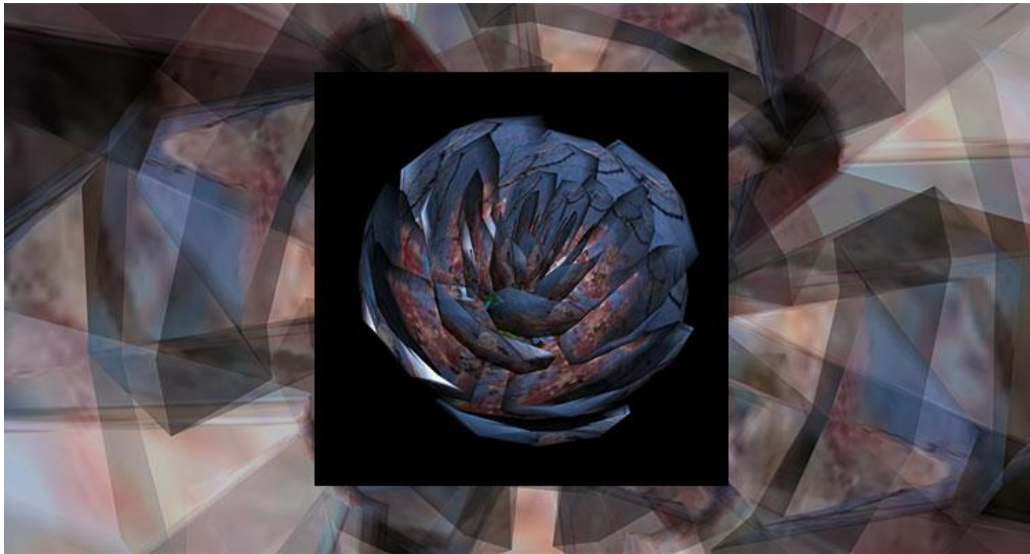


Big Data Requirements in Arts and Humanities

Andrew Prescott,
King's College London



Michael Takeo Magruder, *Data Flower (Prototype One)*, 2010

Is there big data in the arts and humanities?

106
73 Hurley Street
July 30, 1878

Dear Lord Beaconsfield,

I find you are reported in the Times of today to have made last night a reference to a speech delivered by me at Oxford, in which you state that I "described you as a dangerous and even despicable character."

I shall be obliged by your informing me on what words of mine you found this statement.

Letter of Gladstone to Disraeli, 1878: British Library, Add. MS. 44457, f. 166

The political and literary papers of Gladstone preserved in the British Library comprise 762 volumes containing approx. 160,000 documents.



Research

ELECTRONIC RECORDS

IN THIS SECTION

Textual Materials

Finding Aids

Featured Document

Archived White House Website

Audiovisual Materials

Electronic Records

White House Email

Gubernatorial Records

Laws & Regulations

Other Research Resources

Research FAQ

The George W. Bush Presidential Library and Museum holds approximately 80 terabytes of electronic records — the largest electronic records collection of any Presidential Library.



President George W. Bush and Israeli Prime Minister Ariel Sharon laugh together during their joint press conference, July 29, 2003, in the Rose Garden of the White House.

Courtesy George W. Bush Presidential Library and Museum. (P32655-32)

This data includes over 200 million email messages created or received using the White House email system; nearly 4 million photographs created by the [White House Photo Office](#); the files from the shared network drives used by various offices in the White House; and the data from the various records management, scheduling, and appointment systems used during the Administration of President George W. Bush. The vast majority of these electronic records are stored in the National Archives and Records Administration's [Electronic Records Archive \(ERA\)](#). Electronic records not stored in ERA are physically housed by

the National Archives at either the Library and Museum or at the National Archives' headquarters in Washington, D.C.

Print + Bookmark/Share



September 11, 2001

George W. Bush Presidential Library:
200 million e-mails
4 million photographs

Big Data Types in the Humanities

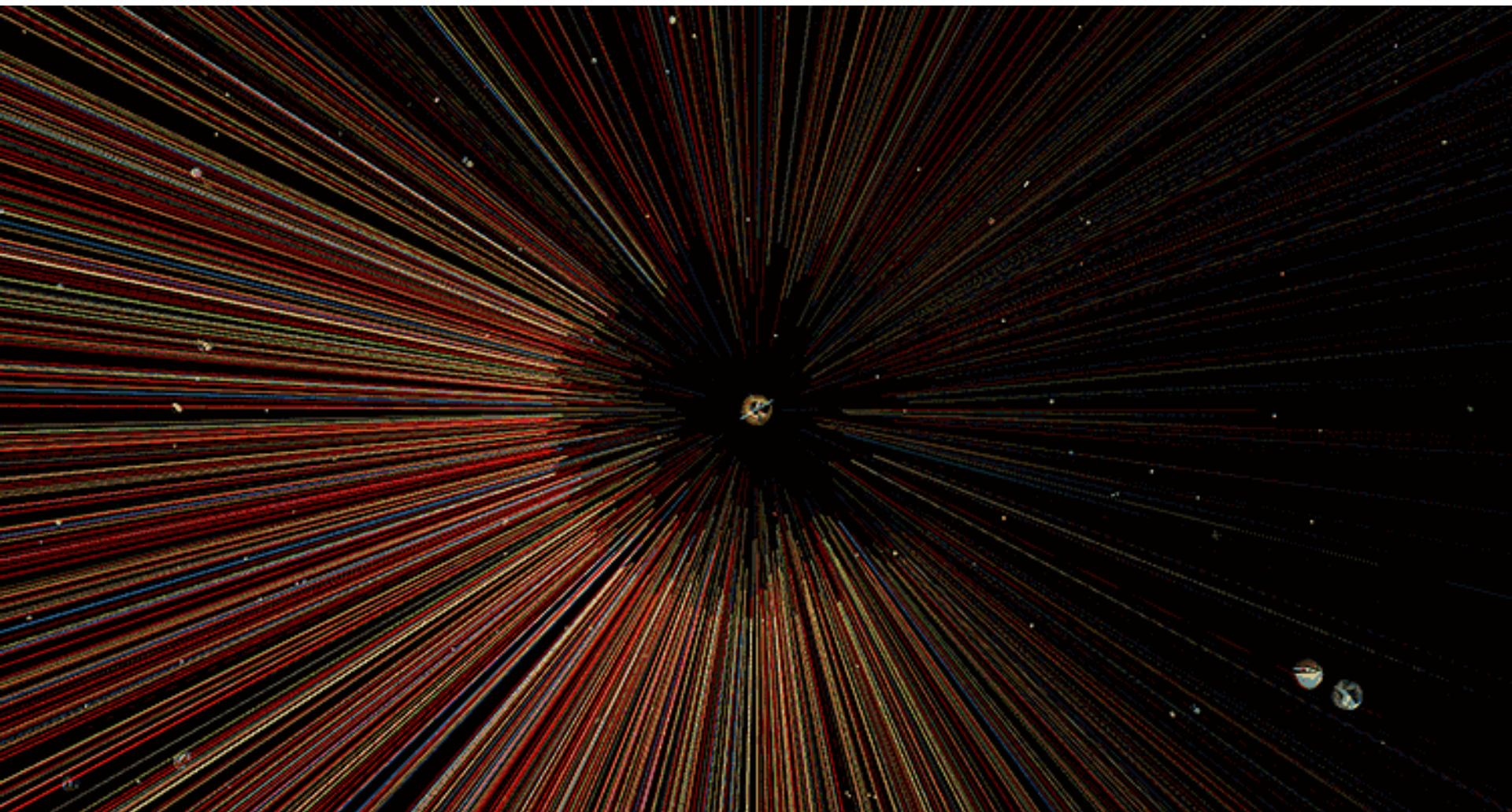
- Strong cross-over with social science data needs: retail, transport, utility, government data all likely to be useful – particularly important is the concern over time (historians want very old data)
- Humanities will contribute to creation of big data sets across disciplines e.g. Old Weather
- Long-standing engagement in some disciplines with big data eg linguistic corpora, library and archive catalogues

Big Data Types in the Humanities

- Book scans: Google, Internet Archive. But big data in humanities not just text.
- Image data – ranging from Rembrandt to born digital
- Sound
- Moving images
- Cartographic data increasingly important (again historical dimension important)

Big Data Types in the Humanities

- Complex data objects will become of increasing importance and will pose very challenging issues
- 3-D archaeological reconstructions
- Virtual environments
- Games
- Social media
- Digital artworks

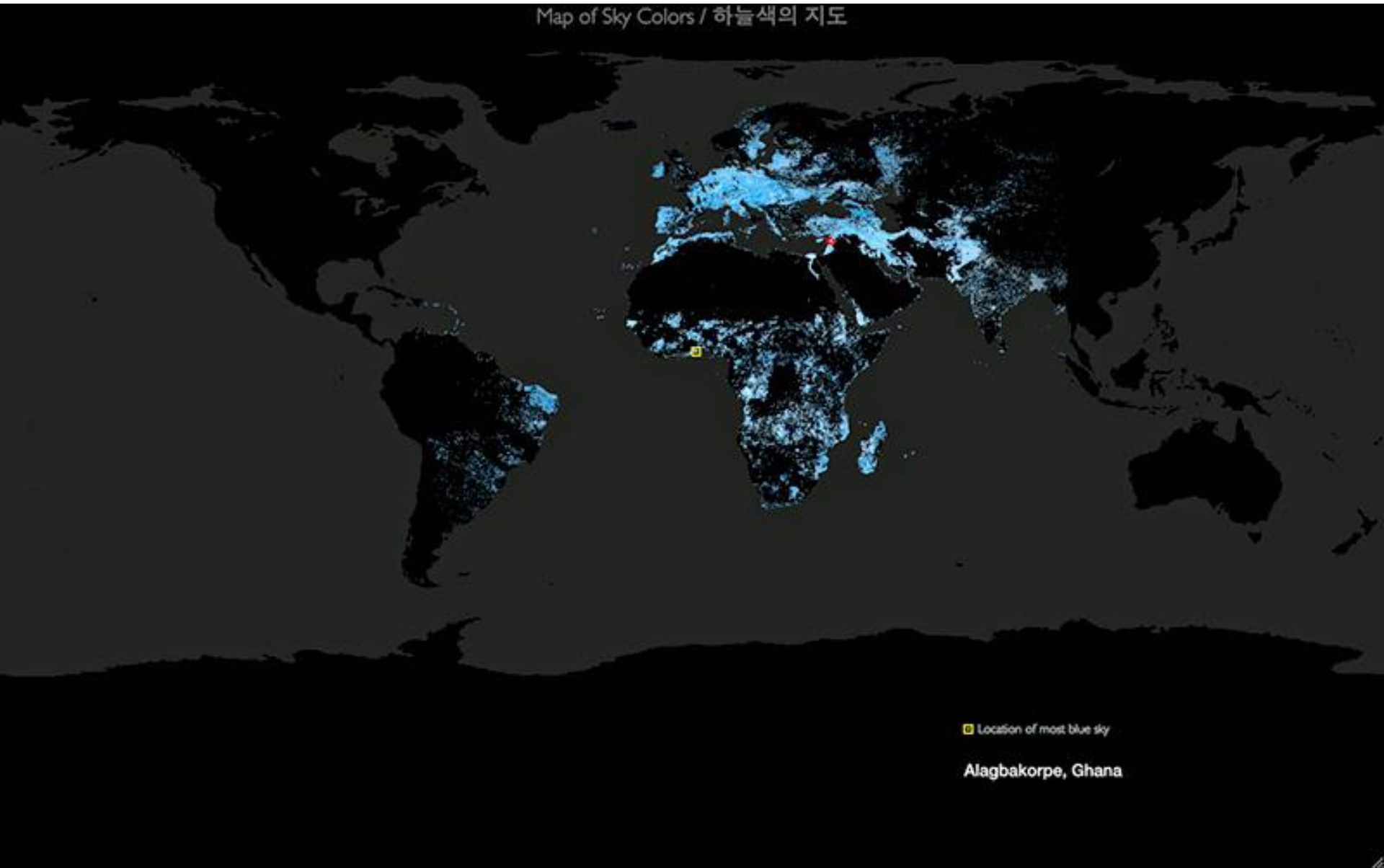


Michael Takeo Magruder, *Data Sea* (2010): <http://www.takeo.org/nspace/ns030/>



Fabio Lattanzi Antinori, *The Obelisk* (2012): Open Data Institute:
<http://www.theodi.org/culture/obelisk-2012>

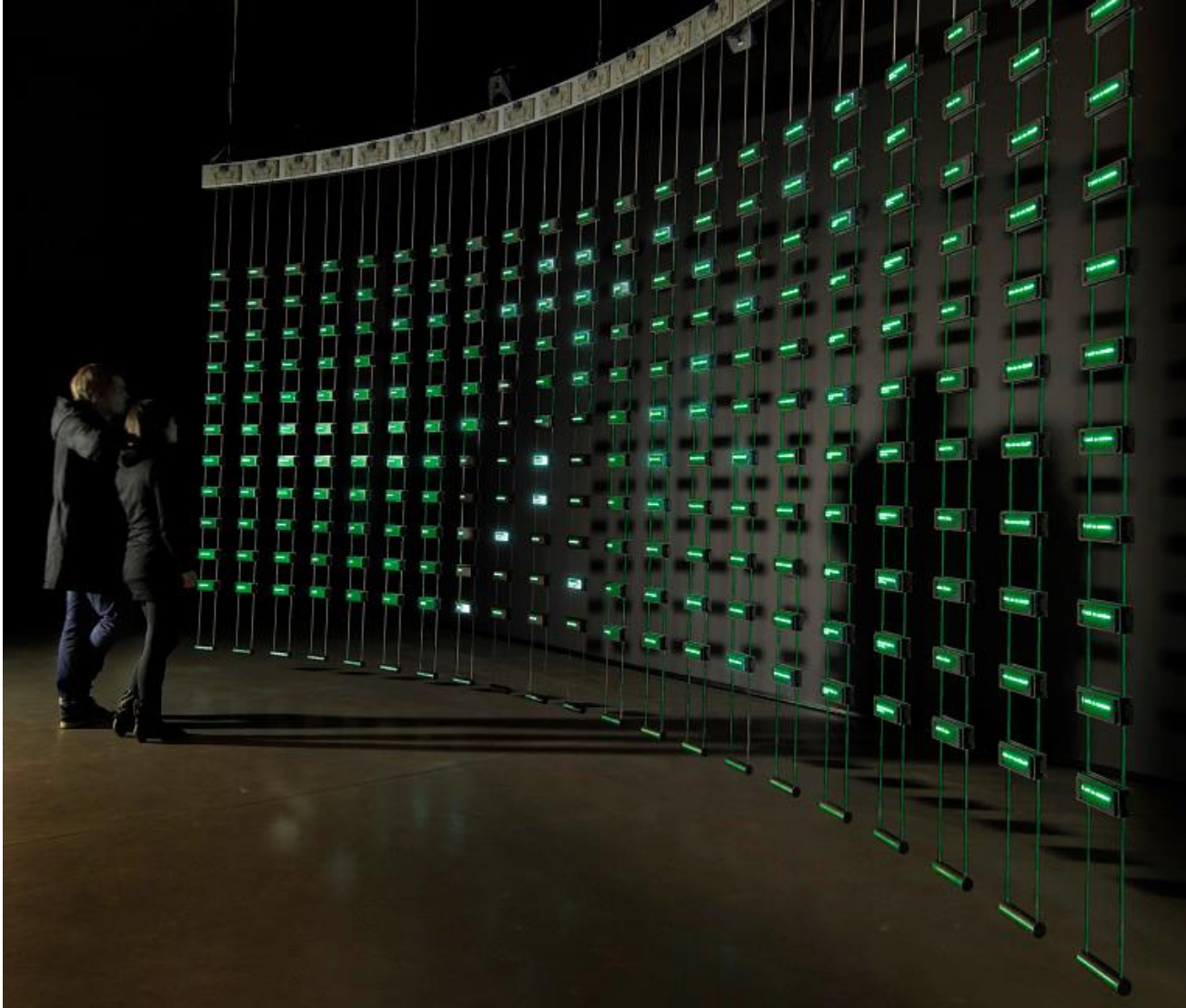
Map of Sky Colors / 하늘색의 지도



□ Location of most blue sky

Alagbakorpe, Ghana

Lise Autogena and Joshua Portway, *Most Blue Skies*, 2010: <http://www.autogena.org/mbs.html>



Mark Hansen + Ben Rubin, *Listening Post*
(2002-6)

Big Data Issues in Arts and Humanities

- There has been a strong engagement with curation issues, but limited achievement with infrastructure
- Data objects are likely to be complex in character (e.g. artworks, virtual worlds) posing curation issues
- Digitised material still significant (family history, newspapers, multi-media), so IP/proprietary issues significant
- Infrastructure issues: skills at all levels, access to equipment and network, culture of lone scholar and lead academic

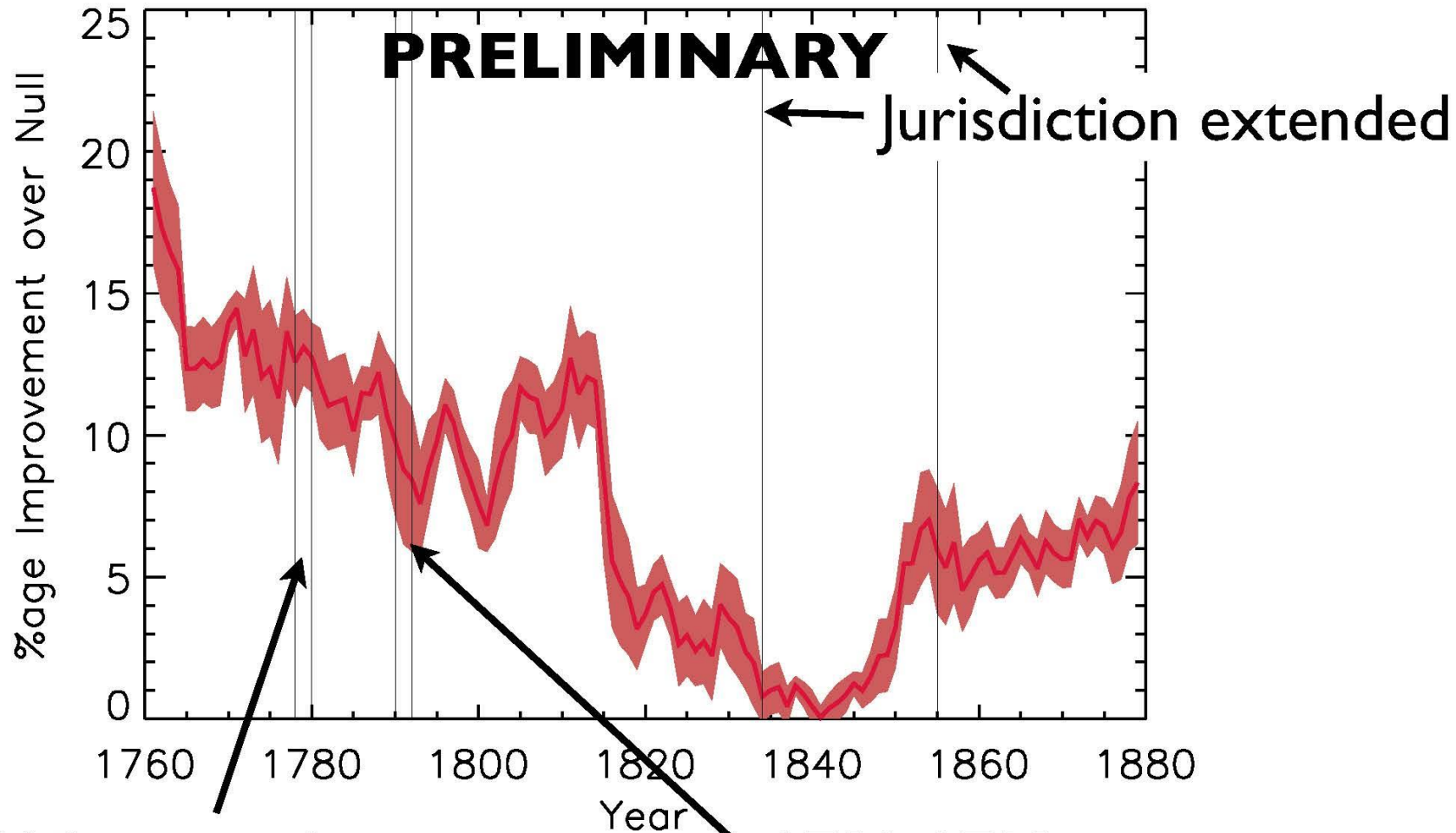
Big Data Issues in the Arts and Humanities

- Complexity deluge rather than data deluge
- Awareness of the issues posed by provenance of data
- Need to investigate layers of data (complicated by need to examine data across time)
- Methodological uncertainty: need to drill down seen as just as strong as looking at wider picture
- No tradition of data-driven research
- Where does probability and prediction fit in?
- Worries about completeness and accuracy (small data mentalities)
- But on the other hand, arts and humanities bring a range of critical tools to bear

Big Data Opportunities in the Arts and Humanities

- Use data to identify problems that we hadn't seen before
- Use new types of data in a more interactive fashion eg film, moving image, sound – this raises new types of research question
- Open up non-canonical material
- Digital arts can help create new types of interface and visualisation: reconnection of arts and sciences
- Tim Sherratt: 'It's about transformation. It's about taking cultural heritage collections and changing them. Changing what we can do with them. Changing how we see them. Changing how we think about them'.

Predictability of Trial Outcome



John Wilkes mandates
“true, fair & perfect”

1790-1792
Censorship

the real face of white australia

home • about



Form No. 21. COMMONWEALTH OF AUSTRALIA. No. 10/383
DUPLICATE Immigration Returns, Birth Registrations and Expatriation

CERTIFICATE EXEMPTING FROM DIGTATION TEST.

I, *Nicholas C. Lockyer*, the Collector of Customs for the State of *New South Wales* in the said Commonwealth, hereby certify that *Kim Chong Nam*, hereinafter described, who is leaving the Commonwealth temporarily, will be excepted from the provisions of paragraph (a) of Section 3 of the Act if he returns to the Commonwealth within a period of *three years* from the date *26 Nov 08*.

DESCRIPTION
Name: *Kim Chong Nam* **RECEIVED**
Age: *43 years* **EX-100**
Height: *5'7 1/2 in* in Both Feet **100**
Build: *Medium* Eyes: *Dark*
Particular marks: *None visible* **100**
(For impression of hand see back of this document.)

Full Face: — Profile: —

Date of departure: *11/11/08* Destination: *China*
Ship: *Empire*
Date of return: *5-11-08* Ship: *Empire*
Port: *Sydney* *J. J. Donohue*
Customs Officer