

Dr Helena Korjonen & Professor Emma Schymanski

University of Luxembourg & members of the Open Science Working Group

A data journey: data obscurity to data enlightenment

How far have we (really) come in 18,000 years?



Overview

- Swift data journey through time
- Identifying persistent & new challenges in our data journeys
- Highlights from the University of Luxembourg



18,000+ years ago: sticks and bones

- Ishango bone, Congo
- Museum of Natural Sciences, Belgium
- Mathematical counting tool
- Lasting format (bone)
- Early humans logging numerical data for later use

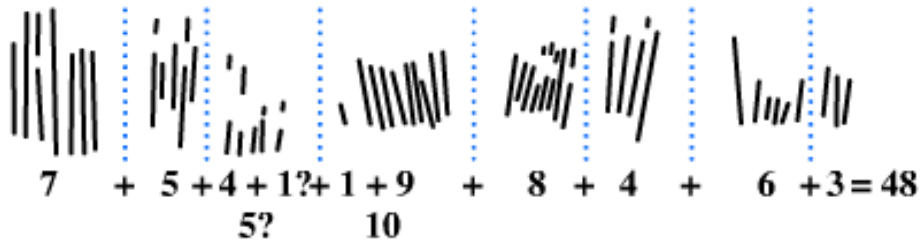


Photo JhowieNitnek (CC BY license)

[https://en.wikipedia.org/wiki/Ishango_bone#/media/File:Ishango_bone_\(cropped\).jpg](https://en.wikipedia.org/wiki/Ishango_bone#/media/File:Ishango_bone_(cropped).jpg)

17,000+ years ago: Record-keeping for survival

- Lascaux cave paintings, France
- Counting, Record-keeping
- Open format?
- Enormous spaces
- Quantify & record the world through repetitive symbols or marks for survival & improve life



5,000+ years ago: Construction & payments

- Ancient Sumerians
- Format long-lasting stone
 - Large & heavy
- Record-keeping of harvests & complex calculations
- Taxes, payments, construction

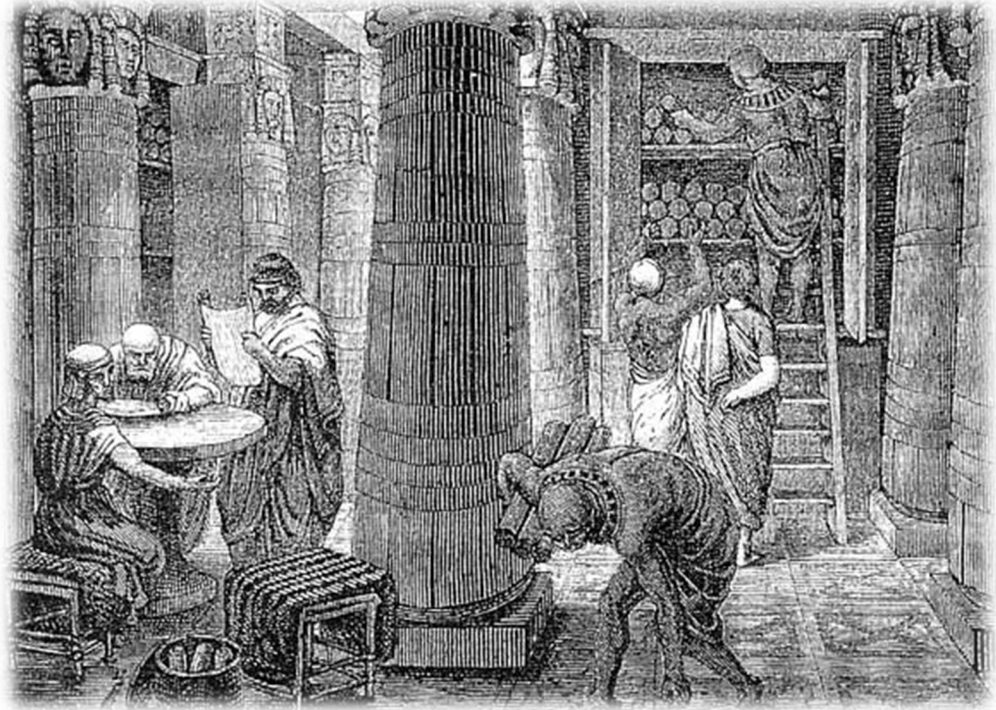


Babylonian mathematical tablet Plimpton 322.

Credit: Christine Proust and Columbia University

2,000+ years ago: Organising knowledge

- First “Ancient world” libraries
- Format fragile scrolls
- Ancient libraries curated and organised knowledge in collections
- Many threats!



<https://www.crystalinks.com/libraryofalexandria.html>

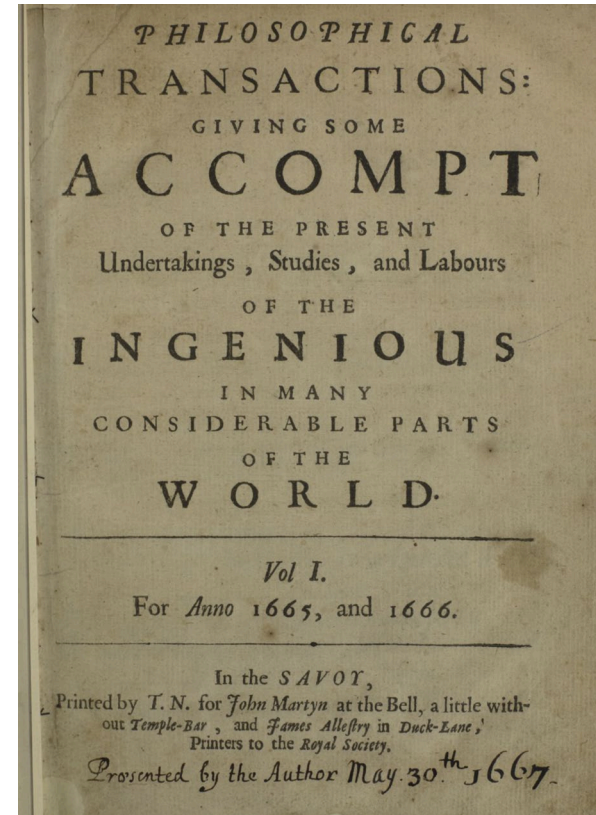
1450s onwards: Population statistics

erslängder, SE/VALA/00419/A I/1 (1753-1785)

- Example from Sweden
- Fragile notebooks
- Churches
- Population registers, land ownership & sensitive data

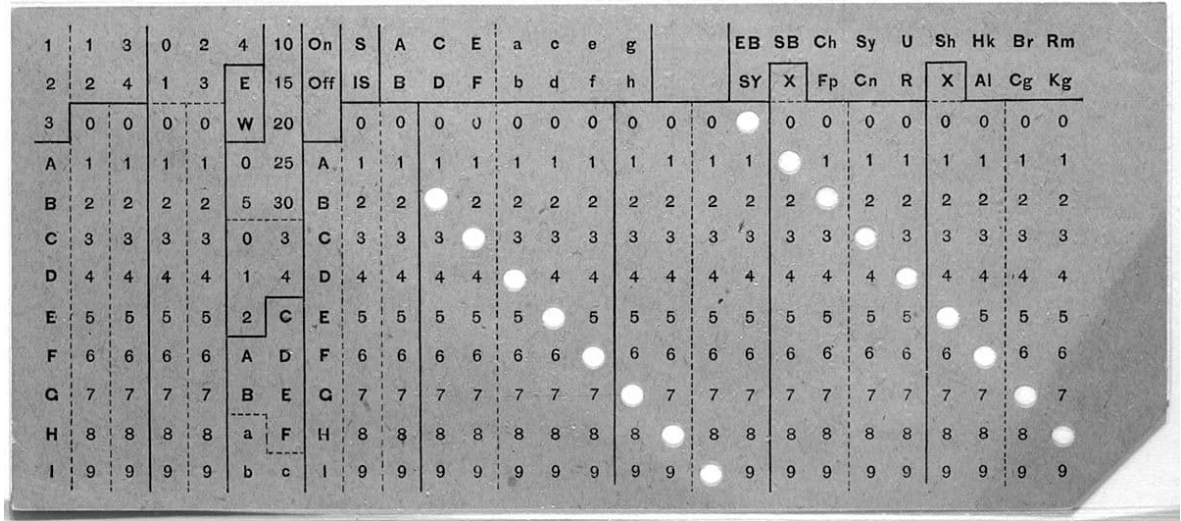
1600s: Shared for the “Greater good”

- Royal Society, England 1665
- Philosophical Transactions journal
- Belief in the greater good & misunderstanding of harm
- The rise of journals and (unethical) clinical trials in the 1600s



1890s: Processing of large datasets

- Herman Hollerith, US Census
- Tabulating Machine with punchcards for data input
- Reduced processing time from 7.5 to 2.5 years



1	2	3	4	5	6	7	8	9	0	On	S	A	C	E	a	c	e	g		EB	SB	Ch	Sy	U	Sh	Hk	Br	Rm
2	2	4	1	3	E	15	Off	IS	B	D	F	b	d	f	h					SY	X	Fp	Cn	R	X	Al	Cg	Kg
3	0	0	0	0	W	20			0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A	1	1	1	1	0	25	A	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
B	2	2	2	2	5	30	B	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
C	3	3	3	3	0	3	C	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
D	4	4	4	4	1	4	D	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
E	5	5	5	5	2	C	E	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
F	6	6	6	6	A	D	F	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
Q	7	7	7	7	B	E	Q	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
H	8	8	8	8	a	F	H	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
I	9	9	9	9	b	c	I	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9

1920s Data storage magnetic tape

- Germany
- Magnetic tape by Fritz Pfleumer
- Audio, video, data backup



Magnetic tape by Fritz Pfleumer

1960s: Early “cloud” storage

- Dr Joseph Carl Robnett Licklider (“Lick”), American psychologist and computer scientist
- A network of computers that could communicate with each other & share resources



SAGE system terminal

https://en.m.wikipedia.org/wiki/File:SAGE_console.jpeg

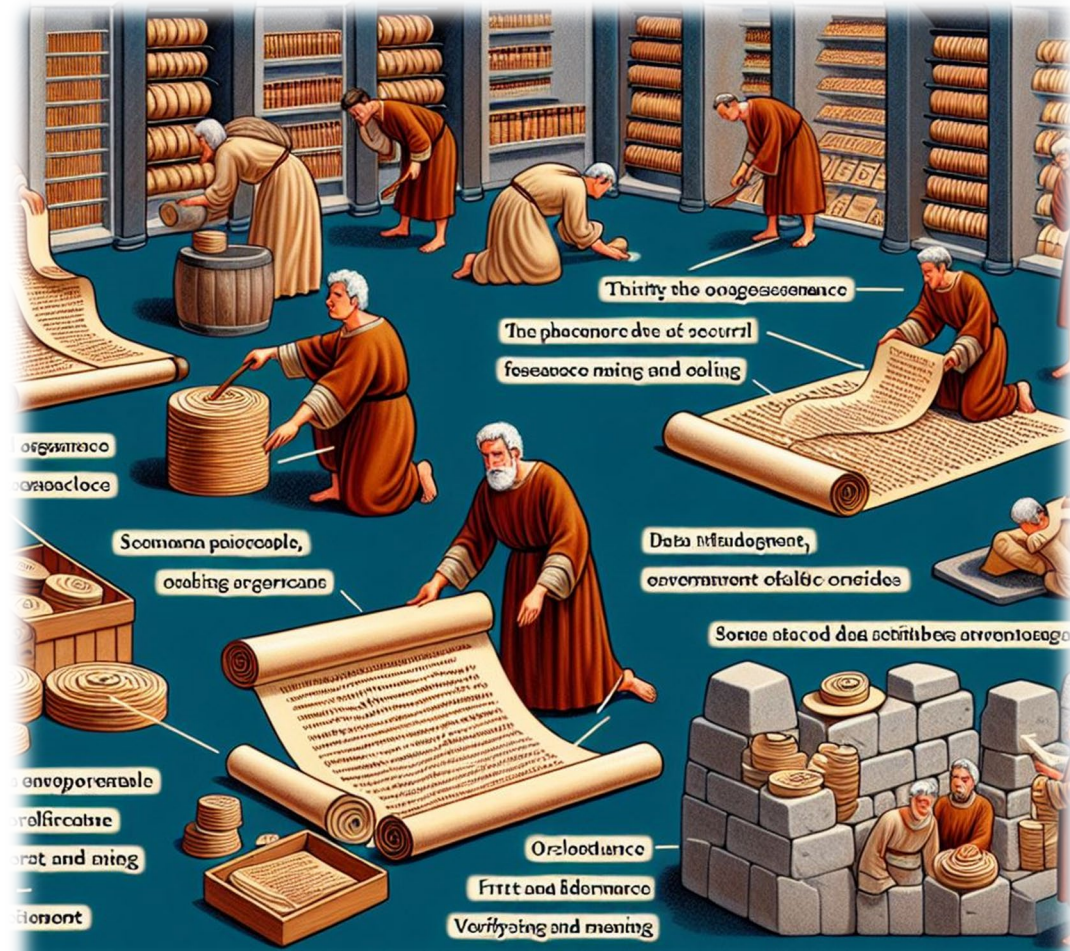
Modern times

- 1960s/70s data **warehousing** & **sharing data** movement begins
- The **Internet** & 'Big Data' 1990s – emerging software and systems for analysis
- **Google search** – sort through huge datasets – the new era of 'data analysis' ('informed decision-making')
- '**Open data**' as a distinct term first appeared in 1995 – US agency's report about geophysical and environmental data disclosure
- **Today: repositories, cloud computing, metadata, ...**



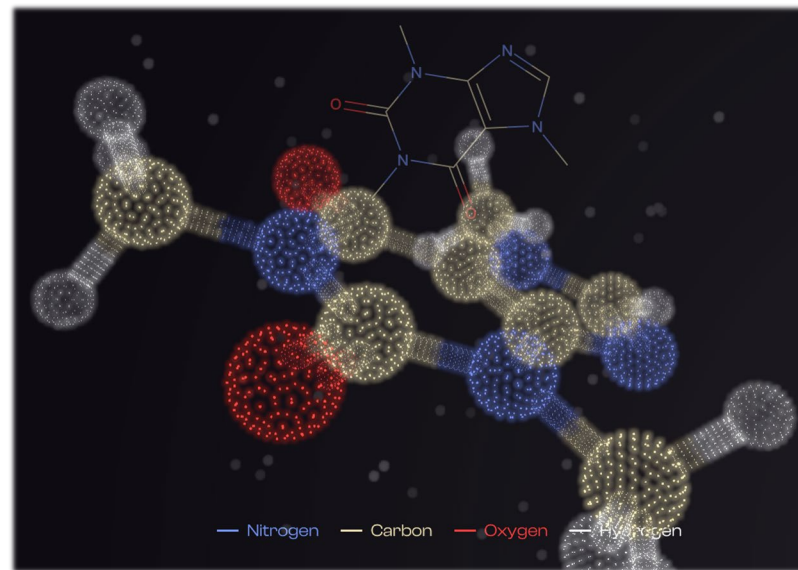
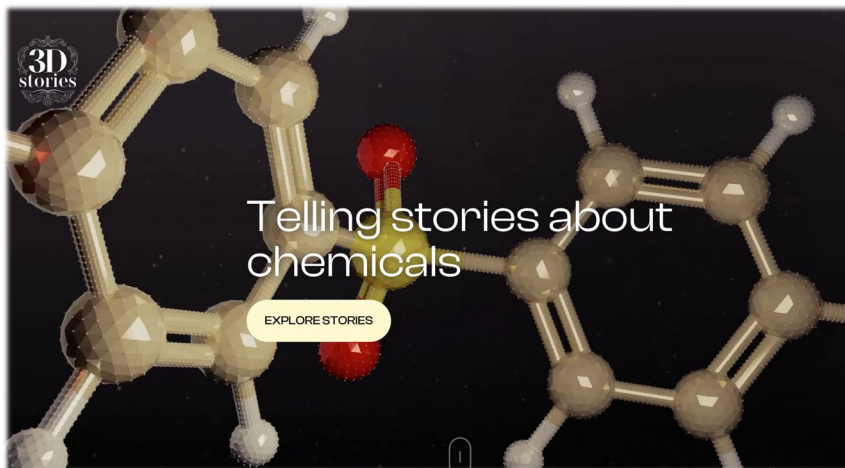
Shared concerns throughout time

- Format/tools go out of date
- Requires skilled specialists
- Ownership & governance
- Storage & archiving
- Threats: Loss, damage, misplacement, destruction
- FAIR principles
- Quality, trust & meaning (the original aims)



University of Luxembourg

- Modern “learning centre” (LLC)
- Open science working group
- ‘Lunchclub’ Community of Practice
- ORBilu – Open repository & bibliography
<https://orbilu.uni.lu/>
- LLC Helpline/support: researchers & students



- PhD inductions & workshops
- Information literacy training
- Love My Data week: March
- Open Access week: October
- Collaboration (UNESCO, Greater Region, Luxembourg)
- Benchmarking



EU

OPEN

DATA

DAYS

19 – 20
MARCH
2025



Thank you!

Dr Helena Korjonen, helena.korjonen@uni.lu

Prof Emma Schymanski, emma.schymanski@uni.lu

Dr Inma Peral, inma.peral@uni.lu

