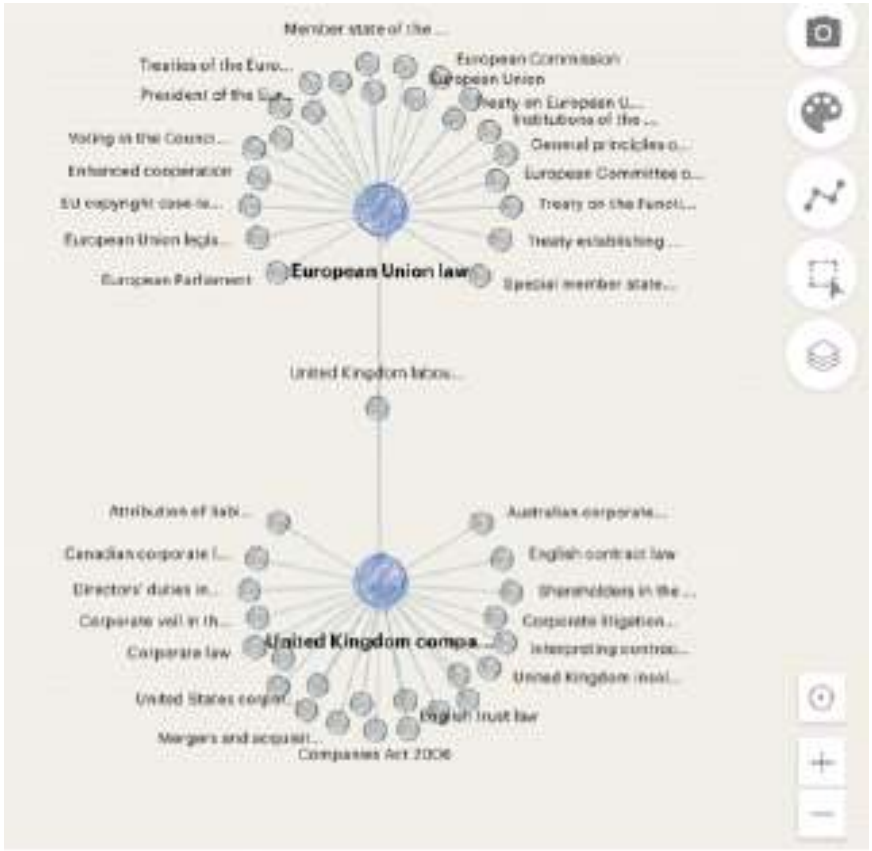
"Yewno's knowledge map allows Kamran to investigate concepts and relationships between concepts to help broaden the keywords he'll use in his research."

Helping research



Helping research

Esther has been studying renewable energy, and wanted to find out how the UK's departure from the EU will affect its policies on wind turbines.



"The mind map layout was really useful, as it provided a great **visual way of looking at links between subjects.** I use spider diagrams for revision purposes so it's a method of presentation that really resonates with me. I ended up finding out lots of information about renewable energy conflict, and the policies of other countries like New Zealand, which will be very useful points of comparison for my research, as well as it being an area that I am generally very interested in.

Yewno showed me the **links between these different subjects without me having to perform extra searches,** which will **save me a lot of time** when it comes to essay writing."

Helping research

Yewno's mission is 'Knowledge Singularity' and by that we mean the day **when knowledge, not information, is at everyone's fingertips.** In the search and discovery space the problems that people face today are the overwhelming volume of information and the fact that sources are fragmented and dispersed. There's a great T.S. Eliot quote 'Where's the knowledge we lost in information' and that sums up the problem perfectly.

Ruth Pickering. Chief Business Development & Strategy Officer

Do You Know About Yewno? By Alice Meadows 7 June 2017
https://scholarlykitchen.sspnet.org/2017/06/07/do-you-know-about-yewno/

Helping research

Making research more discoverable
Content is at the centre of everything a publisher does. Enriching that content delivers significant value across the whole content life cycle. One particular area where the need for content enrichment can add significant value is in **enabling the researcher to find and discover the most relevant content to assist in the researcher's workflow**. Features that can be enhanced using enrichment techniques are relating articles, subject and context navigation, categorisation of content and identification of entities to provide linking to other relevant content.

http://www.67bricks.com/index.php/content-enrichment-industry-insights-1-2016

Improving discovery through relatedness

Publishers are trying to increase the usage of their content but at the same time ensure that they are providing researchers with the most relevant content when they search. One area where content enrichment can assist is in the area of relatedness. Here rather than use just the words in the documents to find other related content you can use **techniques to give more meaning and context to the articles and the relationship they have with other articles**. One approach uses software tools to **extract the meaning from content** to create a content fingerprint and then uses this fingerprint to find related pieces of content.

http://www.67bricks.com/index.php/content-enrichment-industry-insights-1-2016

Helping research

World-leading archive uses GraphDB to build a machine learning assisted knowledge .management platform

The IET looked to Ontotext to deliver **artificial intelligence** technologies into its database for **discovering emerging trends and relationships**. This technology gives customers both a deeper understanding of current developments and more value from the data they have contributed so much towards.

Press Release 24 August 2016 https://www.ontotext.com/company/news/ontotext-selected-unleash-power-institution-engineering-technologys-knowledge/

AI transformation...

Elsevier's Dr. Jabe Wilson answers the question "What does 2018 hold for AI in publishing?"

He points out that while many industries have shifted to digital, the impact is especially dramatic in scientific publishing and R&D "due to the sheer volume of data researchers must sift through." In fact, the desire to help researchers make sense out of all this data is behind **Elsevier's transformation from publisher to "information analytics provider."**

https://www.elsevier.com/connect/jabe-wilson-on-the-future-of-ai-and-scholarly-publishing

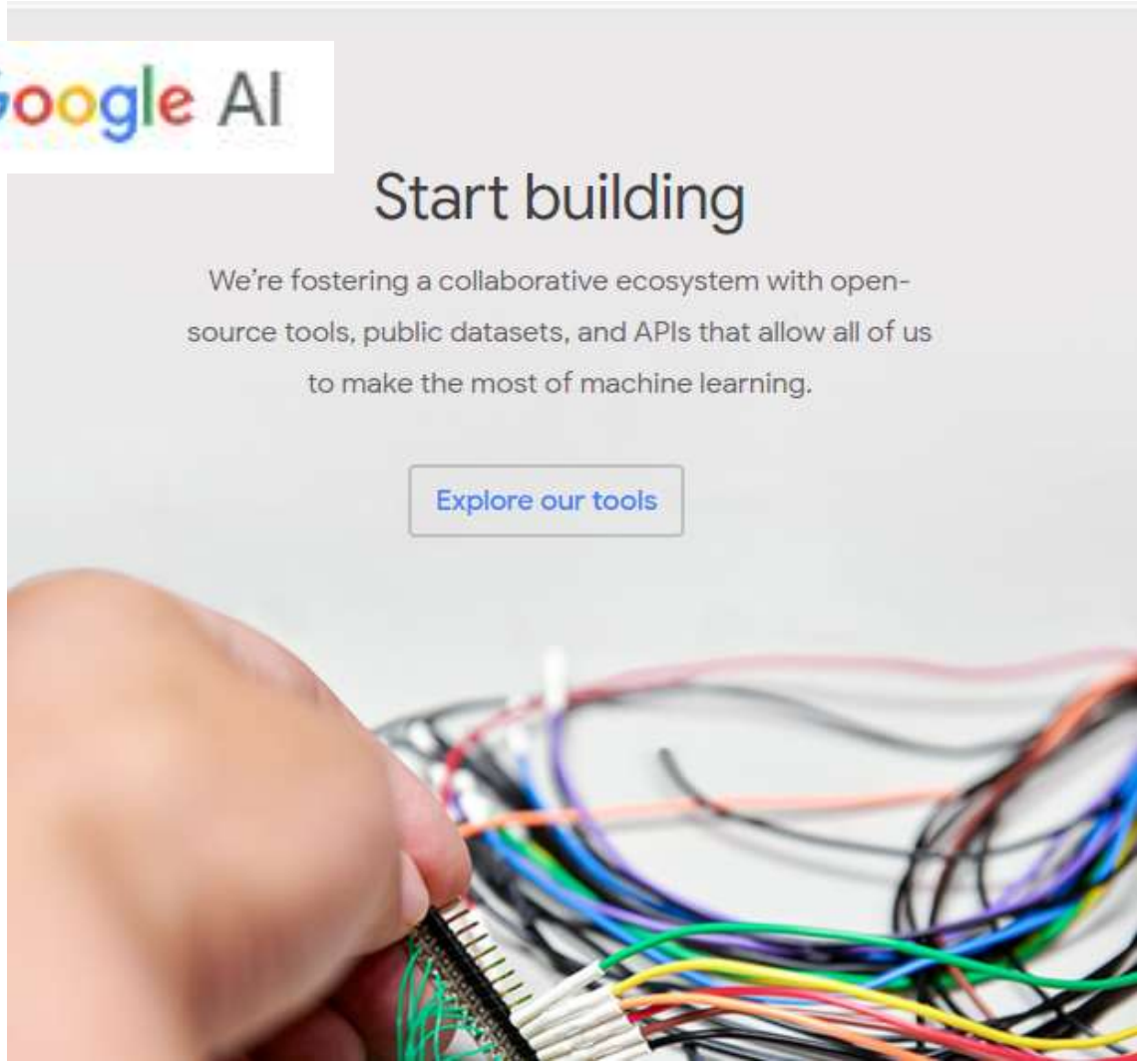
The importance of data

AI works best when large amounts of rich, big data are available. The more facets the data covers, the faster the algorithms can learn and fine-tune their predictive analyses. According to industry predictions, in 2018, AI's greatest limitation — high quality data — will become more evident. Successful machine learning depends on large and broad data sets.

In the next wave of AI empowerment, the algorithms are commoditised, but **whoever owns the data is king**.

The success of artificial intelligence depends on data, Nick Ismail Information Age [blog] 23 April 2018 https://www.information-age.com/success-artificial-intelligence-data-123471607/

The importance of data: Who wants your data?

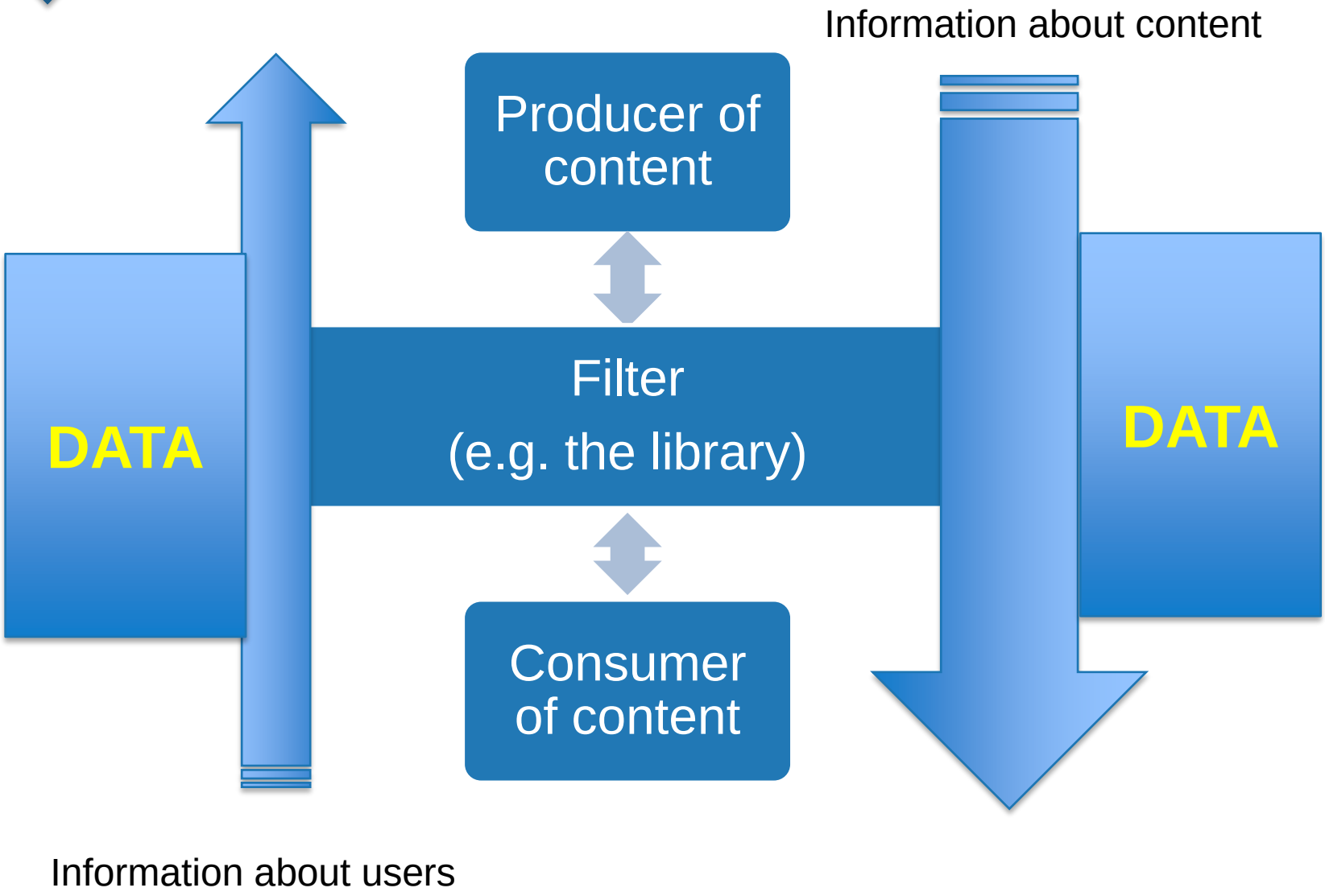


The importance of data

Datasets
HathiTrust makes the texts of public domain works in its corpus available for research purposes. The works fall into two categories: non-Google-digitized volumes, which are freely available, and Google-digitized volumes, which are available through an agreement with Google.

Non-Google-digitized volumes
Description
Approximately 776,200 public domain volumes as of October 2018, primarily, though not exclusively, English language materials published prior to 1924.

Platforms are driven by data…..



Challenges to adoption of big data/AI

Uncovering the human side of data. By Charles Orton-Jones. In Future of Data. Raconteur. 26 March 2019

Data Wars

‘Whoever acquires and controls the data will have hegemony in the future

Indian Prime Minister Narendra Modi
World economic forum . January 2018

Analytics India MAGAZINE

https://www.analyticsindiamag.com/modi-wef-davos-data-control-real-wealth/

Who might do it? What capabilities might be required?

Who is doing it/might do it?

Individual institutions?

Publishers?

Intermediaries?

Sector bodies?

Data companies?

Capabilities required?

Access to good, appropriate data

Machine intelligence – AI

Platform to deliver the service

A workable biz model

The data wars: moving from management information to data driven intelligence

UKSG: April 2019

Ken Chad

Ken Chad Consulting Ltd

Twitter @kenchad

ken@kenchadconsulting.com

Tel: +44 (0)7788 727 845

www.kenchadconsulting.com