Universities and the Challenge of Open Science

Geoffrey Boulton
University of Edinburgh
ISC-CODATA

Trento
October 2018
The Digital Revolution

Beginning of the “Digital Age”

1986

1993

Analogue Storage

Digital Storage

2003

2007

19

Exabytes

280 Exabytes

2014 - 4000 Exabytes

1 Exabyte=10^{18} bytes

19

Information / Data Volume

Earliest written record

Gutenberg

London Library (500,000 books)

Internet Explosion

12 stacks of books from Earth to Sun

0.01% is stored on paper

2 stacks of books from Earth to Pluto

3000 BC 1450 1900 1992 2002 2006 2010
The Digital Tsunami

Acquisition

Storage

Computation & AI

THE WEB & Ubiquitous Communication
Another Gutenberg Revolution

The technologies by which knowledge is acquired, stored and communicated always been essential drivers of human material and social progress.

Johannes Gutenberg
1400-1468

The Digital Revolution

- vast data streams
- vast source diversity
- vast computational capacity
- learning algorithms
- instantaneous communication
- access anywhere anytime
- low cost

A NETWORKED EARTH
Why Open Data?

1. **Self correction**

   *The progress of science is strewn, like an ancient desert trail, with the bleached skeletons of discarded theories that once seemed to possess eternal life.*

   (Arthur Koestler)

2. **Efficiency.**

   *If you have an apple and I have an apple and we exchange them, then we will both only have one apple. But if you have an idea and I have an idea, and we exchange them, then we will both have two ideas.*

   (George Bernard Shaw)

3. **Scaling up**

   *If we had the sense to share all our separate data about all that we know, and combine them in intelligent ways, then science, knowing, would have come of age.*

   (Richard Feynmann)

Why does Feynmann’s comment matter?
“Broad Data”: The imperative for interdisciplinary science

The Broad Data Challenge: “from owning a little data to accessing a lot”
Complexity: system state & dynamic evolution

Simulating system dynamics

Mapping a complex state

Emergent behaviour of a specific 6-component coupled system

Image of brain cells in a rat

Most global challenges are embedded in “complex” systems
BUT: though integrated modelling is well-established
data integration is problematic
Big/Broad Data reveals patterns in nature and society that have been beyond resolution

Example: North Atlantic Ocean Circulation
The human future is an urban future
Today's growing cities are often islands of stability and good governance in oceans of uncertainty, partly because they are better able to adapt to changing realities than entire countries. They can serve as role models, if not vanguards, for the new political-economic models the world needs. Understanding the “city organism” is a vital priority for science & society.

Key linked variables for the urban ecosystem
• Population & its fluxes (formal/informal)
• Energy
• Water
• Sewerage
• Waste
• Food
• Retail
• Transport
• Income
• Housing
• Leisure
• Ethnicity
• Education
• Employment
• Health
• Social services

And universities are the only places where this range of knowledge can be found
SDG Goal 3: Good Health & Wellbeing
How to integrate data to better predict and respond to emerging infections?

- Better prevent and control outbreaks of infectious diseases
- Make data more useable & findable: Pool and standardise individual patient-level data
- Make data available to the health, scientific & humanitarian communities
The “Science International” Accord:
principles of open data
(www.icsu.org/science-international)

Responsibilities
1-2. Scientists
3. Research institutions & universities
4. Publishers
5. Funding agencies
6. Scholarly societies and academies
7. Libraries & repositories

8. Boundaries of openness

Enabling practices
9. Citation and provenance
10. Interoperability
11. Non-restrictive re-use
12. Linkability
Open Science – Open to Whom?

In Europe
“The European Open Science Cloud”
= Open Data + Open Access Publishing

In Africa
“The African Open Science Platform”
= As Europe + networks of collaborative problem-solving with societal partners to foster scientific credibility, practical relevance and socio-political legitimacy.
“At last an authoritative voice has demonstrated the corruption of science, driven by an almost religious certainty, that has propounded a theory that can now clearly be seen to be false, based on unreliable and in some cases invented evidence, ruthlessly used to advocate damaging and unnecessary changes in US policy, “

Public policies

Emotions trump Facts:
why is science a poor persuader on many major issues?

scientists must be both emotionally intelligent and rigorously rational
The Dark Side

- Existential risk
- Political manipulation
- Privacy
- Cyber-crime
- Autonomous weaponry
- Cyber-warfare
Challenges of the Digital Revolution

Challenges for Science

- Integrating data from diverse fields (Broad Data)
- Standards for reproducibility
- Maintaining traceability
- Data to be Findable-Accessible-Interoperable-Reusable
- Open access publishing
- Data for diplomacy
- Privacy enhancing technologies

Challenges for Society created by the digital revolution

- Managing the World’s data
- Machine learning results and interpretability
- Autonomous systems & the human role – pampered, but purposeless, or extinguished
- Implants and the transformation of humanity
- The future of work
- Another “north-south” knowledge divide?
- Science as a public enterprise – or the privatisation of knowledge?
- Education & the role of the University
- The dare side: cyber crime, cyber warfare, political subversion
Questions from Roberto Caso

- what is/should be the role of universities in open science?
- do they have a leading role to play, or is that being taken by commercial platforms with a huge power to control information?
- what is the right role of universities where commercial publishers have major scientist-derived databases and control scientific social networks?
- does open science represent a bulwark of university autonomy and academic freedom?
STAKEHOLDERS

**Government**
- international norms increasingly disrupted or questioned
- levers of power and responsibility divested into private hands
- public good enterprises are shifting from governments to markets
- policy-making subverted by “alternative facts”

**Technology companies**
- behemoths (e.g. Google & Facebook) dominate online presence
- grab all profits from the attention economy
- extending their reach: privatising knowledge & education

**Universities**
- expanded in size but shrunk in purpose
- are they now more than instrumental to government policies?

**Science**
- Open Science as powerful new paradigm
- but how open is “open”

**Citizens**
- locked in self-referential, self-reinforcing echo chambers
- towards nationalistic populism
- inimical to international sharing?

THE WEB

**Expectation:** democratise information; inhibit monopolies; fragment audiences

**Reality:** strengthens monopolies; polarises audiences; indifferent to truth or falsehood
UK Institute for Public Policy Research, 2013, *An Avalanche is Coming: Higher Education and the Revolution Ahead*: "the solid classical buildings of great universities may look permanent but the storms of change now threaten them".

Ernst and Young 2012, *University of the Future*: “A thousand year old industry on the cusp of a profound change. Just as digital technologies have transformed media, retail, entertainment and many other industries – higher education is next”.

World Bank, 2014. Higher education is "the next big private enterprise opportunity".
But universities have crucial attributes

They:

• Have far greater public trust
• They educate the next generation
• They employ the majority of publicly funded scientists
• They conserve and generate knowledge (tested against reality) across the whole spectrum
• Academics hold and curate most of the corpus of rigorous scientific data
Debate:

The instrumental view:
The university as a supermarket, with shelves stocked with:
• specific skills for the labour market;
• parcelled products for business innovation.

But they are precisely the skills that machine learning will replace!

The classical view
Drew Faust: “A university is not about results in the next quarter; it is not even about who a student has become by graduation. It is about learning that moulds a lifetime; learning that transmits the heritage of millennia; learning that shapes the future”.

Ben Okri: “The role of the University is to set up their students for the act of self-discovery.”

Michael Higgins (President of Ireland): The life-enhancing skills such as creativity, clarity of thought and expression are the skills that will be essential to the citizens of the future in adapting to and creatively using a new technological world, making informed choices about what constitutes survival, and what is meant by human flourishing, solidarity or humanity itself.
An Example: a UN Initiative
a digital ecosystem for planetary data: October 2018

Thesis: from governments and funders to investors and markets: creating sustainable business models for private sector leadership.

Antithesis:
• Upstream IP should not be private
• Omits existing time series space data
• Omits earth-derived data
• Purpose: unclear: for global sustainability or private sector profit
• Based on new private sector investment in new space technologies

Synthesis
• Purpose: to support global management towards sustainability (agenda 2030)
• Global governance for key databases in relevant science irrespective of source (geosphere; atmosphere/cryosphere/hydrosphere; biosphere; anthroposphere)
• No upstream IP – private value reaped downstream – incentives for upstream investment
• The private sector should work out its own business model

Spin-off principle? - Access to knowledge a human right
Basis for White Paper to UN General Assembly
Valediction

• Knowledge and science are quintessentially public enterprises;

• The world faces a crisis of sustainability and a converging set of environmental, socio-economic, political and cultural challenges for which the digital revolution offers the potential for powerful new solutions;

• Many practices of researchers and institutions are merely adaptations to an outmoded print technology and are inimical to optimal exploitation of the digital revolution and to the public good;

• Universities have unique potential, and a responsibility, to respond to these challenges – they should do so with boldness, leadership, responsibility and a greater sense of themselves (enabled by autonomy & academic freedom)

Failure to confront the challenge would be seen as an existential failure