



## **Preservation and Big Data: broad strokes and emerging trend**

**(aka my guide to the future)**

Three narratives of the future

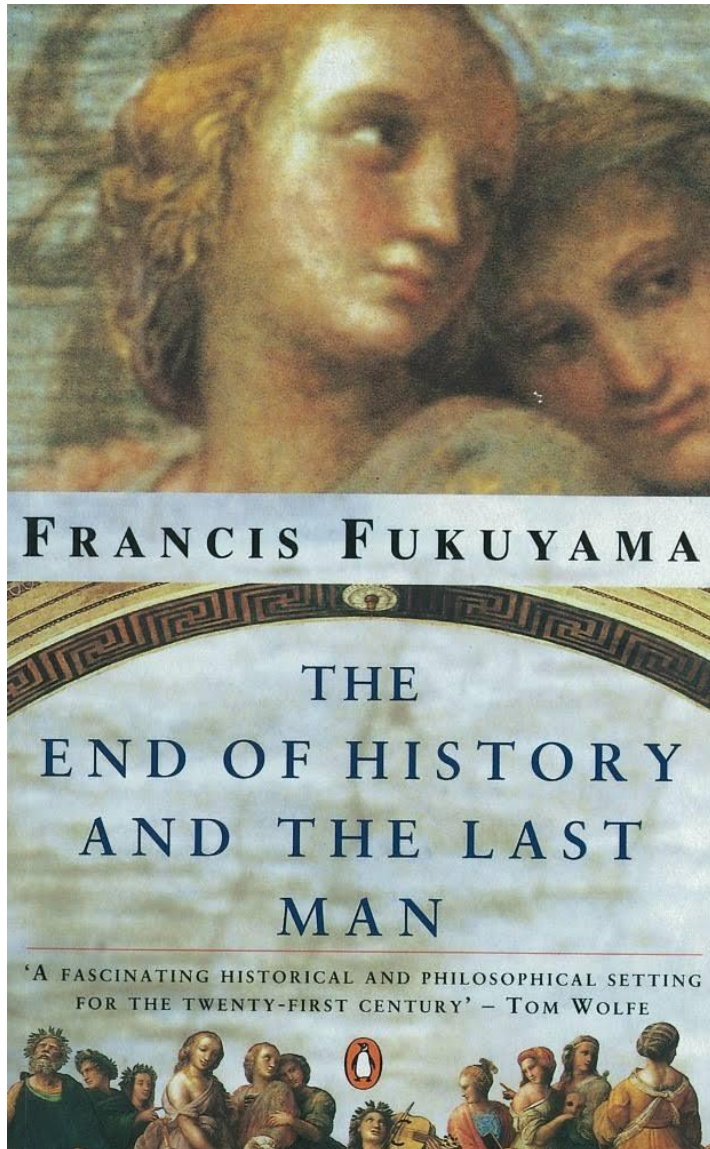
15<sup>th</sup> September 2008

Data: abundant, inexhaustible, frictionless

Obsolescence as an economic function

And transactional data?

Three futures.

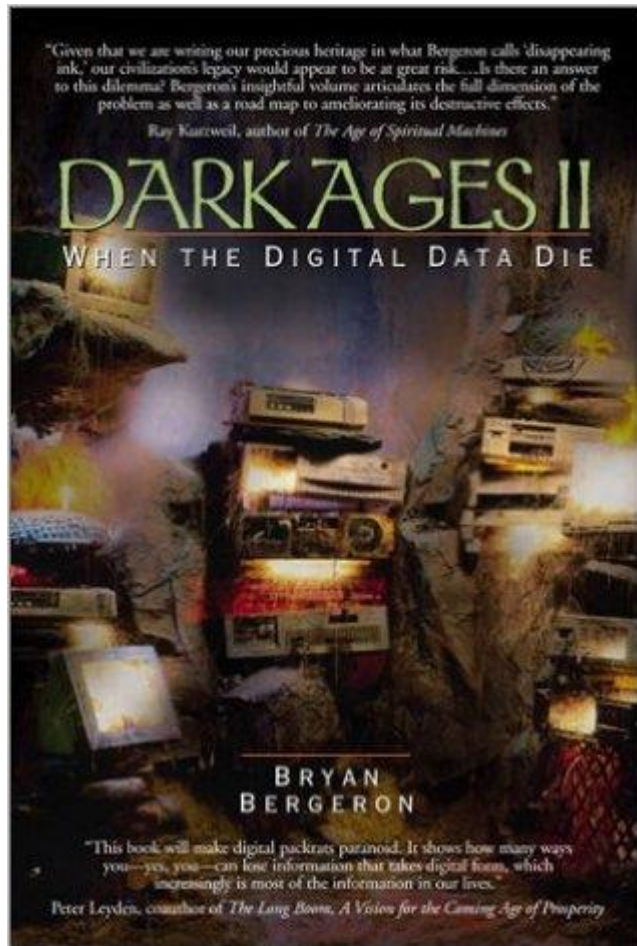


## 1990's: the End of History

*What we may be witnessing is not just the end of the Cold War, or the passing of a particular period of post-war history, but the end of history as such: that is, the end point of mankind's ideological evolution and the universalization of Western liberal democracy as the final form of human government.*

### **Verdict:**

*Wrong!*

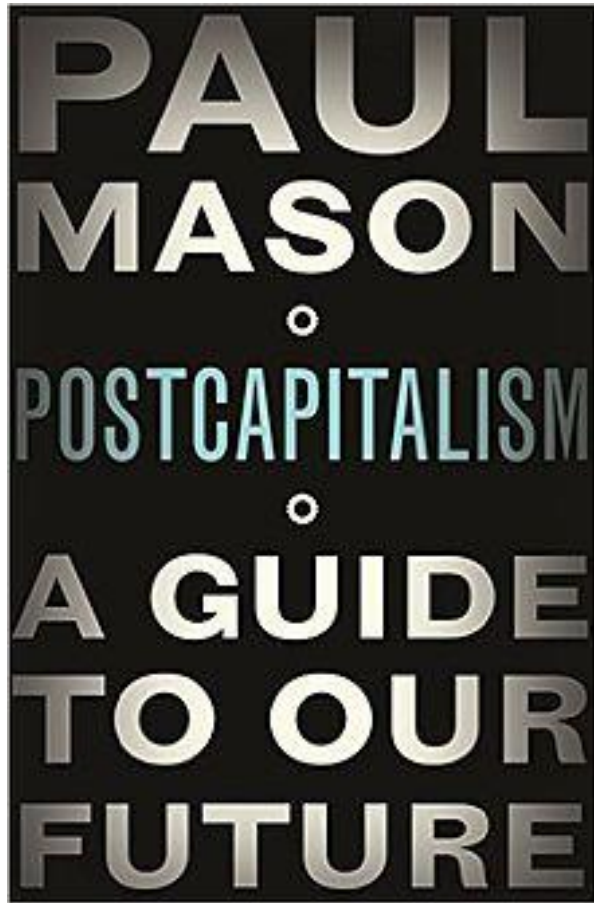


## 2000's: the Digital Dark Age

*'data is at far greater risk than we've ever imagined ... a frightening future, where so much critical information is lost that civilization itself could collapse.'*

**Verdict:**

*Still waiting*



## 2010's: Post Capitalism

*"An economy based on information, with its tendency to zero-cost products and weak property rights, cannot be a capitalist economy."*

### Verdict:

*I don't think so: here's why*



# From the End of History to 15<sup>th</sup> September 2008

*“Capitalism’s failure to revive has moved concerns away from the scenario of a ten year stagnation caused by overhanging debts towards one where the system never regains its dynamism. Ever.*



- Fiat money
- Financialization
- Global imbalances
- Information Technology

Data is the new oil.



**Ann Winblad**

Co-Founder, Managing Director Hummer Winblad Venture Partners

San Francisco Bay Area | Venture Capital & Private Equity

Current Richard M. Schulze Family Foundation, OptiMine Software, Inc., Sonatype

Previous Voltage Security, Karmasphere, SVForum



**Peter Sondergaard**

SVP, Head of Research, Gartner Inc

Greater New York City Area | Information Technology and Services

Current Gartner Inc

Previous Gartner, IDC Scandinavia

Education Københavns Universitet

2nd

<https://www.linkedin.com/in/annwinblad/>

# Data as the new oil: is (big) data the light at the end of the tunnel?

*Abundant*

*Valuable*

*Inexhaustible*

*Frictionless*

*Additive*

@williamkilbride



**Abundant?**

*'It's not like finding a needle in a haystack: it's like finding a needle in Germany.'*

(Paul Miller, DPC Briefing Day on Linked Open Data, July 2013)



G. Bembridge 2012 <https://www.flickr.com/photos/tipsfortravellers/>

## Frictionless?

*Julian Assange*

*Edward Snowden*

*Alexandra Elbakyan*

...

*Institutional Repositories*

*EDRMS*

*Data Sharing*



## Valuable?

Facebook Inc

NASDAQ: FB - 11 Mar 16:00 EST

109.41 USD ↑ 2.09 (1.95%)

1 day 5 day 1 month 3 months 1 year 5 years max



Open 108.48  
High 109.42  
Low 107.92

Mkt cap 310.86B  
P/E ratio 84.82  
Div yield -

Google Finance - AOL Money - Yahoo Finance

Disclaimer

*Cash Value: \$6.3bn*

*Share Valuation: \$104bn*

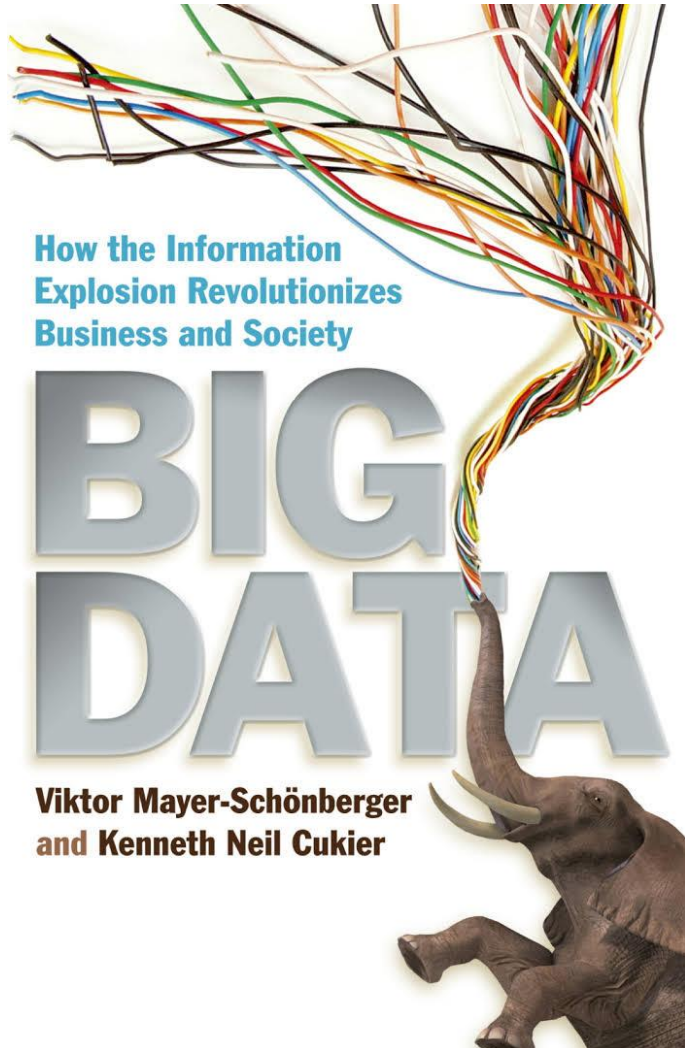
*Intangible assets: \$97.7bn*

*Pieces of data: 2.1trillion*

*Productive capacity: nil*

*Production costs (to FB): nil*

## Additive?



*“...The crux of data’s worth is its seemingly unlimited potential for re-use: it’s option value. ...most of data’s value lies in its use not in its possession”*

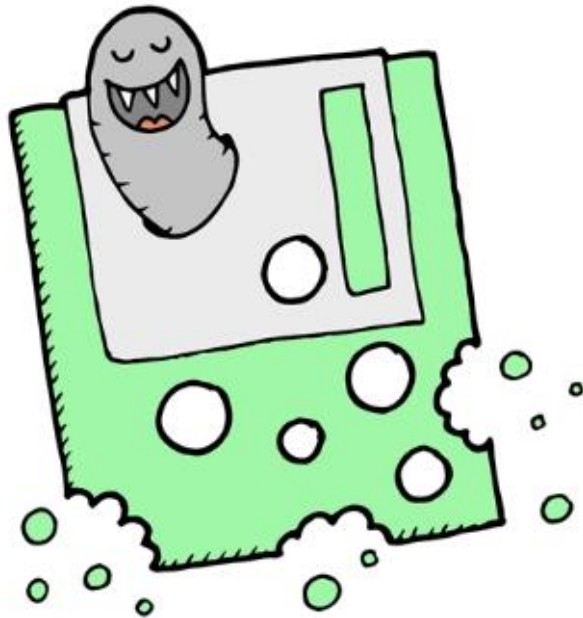
(Mayer-Schonberger and Cukier 2015)

...Digital  
Preservation  
Claxon...



Inexhaustible?

**Digital preservation makes bleak reading**



*<Enter your details here>*

.....

.....

.....

## Inexhaustible?



*“...He (Evan Spiegel) likes to live in the present”*

(Wall Street Journal June 2015)

# What's the problem?

Media failure

Media obsolescence

Software failure

Software obsolescence

Inaccessible file format

Format migration errors

Loss of media

Disaster

Broken dependency

Failure of systems integration

Data Theft

Virus or malware

Poor documentation

Poor rights management

Lost encryption key

Poor version control

Inadvertent deletion

Malicious deletion

Poor file-naming

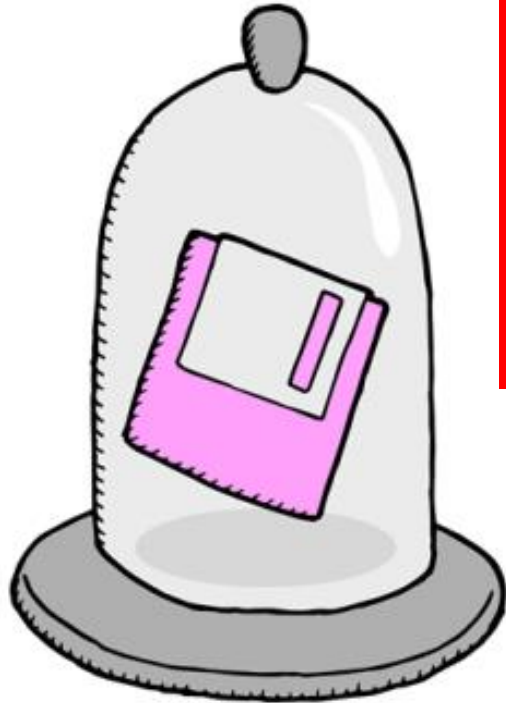
Corruption

Link rot

...

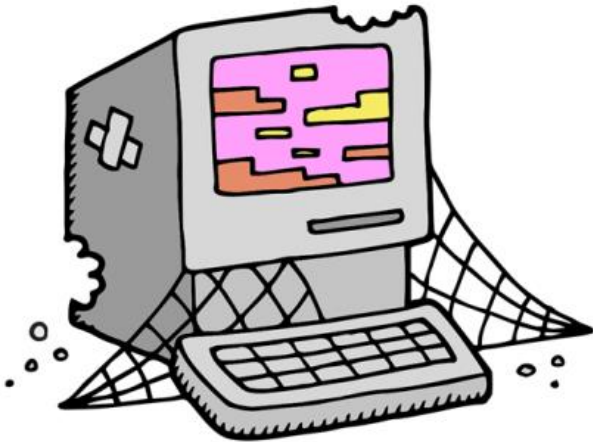
Actual Experience?

**CORPORATE ABANDONMENT  
IS AS BIG A THREAT  
TO THE DIGITAL ESTATE  
AS OBSOLESCENCE**



Nook	Friends ReUnited	Yahoo 360
GoogleCode	Yahoo Mail Classic	del.icio.us
GeoCities	Blipfoto	MyBlogLog
GoogleWave	MySpaceBlogs	BeBo
knol	...	...

## Interdependence as risk



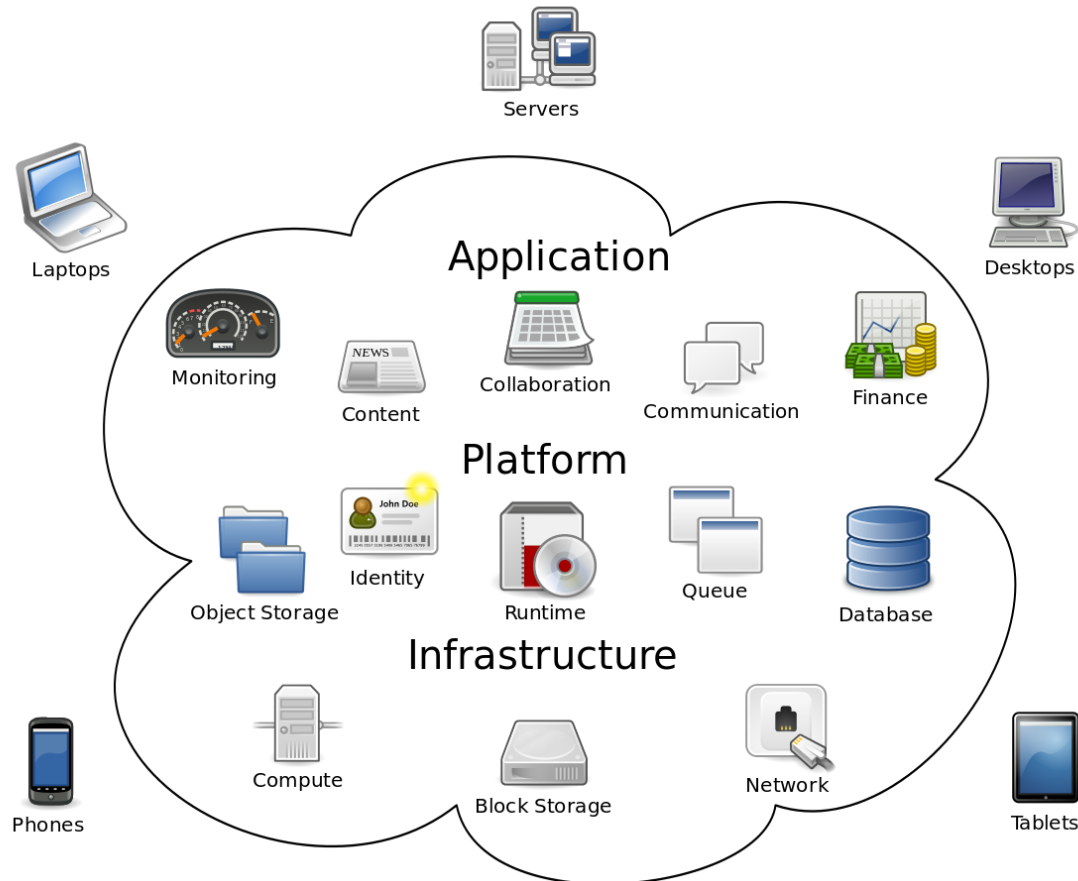
✋♂ ♀❖♂□△◆⚡⌘■♂ ⌘◆  
☹ ◆♂□❖⌘♂♂♂  
♂❖♂□△◆⚡⌘■♂ ⌘◆ ☹  
♂□■◆□☹♂◆  
♂❖♂□△◆⚡⌘■♂ ⌘◆ ☹  
❖◆●■♂□☹♂⌘◆△

or

**If everything is a service,  
everything is a contract,  
everything is a vulnerability**



# Cloud Computing



Cloud computing is the delivery of computing as a service rather than a product



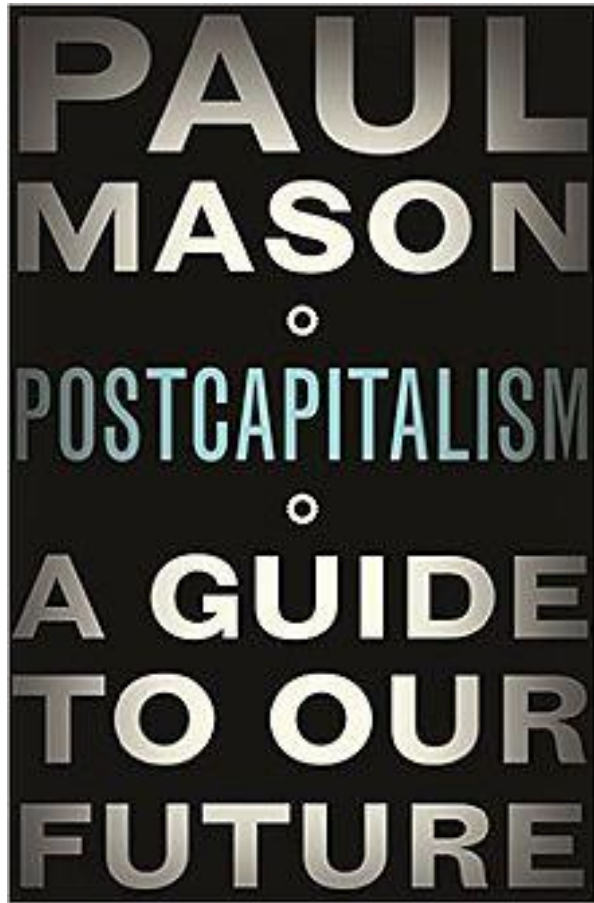
Why this matters?

Long chains of interdependence

When a small glitch or upgrade changes our view on data and the real world impact this has

- Stephen House
- HSBC
- Google Fonts

Obsolescence, link rot and platform instability are practical expressions of neoliberal economics.



## 2010's: Post Capitalism

*"An economy based on information, with its tendency to zero-cost products and weak property rights, cannot be a capitalist economy."*

### Verdict:

*There are problems with the economic analysis: but more than anything, he takes the data for granted. No clear sense of what data is, how it is created and how it exists.*

# From the End of History to 15<sup>th</sup> September 2008



*“Capitalism’s failure to revive has moved concerns away from the scenario of a ten year stagnation caused by overhanging debts towards one where the system never regains its dynamism. Ever.*



- Fiat money

- Financialization

- Global imbalances

- Information Technology



The Future: Transactional data.

Unless we figure out how to secure this data for access and re-use in the long term we simply cannot derive long term value from it.

Ignorance is no longer an excuse. Sara has taken care of that.

The implications of failure are not a digital dark age. The implication of failure is to extinguishing the one glimmer of light in the darkness already upon us.

Agencies that invest in the development of data but which don't have policies regarding the long-term value and exploitation of these data sets are derelict in their duties.



## Preservation and Big Data: broad strokes and emerging trend

Three narratives of the future

15<sup>th</sup> September 2008

Data: abundant, inexhaustible, frictionless

Obsolescence as an economic function

And transactional data?