



# Automated Intelligence (AI), the Internet of Things and Archives

TEACHER RESOURCE AND GUIDE



Supported using public funding by  
**ARTS COUNCIL  
ENGLAND**



# Automated Intelligence (AI), the Internet of Things and Archives

In this document we explore the following three issues:

- 1) Everything digital is incredibly brittle and fragile and destined to break.**
- 2) Digital content is increasing centralised, censored and vulnerable.**
- 3) The roll out of Automated Intelligence (AI) means digital content will likely be used in unexpected and unprotected ways.**

.....



Supported using public funding by  
**ARTS COUNCIL  
ENGLAND**



# 1. The Incredible Fragility of Digital.

To summarise – just about everything digital is destined to break

- **Storage failure**

**Most devices designed to hold digital material fail and fail relatively quickly.**

*“Different media have tremendously varying shelf lives,” explains Dr. Altman, “It’s possible to get an archival optical media and write it in a professional way, so that you would expect it to last 100 years, if it’s stored in the right place. But recording on other things like a random hard drive, a CD, or a flash drive and sticking it on the shelf and you could come back in three years to find significant degradation of data.” Making a backup on a separate disc, or flash drive, or a backup tape — all of these are just fine, but you’re going to need to do it every month and update it,” suggests Dr. Altman*

<https://www.digitaltrends.com/computing/how-do-you-preserve-digital-data-forever/>

- **Externalisation, Corporate Failure and unknown destinations**

Furthermore there is an increasing reliance on external corporations to provide storage platforms and locations, such as ‘The Cloud’ or Instagram, Facebook. This is particularly troublesome as many people upload directly and thus do not even have a duplicate. There is of course no corporation that has lasted forever and very very few last more than 100 years.

*The average lifespan of a company listed in the S&P 500 index of leading US companies has decreased by more than 50 years in the last century, from 67 years in the 1920s to just 15 years today, according to Professor Richard Foster from Yale University.*

<http://www.bbc.co.uk/news/business-16611040>

This means there is significant risk in regards to the knowing what will happen to material, including ownership, appropriation and even its existence. Vine, which collapsed in 2015 has



Supported using public funding by  
**ARTS COUNCIL  
ENGLAND**



for the time period created an archive of the content but there is no further way to interact with the material and no promise to not close it down entirely.

- **Digital is one system. Don't put all eggs in one basket.**

The digital world is based on a single system with ten or less languages which makes it incredibly vulnerable to centralisation, unexpected changes, viruses and also misuse. The system was created and is owned by corporations who have significant and ongoing reach. For example, Microsoft's Windows XP is famously "no longer supported" effectively speeding up the decline and death of the files, software and hardware associated with it. Most software is bought on licence and can be withdrawn at the corporation's pronouncement.

Imagine for example that the Y2K bug wasn't actually sorted out and everything digital from 1976 until this moment was built with a virus in its 'code' in its DNA. Perhaps everything from 1976 this moment will completely disappear? All digital records may disappear? Even the digital archive caught on fire recently. (<http://www.bbc.co.uk/news/technology-24848907>)

.....

## 2. The centralisation and censoring of knowledge on the internet

- **Monopolisation**

We are living in a context today where around 3-5 companies now own the over 90% of the area they operate in – construction, medical equipment, aeronautics, agriculture, pharmaceuticals etc and so on. Sprint, Verizon and AT&T are, for example, the three main companies that own the servers that make up the Internet. Four companies own the domain registration platforms and so it goes on.



Supported using public funding by  
**ARTS COUNCIL  
ENGLAND**



When it comes to searching and providing information, nobody comes even close to competing with Google or should we say Alphabet. Alphabet is close to complete and total domination, with even libraries using Google software for their catalogue searches and organisation. The control and the flow of information are unfortunately increasingly controlled, censored and more concernedly; centralised.

The result of this means information on the internet is increasingly vulnerable to censorship via manipulation of search algorithms, channel and website deletion by corporations and governments. In the last 2 months 19 of the 20 people we regularly listen to on YouTube have been “removed for breaking community guidelines”. Three websites we have regularly read have been removed for the same appalling, unreasonable and un-defendable ‘reason’. There is a noteworthy risk that one’s work could be ‘disappeared’ with no recourse. For example, YouTube has not yet been taken to court over the deletion of videos and channels.

- **Breakdown of Institutions**

The other issue is the breakdown of institutions which leaves little confidence and trust in them doing the right thing. For example, 1) VW deliberately ran a decade-long long conspiracy for which it needed to pay \$4.3billion penalty and 8 senior staff members were jailed, 2) The ‘Big Four’ auditors, who are meant to prevent and reveal corruption have been proven to be involved in these crimes themselves 3) FIFA, the worlds governing body for the sport of football was found to be involved in deep and widespread corruption that included match fixing, corrupt awarding of world cups and bribery. 4) The labelling of food by supermarkets is increasingly difficult to trust, for example TESCO was selling horse meat as beef. Grains and rice have been replaced with plastic and so on. And in India, Nestle was selling noodles via the brand Maggi, which contained seven times the permissible level of lead in them.

### **3. The roll out of Automated Intelligence (AI) and**



## 4. the Internet of Things

Artificial Intelligence (AI) is best understood as Automated Intelligence. Automated means it happens without direction, it is automatic. Computers conceived and developed by corporations have been designed to communicate with each other, which up until this point in their development has been useful.

However the combination of the facts above alongside the roll-out of AI and Big Data means we no longer know how information, particularly personal and political information will be used or misused. Placing information, particularly information which links people to anti-elitism, dissent and non-conformist thinking, would be irresponsible and negligent today.

**There is no way to predict how this information could be used.**

As we mentioned, computers are designed to communicate and the data hacks and breaches reveal that no digital information is able to be protected. Here are some recent examples:

- 1) There has been a list of high profile cryptocurrency hacks,
- 2) Over 140 million financial details (credit scores etc) were hacked from Equifax,
- 3) the Health Insurance information of over 2m people were hacked from Accendo Insurance,
- 4) the emails from 3billion Yahoo accounts were also hacked but the breach was not confirmed until three years later.
- 5) Recently a ransomware infected 47 NHS Trust databases and is understood to have 560 million login credentials worldwide.



Supported using public funding by  
**ARTS COUNCIL  
ENGLAND**



6) This is a list of '26 of the most infamous data breaches'. It is interesting to note in this list '26 of the most infamous data breaches' <https://www.techworld.com/security/uks-most-Infamous-data-breaches-3604586/#r3z-addoor> the breaches are increasing in size and regularity. As shown from the examples above the breaches pertain to all aspects of our life; financial, personal, health, schooling and so on.

It is evident there is significantly less rather than more aptitude and ability today to protect our digital information. The corporate roll out of AI and The Internet of Things furthermore makes this situation ever more problematic.

**As a significant percentage of our archive is other people's content, our solution to this very difficult situation outlined above is to provide 'an envelope' which highlights the information that is available on paper within the Bishopsgate Institute.**

.....

## **Some things to consider when establishing an archive:**

### **DIGITAL ARCHIVE**

#### **The Pros**

- Easy to save and store
- Unlimited storage space.
- Easy to share and make use of.
- Little physical labour required.
- Documents are searchable.
- Digital copies are less likely to be damaged.

#### **The Cons**

- Technological obsolescence: Today's media may be obsolete in 10 years. (Like floppy disks.)
- Scanning old documents can be labour-intensive.



- Good archival formats don't require anything but the human eye to access those records. With digital, you need a digital viewer.
- Reading images is format-dependent. You need to be able to open the saved file formats 10 years from now.
- Few archival records will be natively digital, so digital copies are just that, copies.
- Materials are catalogued together in the archives – putting materials online can take the document out of context

## PAPER/PYSCHIAL ARCHIVE

### The Pros

- Easiest to view.
- Most archive records will be in the original, unaltered form on paper. Original source documents are VITAL for archival purposes.
- Physical files will last longer than people and won't get lost in the world of internet – when you scan stuff it can go missing/ not be scanned/ miss something.
- You can't hack and change physical archives / protection from accidental deletion

### The Cons

- Most expensive to store.
- Harder to search and share.
- Can be fragile
- Can be difficult to stabilize if on acidic paper
- Subject to insect, mould and other biological degradation.

### Risks and considerations if you choose to deposit your archive(s) with an institution:

- While rare archives have experienced accidental and deliberate fire and water damage. In 2013 the Internet Archive in San Francisco suffered significant fire damage.



Supported using public funding by  
**ARTS COUNCIL  
 ENGLAND**



- An institution might close due to changes in funding/political interests. The purpose built Women's Library was shut down and the archives were taken to LSE Library. A nation may be invaded or born and archives shift to a new geography.
- A change in directorship can mean a clear out and reprioritisation of archives.
- 'Stock stacking' can result in archives being discarded and deleted.
- Key words, phrases for labelling can change over time and not be updated.
- Bias directorship could purposefully hide, obscure, censor and mislabel.
- Staffing resources in the future are unknown. Staff can be overworked and underpaid and sometimes disinterested and lackadaisical. Perhaps not able to label, key word and input properly and thus remain unknown.
- Collections are vulnerable to being stolen by corrupt or overworked staff.
- Collections are vulnerable to being forged or having forgeries added as was done with British WW2 files at National Archives.

#### **General considerations:**

- 'Gossip hunters'. Journalists, authors etc who mine archives for gossip/drama.
- Ownership of materials. Future financial benefits from archives use?
- The commercialisation of materials.
- The repackaging, co-opting and colonisation, commercialisation of ideas etc.
- Misunderstandings by future generations.
- Maintaining agreements for private, time delayed etc.
- Biases in software, particularly related to languages and labelling, searches.
- Post-internet and post digital age where everything will be different.