🏠 MUCH ADO ABOUT DESCRIPTION

Federated Searching, Finding Aids, Indexing, AI, & All That Stuff
Banging on about things I still care about (for some reason).

MUCH ADO ABOUT DESCRIPTION

ॐॐॐॐॐॐॐ

❧❧❧❧❧❧❧

Maps
A funky way of describing the world around us.

❧❧❧❧❧❧❧

The DAD: the Daddy-of-All-Descriptions
The indispensable need for a Datum to understand the Data

❧❧❧❧❧❧❧

Federated Searching, Finding Aids, Indexing, AI, & All That Stuff

2015, September 3: Ambient findability

This term has just come up in discussion on the NZ List. I hadn't come across this book before but (for obvious reasons) I think I'll buy it if it's still available. Of course, we've known about situating knowledge in ambient space for quite a while.

Ambient Findability: What We Find Changes Who We Become by Peter Morville (Goodreads Author) 3.74 of 5 stars 3.74 · rating details · 821 ratings · 72 reviews

How do you find your way in an age of information overload? How can you filter streams of complex information to pull out only what you want? Why does it matter how information is structured when Google seems to magically bring up the right answer to your questions? What does it mean to be "findable" in this day and age? This eye-opening new book examines the convergence:
- How do you find your way in an age of information overload?
- How can you filter streams of complex information to pull out only what you want?
- Why does it matter how information is structured when Google seems to magically bring up the right answer to your questions?
- What does it mean to be "findable" in this day and age?

This eye-opening new book examines the convergence of information and connectivity. Written by Peter Morville, author of the groundbreaking *Information Architecture for the World Wide Web*, the book defines our current age as a state of unlimited findability. In other words, anyone can find anything at any time. Complete navigability.

Morville discusses the Internet, GIS, and other network technologies that are coming together to make unlimited findability possible. He explores how the melding of these innovations impacts society, since Web access is now a standard requirement for successful people and businesses. But before he does that, Morville looks back at the history of wayfinding and human evolution, suggesting that our fear of being lost has driven us to create maps, charts, and now, the mobile Internet.

The book's central thesis is that information literacy, information architecture, and usability are all critical components of this new world order. Hand in hand with that is the contention that only by planning and designing the best possible software, devices, and Internet, will we be able to maintain this connectivity in the future. Morville's book is highlighted with full color illustrations and rich examples that bring his prose to life.

Ambient Findability doesn't preach or pretend to know all the answers. Instead, it presents research, stories, and examples in support of its novel ideas. Are we truly at a critical point in our evolution where the quality of our digital networks will dictate how we behave as a species? Is findability indeed the primary key to a successful global marketplace in the 21st century and beyond. Peter Morville takes you on a thought-provoking tour of these memes and more -- ideas that will not only fascinate but will stir your creativity in practical ways that you can apply to your work immediately.

Note (2025): This book chimes into other postings on the theme of *structure vs ubiquity*.

2015, September 3: Location data

Article in the Conversation about uses being made of location data.

Not being a part of cutting edge thinking taking place on description nowadays (assuming such thinking is taking place somewhere or other) I don't know if this idea is already passé, but it occurs to me that utilising location data instead of taxonomical control over location access points and metadata in archival descriptions (any kind of recordkeeping for that matter) would be a good idea. The problems of controlling the language of place (name

change, scaleability, variant usage, dialect, cultural mores, etc.) are notorious. To replace all that with a link to the very place itself would cut out a lot of that – except, of course, when you are mistakenly taken to Hobart, Wisconsin, instead of Hobart, Tasmania.

2016, September 12: Posting hyper links

Court of Justice of the European Union
| Violations of copyright | No 92/2016 8 September 2016 |

Judgment of the Court in Case C-160/15 | The posting of a hyperlink on a website to works protected by copyright and published without the author's consent on another website does not constitute a 'communication to the public' when the person who posts that link does not seek financial gain and acts without knowledge that those works have been published illegally ....

Phew!

<<**Andrew Waugh: I don't know what was in the rest of the Archives Professionals list item, but this is not a completely accurate description of the court's finding.**
        See http://curia.europa.eu/jcms/upload/docs/application/pdf/2016-09/cp160092en.pdf **for the details.**
        The defendant was a Dutch media company. They published an article that linked to two photographs. The web site that had published the photographs did not own copyright to the two photographs. The copyright owners asked the defendants to remove the links. The defendants refused. The copyright owners served a takedown notice on the web site that was publishing the photos (probably at the same time and independently of the contact with the defendants). The web site complied. The defendants then edited their article to point to another site that was still hosting the photos. The copyright owners served a takedown notice on the second web site (which complied) and the defendants (which didn't). The copyright owners then sued the defendants for copyright infringement.
        The media company was apparently relying on a well established principle that making a link to content on the web does not infringe copyright. This principle was established when media companies took aggregation companies to court over linking to their news articles (the media companies lost).
        The court held that linking to an unauthorised publication is quite different to linking to an authorised publication. A person or organisation that knowing links to an unauthorised publication is liable for copyright infringement.
        This raises the question about how a person is to know if a publication on the web is authorised or not.
        First, the court decided that if the linker knows (or reasonably should know) that the publication is unauthorised, then the linker is liable. This applies to everyone - private individual or company - and it doesn't matter if you link for profit

or not. In this case the defendants had been served twice with a takedown notice, so should clearly have reasonably known that the publication was unauthorised. The court was also at pains to note that if a link allows you to bypass a paywall (or other barrier intended to keep content only for authorised people), you should know that the publication is unauthorised.

Beyond this, the court held an additional barrier for people or organisations linking to items for profit. In this case, the court said that they have the resources to investigate whether the original publication is authorised or not. In this case the onus of proof is reversed. If you link to an unauthorised publication, the court will assume that you knew it was unauthorised, and you will have to rebut this assumption.

As a not-for-profit linker, you are given a small amount of leeway. You don't have to show that you checked whether a publication was authorised or not. But you will be liable for copyright infringement if the court finds that you knew (or should have known) that the publication was not authorised.

This is quite a change in the previous position.

(Bear in mind that I'm not a lawyer, nor do I play one on TV.)>>

2018, October 24: Documenting Australian Society

<<Michael Piggott: The following notice posted on behalf of the UNESCO Australia Memory of the World Committee.>>

UNESCO Australia Memory of the World Committee announces a Summit. The AMWC web site lists the composition of the Committee and sets out its mandate –

- develop the Australian Memory of the World Register
- propose nominations to the Memory of the World International Register
- promote the Program through publications and presentations
- encourage and seek government and private sector sponsorship for specific projects and activities
- participate in selected heritage activities and be an advocate for the documentary heritage sector.

Presumably this initiative is a "selected heritage activity".

The background paper asks (p.2) "What is the problem…?" Good question. It appears that the focus is on a lack of co-ordination in the collection and preservation of and access to documentary heritage and documentary heritage is defined as documents of significant and enduring value. What about documents of vital immediate concern that have no lasting and enduring value? All this translates into a concern with the activities of collecting bodies (given a wide definition). Apart from a detour into the process of appraisal, there are only one or two (possibly ambiguous) references to "documentation strategies" (p.2):

> Are there time periods, issues, communities, minorities and phenomena which urgently need targeted documentation strategies? (p.2) … If a specific issue remains unrecorded, a community undocumented or a nationally significant individual's recollections not captured, downstream processes like digitisation and metadata tagging are irrelevant …

but the term is not included in the scope and definitions section (p.3). The ordinary meaning of "documentation strategy" extends to gaps in the formation process but the whole tenor of this document suggests that the intended meaning is gaps in the collection process of what is already formed. In short, the "problem" is how to improve and co-ordinate the gathering of resources not their formation. The focus is on institutions and processes involved in the formation of collections rather than the formation of documentary heritage in the larger sense, much less on the place of documentation in Australian life. If that's the way you frame the problem, then that determines the kind of solution you'll come to. If that is a misreading, then it is certainly one that fits with the language used and it is one that needs much re-drafting to clarify:

… a wide range of organisations, institutions and initiatives that are committed to enabling the long-term preservation of and access to (p.3) … a multiplicity of institutions, initiatives and programs (p.4) … the collective stock of documentary heritage is preserved and added to primarily by publicly-funded, university, community and other libraries, archives and related institutions. Material of unknown quantity and significance is also held in private hands, including in company and organisational archives that may or may not provide public access (p.4) … Complementing [them] are over three thousand local library and local museum collections, historical societies, halls of fame, mechanics institutes, oral history groups, heritage centres, universities' archives and special collections, pioneer associations and keeping places (pp.4-5) …*Significance 2.0* … is an important and useful tool for assessing the value and utility of existing documentary holdings and potential future additions to the distributed national collection (pp.5-7) …

All of which colours the very first of the major issues (p.7) viz. "… what must be documented, identifying high-risk areas that need attention, setting priorities and who to involve/engage in the process?" Very good question.

Gathered   **Ungathered**

I think the prior question is identifying the mind-set that went into the drafting of this document and attempting to suppress it in the discussions that ensue. Collecting may be part of documenting Australian society but it certainly isn't all that's involved. Not by a long chalk. How is Australian society documented? By whom? By what means? How do we identify what exists and gaps in the formation process? How do we fill gaps in the formation process? How is documentation that exists made accessible (irrespective of whether or not it's been gathered)? What ought to be accessible, on what basis, under what principles? How do we keep that knowledge, about the knowledge, up-to-date and accessible?

Don't assume that collecting is an answer to anything of any great importance. If "collecting" is the answer, what was the question again?

2023, June 3:

A Saturday reflection

Because I am out of things these days, I must be forgiven if I draw attention from time to time to things everybody already knows about anyway. In 2014, I referred to SNAC (<u>Social Networks and Archival Contexts</u>) in my <u>presentation</u> to ASA/ARANZ Conference in Christchurch (Figure Ten). A little later a friend asked me if I'd ever heard of it. *Yes*, I said stuffily, *and I approved of it*. And she had been in that audience!!! So much for the imperishability of authorship.

SNAC is collaborative, open-ended, and its not one of those projects people have to fret over until it is top-heavy and collapses. It does what it can, and does it very well.

> **[It] is a free, online resource** that helps users discover biographical and historical information about persons, families, and organizations that created or are documented in historical resources (primary source documents) and their connections to one another. Users can locate archival collections and related resources held at cultural heritage institutions

around the world. **SNAC is an international cooperative** including, but not limited to, archives, libraries, and museums, that is working to build a corpus of reliable descriptions of people, families, and organizations that link to and provide a contextual understanding of historical records. A complete list of SNAC Cooperative members can be found here.

The list of members includes mostly research collections but also Archives nationales de France, Arquivo Nacio al, Brasil, Library of Congress, National Archives and Records Administration, New York Public Library, Utah State Archives. Nothing from Asia, the Pacific, or Australasia that I could see. It may be to "help users" but it is complied by describers and could obviously also be useful for our work.
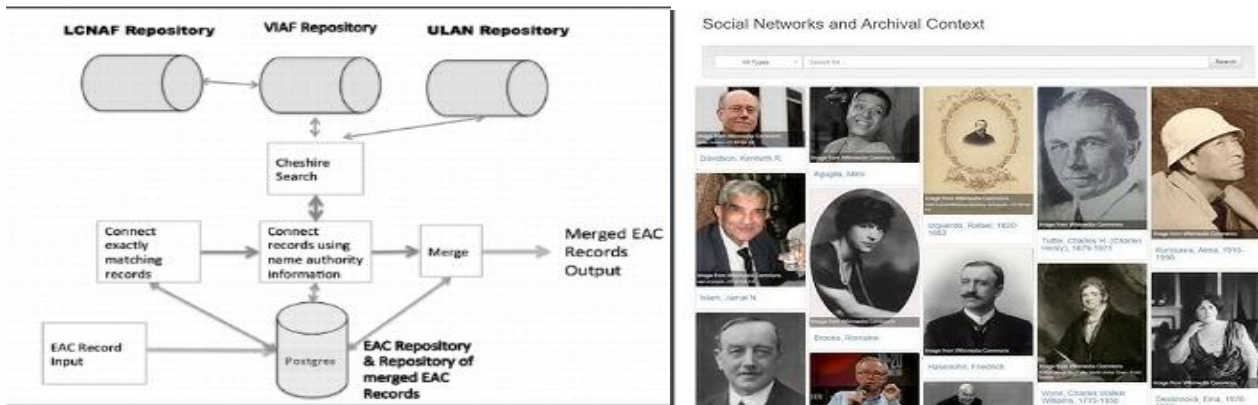
SNAC (snaccooperative.org)

Having despaired of anyone showing interest in the *Modest Proposal* (let alone taking me up on it) I would just like to say that SNAC could, in my view, be used to achieve many of its purposes. It seems to me that it would also be a useful framework tool in the hands of those folks I hear about periodically from over the horizon nattering about *Documenting Australian Society*. The great thing about SNAC is (so far as I can see) that participation **wouldn't** need meetings, plans, collaboration, sponsorship, or direction (all those things that seem to stand in the way of getting anything done and things I tried to wring out of the *Modest Proposal*). If they'd have us, some of us, even one of us, of could just sign up and get things going – it would only take one to kick it off. **No one has to decide what the framework will be and, as it grows, a network of relationships is built up**.

Inauguration rather than innovation for once. We have, of course, online resources already such as ADB but this could be a big, juicy one for us all to help build as well as use - one that would also include data on corporations and could include individuals beneath ADB's notice. It should be of interest (if it isn't already) at least to those of us with a taste for that sort of thing. Please explore.

2023, August 1:

Another resource that some archivists may not be aware of is VIAF (Virtual International Authority File). Well-known to librarians et al.

2018, December 23: For the want of a comma

One for the silly season – something to take our minds off Brexit, ISIS, and falling stock prices. When I worked (briefly) in the DSS Legal Unit, my boss and I were responsible inter alia for proof-reading all legislation, subordinate legislation, and instruments (there was a lot of it!) Those were pre-digital days when documents were typed and proof-set. The documents had to be read for sense, grammar, and spelling but also for punctuation. On one memorable occasion we missed a comma and this became known thereafter as the Back/Hurley comma. Now, an article on the BBC News site reminds me how important punctuation is in formal documents.

> ... A dairy company in the US city of Portland, Maine settled a court case for $5m earlier this year because of a missing comma. Three lorry drivers for Oakhurst Dairy claimed that they were owed years of unpaid overtime wages, all because of the way commas were used in legislation governing overtime payments. The state's laws declared that overtime wasn't due for workers involved in "*the canning, processing, preserving, freezing, drying, marketing, storing, packing for shipment or distribution of: 1) agricultural when produce; 2) meat and fish products; and 3) perishable foods*". The drivers managed to successfully argue that because there was no comma after "shipment" and before "or distribution", they were owed overtime pay. If a comma had been there, the law would have explicitly ruled out those who distribute perishable foods ...

It is common practice to omit commas before "and" and "or" but I have always tried to apply a rule that one should be included in a list of three or more. I believe I got this out of the great Strunk & White style guide when I was still at school. I was pleased, therefore, that the article continued:

> ... Arguments have been fought over the value of so-called **Oxford commas** (an optional comma before the word "and" or "or" at the end of a list). There might be good arguments on either side of the debate, but this doesn't work for the law because there needs to be a definitive answer: yes or no. In high-stakes legal agreements, how commas are deployed is crucial to their meaning. And in the case of Oakhurst Dairy against its delivery drivers, the Oxford comma is judged to have favoured the latter's meaning ...

It is so pleasing to know that, in this age of poor grammar and sloppy spell-checking, stuff like this still matters – at least where the exact meaning of documentation counts. In imaginative writing, ambiguity can be a virtue, but the article makes the point that deliberate ambiguity in formal documentation can also be achieved by the creative use of punctuation:

> ... Getting different countries to sign up to the same principles can be challenging, particularly for climate change agreements. Early climate change conventions included this line: "*The Parties have a right to, and should, promote sustainable development.*" The sentence ensures those signing the agreement have the ability to promote sustainable development – and should do so. But in its original draft, the second comma was placed after "promote", not before it: "*The Parties have a right to, and should promote, sustainable development.*" Some countries weren't happy with the original wording because they didn't necessarily want to be locked into promoting sustainable development. Moving the comma kept the naysayers happy while placating those who wanted stronger action. "By being slightly creative with punctuation, countries can feel like their interests have been addressed," explains Stephen Cornelius, chief advisor on climate change with the WWF, who has represented the UK and EU at UN climate change negotiations. "You're trying to get an agreement that people can substantially agree with." ...

2019, January 22: ARANZ is looking for information on records held ...

ARANZ is conducting an historic records search in relation to a NZRC into institutional abuse (see below). It is good that they are already looking at maintaining the data base along the lines of Find and Connect instead of "retiring" it like AHRR. Projects like this give rise to a multitude of issues that I hope were at least touched on at the Canberra Summit last month before they passed the Canberra Declaration.

- Which is better: a targeted db that is theme specific (e.g. institutionalisation programmes) or a generic db that can contain such data and make it usable alongside other data (assuming a generic db could be funded and maintained)?
- What is the on-going viability of theme based projects? Presumably, the victims of institutional abuse will eventually all die and the puff behind projects such as Find and Connect, insofar as they derive from the search for reparation and justice, will expire with them. What is the continuing rationale for funding and maintaining such projects beyond that?

- How are the boundaries of theme specific data to be defined and who is to define them (do we maintain a register of functions/mandates)? What is meant by "care of children and other vulnerable people"? I have always thought that the refined tortures I endured as a child for two bleak and ghastly years at boarding school are worthy of someone's investigation but I doubt it could be part of this.
- Can standardisation be achieved in the data structure to assist cross fertilisation between projects, certainly, but more importantly to clarify meaning (e.g. "type of institution", meaning of "care" and "vulnerable")? Where does it all fit into the descriptive standards debate?
- Should the db be limited to a "tiny proportion" of the "vast quantity of documentary evidence of life … [that] is produced every day" and forms part of a "distributed national collection" (Canberra Declaration 3 & 4) or should it also document ungathered (never-has-been-and-maybe-never-will-be-collected) documentary evidence as well?
- How is the tension between a focus on collections (where stuff is to be found) and originators of the stuff to be handled (do we maintain a register of corporations and institutions as well as a register of "collections" + a register of ungathered records)? How can these perspectives best be represented in the db?
- In short, should such projects proceed to "identify a suitable schemata/framework for mapping and planning … the documentary heritage universe" or (in the delusion that it amounts to covering the same turf) merely survey "the existing state of … documentary heritage **holdings** [my emphasis] to identify strengths, overlaps, weaknesses and gaps" (Canberra Declaration 11).
- How do we describe corporations and institutions whose mandate is/was broader than institutional care of children and other vulnerable people? How are the records descriptions to be contextualised – at a granular level (in Flatland) or structurally (in a R/keeping Multiverse)? How is all this to be "represented" contextually and historically? Should the db focus on themes or functions?
- And so on, and so on, and so on ….

2019, March 2: [Rolling around in the muck](#)

Came across an enjoyable little website from the U.S. called [MuckRock](#) which proclaims itself to be "a non-profit collaborative news site that brings together journalists, researchers, activists, and regular citizens to request, analyse, and share government documents, making politics more transparent and democracies more informed". It's a big site that will repay further exploration. At first blush it looks like a somewhat racier and more pro-active version of the [National Security Archive](#) and they have a newsletter to which you can subscribe

It gives rise to some reflections about archives and archivists in the digital world. As more documentation comes online (and total digitisation is, after all, the proclaimed goal of our own and other governments) it raises the questions like:

- Are online enablers such as MuckRock, the reference archivists of tomorrow?
- Are our tools and our methods to give way to theirs?
- Are there, at any rate, lessons we can learn from them to improve our tools and methods?

🏛 MUCH ADO ABOUT DESCRIPTION

They claim to have been instrumental in forcing the CIA to establish its online archive – the Freedom of Information Act Electronic Reading Room. Lots to reflect on there. I cannot forebear, for example, from commenting on how much more it is like a Search Room than a Reading Room. It seems (quaintly) to be applying something like a 30-year rule to "other CIA release programs" and lots of stuff is redacted anyway but it would provide plenty of scope, I imagine, for the clever re-use of even redacted material that Tim Sherratt does so brilliantly. Indeed, MuckRock seems to go in for that sort of thing also.

Compare the two sites. The CIA site presents the material according to a fairly unsubtle agency agenda. MuckRock re-presents it according a research agenda. (Trying to use the most neutral descriptions I can here without impugning anyone's motives). Neither, I would say, would be models for Acton's notions of objectivity. To return to the issue posed by the museums guy during the GLAM session in Washington (20 Aug last), how can such sources be trusted in a post-truth era? Do archives (and other players in GLAM space) take their online places alongside such sites or do we try to re-establish our bona fides (our brand, if you like) as witnesses for truth?

The CIA's is a single agency site, of course, with a focus on recent events but it gives rise to the question: with resources like this what is NARA for? To make and oversee policy? To establish high-level contextuality enabling navigation across sites (rather than within them)? To keep the bastards honest somehow? Certainly not to be a gatherer and provider of access (at least not for stuff users are really interested in). Perhaps we will be custodians of last resort for the undigitised left-overs. In short, how much of traditional archival methods still need to be practised?

MUCH ADO ABOUT DESCRIPTION

Are we digitising the old stuff we hold in a desperate (and possibly doomed) attempt to stay relevant? When a real digital divide separates the digital and the digitised (on the one hand) and the undigitized, probably-never-will-be- digitised (on the other) will unregenerated archival methods be so different from online ways of finding (and different ways even of describing) records online that we will be relegated to the role of mere custodians of an undigitised detritus from earlier times? Admittedly, there'll still be a lot of it, but I still like to think that what's happening on both sides of that divide requires the mind and skills of a recordkeeper. Time will tell whether that is a forlorn hope. Sometimes I wonder if it's even an aspiration that we all share.

2019, July 10: Periods in description

From time to time I have speculated about the use of "time periods" as access points in descriptive systems. They would fit firmly into the browsing/filtering rather than the targeting approach. I know AWM uses them to categorise material by "Conflict"

Napoleonic Wars, 1803 –1815 (1)
South Africa, 1899-1902 (Boer War) (3)
First World War, 1914-1918 (125)
Second World War, 1939-1945 (261)
Vietnam, 1962–1975 (21)
Indonesian Confrontation, 1962–1966 (1)
East Timor, 1999–2013 (5)
Egypt [Sinai] (MFO), 1982-1986, 1993 – (1)
Afghanistan/Pakistan (UNMCTT), 1989-1993 (1)
Cambodia (UNAMIC), 1991-1992 (1)
United Nations Protection Force, Croatia, Bosnia–
Herzegovina and Macedonia (UNPROFOR) 1992 – 1995 (1)

And I have just come across something similar on the National Archives (UK) website

All time periods
Medieval 974-1485
Early modern 1485-1750
Empire and Industry 1750-1850
Victorians 1850-1901
Early 20th Century 1901-1918
Interwar 1918-1939
Second World War 1939-1945
Postwar 1945-present

Happily, our written records cover a much shorter span although, when I was seconded to PROUK (as it then was) in the '70s, I was told they had only one piece that pre-dated 1000 (presumably that accounts for 974). You might say that all you have to do is fit the date field in a description within a date span specified as search criteria. But it's not that simple. Periods
- Suggest patterns of search for the user;
- Assign stuff according to purposeful descriptive intent;
- Avoid mismatch because of archival confusion over dating the artefact vs dating the content (cf. PS below);

- Provide for enclosures and estrays that are not accounted for in the date(s) recorded in the description - even though they should be;
- Allow for thematic overlap (different themes in the same or over-lapping periods).

Does anyone know of other examples?

P.S. Of course, dating digital records makes distinctions between date-of-the-record, date-of-the-rendition, date-of-use, and date-of-the–content even more important (and harder to conceptualise). Come to think of it, if you add microfilming into the mix, those conceptualisations are equally valid with pre-digital records.

2020, January 9: … Designing the Archive ASA-ICA 2019 Conference

This posting was in response to one that appeared on another List (to which I do not subscribe).

> **<<Chris Hurley argues that parallel provenance is unresolved provenance, ie the provenance of a record simply hasn't been accurately identified yet.>>**

Sigh. Not quite … It would be more accurate to say that *Chris Hurley argues that parallel provenance is unresolved context.* To put it (yet again) as simply as I can, parallel provenance is a problem (and not a solution) that arises when the ambience is too narrow to encompass the provenance. It ceases to be a problem (i.e. the problem is resolved) when the ambience is broadened and reaches the boundary of the multiplicity you have identified. It's about getting the context right by adjusting the ambience to fit the multiple provenance. What has to be "accurately identified" is not just "the provenance of a record" (in whatever multiple forms it may take). That you have already done when you come to consider whether or not parallel provenance arises. When you've accurately described the provenance, you'll find that what you've got is one of two things -

- **either** a description that is completely and accurately contextualised in which the provenance is adequately comprehended by your ambience (**multiple provenance**)
- **or** a description that is partially and incompletely contextualised because your ambience can't contain the multiple provenance you've identified (**parallel provenance**).

Provenance	Time period	Ambience
Multiple	Different	Same
Simultaneous multiple	Same	Same
Parallel	Same	Different

> **<<For me, it contains the potential for enduring provenances, that co-exist in the one collection and not only don't need to be resolved, but shouldn't be, because to resolve them is to eliminate (the facilitator's word again) one of the originating systems from view.>>**

It follows that I have no substantial objection to this view, though I would quibble with "potential". Description is about depicting reality, not about imposing a view on it. There is, of course, always an element of organising perception - as with classification in the natural sciences. But ultimately our descriptions are meant to be accurate rather than artistic. If you're not already describing multiple provenance then you're doing it wrong. The only potential involved is to get it right in future.

2020, April 20: Collaboration and leadership

On the NZ list, chatter following the announced closure of the Community Archives website has prompted a statement from the NZ National Archivist which reads (in part):

> … As Chief Archivist, I am aware of the sense of a lack of leadership for community-held collections. There is also duplication of effort and offerings which indicates we could benefit from a joined-up approach. To that end, I have, along with the National Librarian Bill Macnaught, been working closely with the leaders of documentary heritage organisations which have national mandates to set up and share the mechanisms for collaboration at a national level. Those organisations include Te Papa Tongarewa, Auckland War Memorial Museum and Ngā Taonga Sound & Vision. This work is known as the National Documentary Heritage Strategy (NDHS) working group.
>
> Our initial roundtable discussion in March 2019 supported in principle the notion of collaborating on a national documentary heritage strategy, with the aim of developing a collective action plan. This work includes a community-held collections focus. A number of further meetings have been held. The group is re-forming and re-focusing as the new chief executives at Te Papa and Ngā Taonga settle into their roles. I hope to be able to share more of this work in the coming months, and that a more coherent strategy provides the basis for more effective use of our current and any new resources. I note that, while the work has not concluded, Ministers have also started considering how to strengthen the contribution that national libraries and archives to New Zealand culture through the National Archival and Library Institutions Ministerial Group …

The Community Archive began many years ago as the National Register of Archives and Manuscripts (NRAM). The Chief Archivist's statement indicates the justification for its closure beyond technological obsolescence (in short, the reasons for not maintaining and migrating it to a new platform) and these are more interesting because they go to the essence of the thing rather than the mechanics.

> After 10 years since its inception, the website was increasingly difficult to maintain by us and by contributing organisations, and usage was very low by current standards. It does not connect people directly with the archives themselves and provides functionality readily available through other options.

Not sure what the other options may be but this makes depressing reading for a proponent of the *Modest Proposal* (MP). Does it demonstrate that if we built it they would not come?

- The *Modest Proposal*, first and foremost, would provide a **structure** (multi-layered and multi-faceted) into which contributions would be contextualised rather than an assemblage of offerings from participating contributors. The work to develop this contextual framework (as with any finding aid) would be far more challenging than the mere task of adumbration and compilation. Because, as with any finding aid, the value-add would lie not just in hosting content but in conferring meaning it would not be uncontested. Hence the need to build it on the principle of parallel provenance.

- The *Modest Proposal* would provide **direct access** to the source descriptions so far as the capability of the native programmes permitted. Indeed, the core proposition is that the MP gateway should displace the front-end for the majors so that entry into their own descriptive efforts would be via MP rather than any home-grown (necessarily partial) view of what they are describing. MP would give their descriptions the same larger contextual framework within which to work as the one provided for community archives. At the same time, it would offer smaller players without the capacity to join up online the opportunity to contribute into an NRAM-like framework. Usage would not be very low because there would be nothing else to use.

- The *Modest Proposal*, therefore, **requires much more** from the national institutions than mere support, co-ordination, and encouragement for the efforts of others. It requires them to submit themselves along with everyone else to the same architectural framework and to thus truly integrate the national descriptive effort into a seamless search and discovery mechanism. This could incorporate built in labour-saving efforts (viz. SNAC) for sharing contextual data rather than developing ontological schema for distributed use. Thus eliminating duplication of effort.

But a "national documentary heritage strategy" is better than nothing, I suppose.

2020, September 15: Resource request … 7th Biennial Conference …

<<Deborah Lee-Talbot: I'm a PhD candidate at Deakin University. Part of my thesis is concerned with the Australian Joint Copying Project and the role of Phyllis Mander-Jones in the creation of this unique archive. Does anyone have access to a copy of the following paper that they can share, please? Terry Eastwood, 'Reflections on the Development of Archives in Canada and Australia', in Papers and proceedings of the 7th Biennial Conference of the Australian Society of Archivists, Inc., Hobart 2-6 June, 1989', pp. 75-81 …>>

<<Joanna Sassoon: A couple of fairly fundamental terminology questions: Is the AJCP unique, and is the AJCP an archive?>>

<<Andrew Waugh: … Deborah doesn't say that the AJCP is unique, but that the resulting collection of information is unique. Which, of course, is self evidently true.
      As to whether it's an archive, I would say it's not. It's a carefully curated, selected, collection of access copies of documents held in other repositories (some of which are archives). It meets none of the criteria for being an archive. At its most fundamental, that is why the AJCP was thought of by the NLA, not the NAA (it's a library mindset), run out of the NLA, and the result is held as part of the NLA collection.
      It's also a great example of the technological contingent nature of providing access. If the AJCP was being run today, Australia would be paying the source repositories to scan and make available electronically the original documents. We wouldn't need to hold access copies in Australia.>>

<<Deborah Lee-Talbot: Thank you for this post, Joanna. You've cut right to issues that I examine in two of my chapters. From a discussion about the creation of the AJCP, issues of materiality, duplication and the positioning of the AJCP within a Pacific context I do believe it is unique. Regarding the second question, is it an archive, that is a work in progress. Perhaps after I receive the latest round of books from Deakin Library, I will have a better term to apply, or I will have more evidence to add certainty to this assertion.>>

<<Deborah Lee-Talbot: Thank you for your perspective, Andrew. Can you please provide information at the criteria you are using to define an archive? Focusing on the Australian context, I am currently using the simple definition of 'non-current records deposited or selected for deposit in an archival institution'.>>

<<Andrew Waugh: I'm paraphrasing, of course. For a recent discussion of the state of archival thought, I'd suggest 'Encyclopedia of Archival Science', Duranti & Franks, Rowman & Littlefield, 2015, ISBN 978 0 8108 8810 4. It doesn't have a definition of 'archive'. Look at the entries for 'Archival Bond', 'Appraisal', and 'Archival Collection'. In particular, in the later, "Note that an archival collection is not to be confused with an artificial collection, which is a set of individual items with separate provenance brought together by a collector around a theme." The key, for me, is the conscious *selection* of the AJCP from the original fonds …>>

<<Catherine Robinson: From 'Keeping Archives', 2nd edition ISBN 1 875589 15 5 .
      Archives – "Those records that are appraised as having continuing value. Traditionally the term has been used to describe records no longer required

for current use which have been selected for permanent preservation. Also referred to as permanent records."

The AJCP is a selection of copies of records and managed as a collection. The archives are held in the repositories in the UK.>>

<<Joanna Sassoon: ... And those pillars of archival theory provenance and original order .... And of course what is filmed from within a file is also a selection ... and you never actually know what has not been filmed unless you return to the original file. And since many of the archives (in the M series at least) are in private hands what was not selected becomes less easily known as the decades pass.

Whatever practices went with filming which are at times open to question, Mander Jones' foundation bibliographic work is astounding and I treasure my battered secondhand copy I was lucky to purchase from a dealer who had no idea of its value ... The other question, relating to the private collections filmed is, now that a selection has been digitised, will anyone want to collect the originals?>>

<<Deborah Lee-Talbot: Thank you, Andrew ... I appreciate the generosity of you, Catherine and Joanna in the provision of materials and time towards this discussion. To be able to discuss ideas and research, especially at the moment, is welcomed.

Joanna, I recognise the value of what you are saying regarding the selection process. The selection of the AJCP M-Series materials was the consequence of a strict criteria as set by Mander-Jones: to be copied the record had to be produced as a consequence of direct experiences in the Pacific, New Zealand or Australia. A record would be excluded if the author produced a field journal in, say, Kew about a botanical species from Australia. While this certainly meant information was excluded, it raises an opportunity to consider how a collection was created specifically for this region, from materials produced within this region, yet from archives held in London. I too deeply value my copy of Mander-Jones book, I was quite delighted when I acquired my copy last year from the Royal Historical Society of Victoria book sale.

As to the digitisation of the collection, I think of this a lot while writing my chapters. For some historians, no matter how clear and concise an image is, the digital is never good enough. They will certainly covet the originals. For others, myself included, the ability to access digital collections to perform historical social and cultural analysis has been greatly welcomed.>>

<< Is the AJCP unique, and is the AJCP an archive?>>
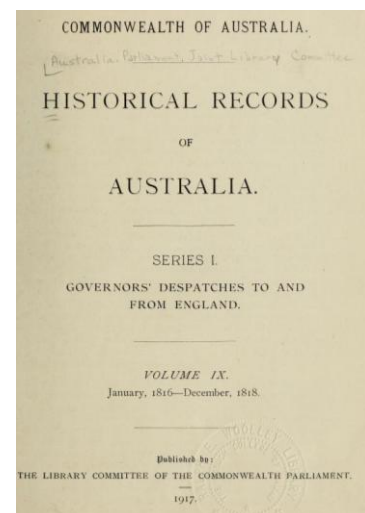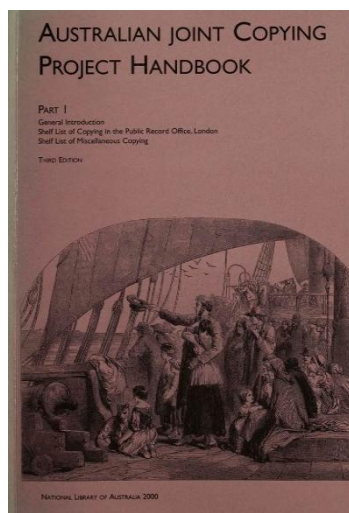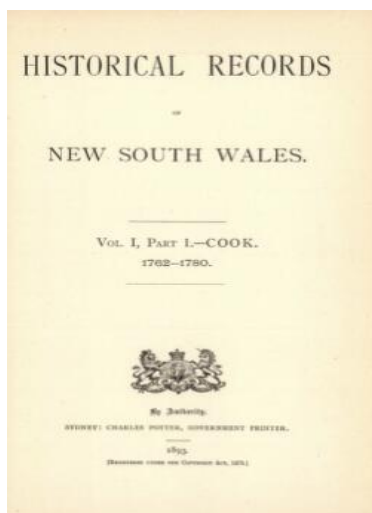
I suppose you could say that the AJCP materials are an archive of the copying project itself – of a piece with *Historical Records of NSW* and *Historical Records of Australia* – an even earlier technology and possibly not of the same quality. The whole question of records publication (from facsimile to calendars is a deep one). In the 19th and early 20th centuries, governments would publish blue books comprising copies of huge quantities of official

papers on topical or politically "hot" issues. This was a propaganda effort by the originators of the documents. There was also a vogue for running off facsimiles of lurid documents and selling them as mementos. When Keeper, I went on a wild goose chase to look at a possible Kelly estray in Dromana. It turned out to be a 19th century facsimile offprint of an official document. I had to break it to the good souls that they weren't sitting on a fortune.

It was my understanding that Pacific Manuscripts Bureau did this kind of work,

> The Pacific Manuscripts Bureau copies archives, manuscripts and rare printed material relating to the Pacific Islands. The aim of the Bureau is to help with long-term preservation of the documentary heritage of the Pacific Islands and to make it accessible.

I also have a vague recollection (very vague) of sitting down to dinner at Glenda Acland's house in Canberra sometime in the 1970s with a visiting archivist from the Bundesarchiv in Koblenz. His project involved locating the surviving archives of German Colonies dating from the period of the Second Reich (1870-1918). This was during the Cold War and many of the archives of the German Foreign Ministry and Ministry for Colonial Affairs were in East Germany at Potsdam and virtually inaccessible to West German archivists and scholars.

The Project involved copying surviving colonial records to reconstitute a kind of mirror-image of correspondence series between the Ministries and the colonies. A kind of Faux Fonds. Instead of the Home Government records (Ministry out : Colonies in) the microfilms would comprise the reverse (Ministry in : Colonies out). Don't know if anything ever came of it; it may have simply been a terrific junket to assess feasibility.

He would have come to Canberra where the German New Guinea records were being microfilmed prior to repatriation to the newly independent PNG. At that time Hilary Rowell would have been involved and she might be able to help. If you find this interesting, I'm sure the visit would have been documented in some file (or files) at NAA which would now be in the open period (if it/they survive).

2020, September 16:

<<Deborah Lee-Talbot: That's one way of looking at it. I'm starting to think approaching the AJCP as a historical artefact is a way to go in this project. In regards to PMB; the AJCP started acquiring Pacific materials and then Maude started the PMB to ensure a specific collection was created. There's a little crossover between the two. Thank you for sharing your recollection. I've started looking into whether anything did come from this visit … With your mention of 'Faux Fonds' I was reminded of the UC San Diego PNG Patrol Reports which were digitised. But there's no reflection, just a straight copy …>>

2023, January 31: Indexing

Went into Sydney yesterday and disgraced myself once more at the excellent Abbeys Book Shop. Picked up a title I wasn't looking for that others may not know about either:-

Dennis Duncan *Index, A History of the* pb Penguin 2022

I didn't find the contents lived up to the title but it provided an enjoyable 40 minutes on the train back to Gosford before turning to some of my other purchases. I spent many hours in 1969 swotting the *ALA Filing Rules*. Duncan did succeed in bringing back some of the memories, e.g.

> Letter-by-letter: Newman, Paul / newspapers and news-sheets / *New Tenures*
> Vs
> Word-by-word*: New Tenures* / Newman, Paul / newspapers and news-sheet

I once had a twenty-minute lecture from Peter Scott about what he called lexicographical order (by which I think he meant word-by-word but I stopped listening after the first five

minutes). I've always had an interest though (<u>Teacher to Young Self</u> in the middle of a two-hour tutorial on a hot afternoon in 1969 at the Kensington campus: *Mr Hurley, if you ask one more question, I'll scream!*). I used scan indexes in reference works for fun (I have strange ideas of fun) – the *White Pages* Telephone Directory, for example (do they still have those?) to see what principle governed the arrangement of entries.

The next book I opened was Don Hollway *At the Gates of Rome; the Fall of the Eternal City, AD 410*. Going first to the index, I found little nourish my interest (books have short indexes) but these entries provided some grist:

> Julian, Emperor 68
> Julian Alps 100,198
> Julii 65
> Julius 59-60, 61
> Julius II, Pope 247

From the mists of memory, I'd say the second (Julius/Julius II) is an example of the nothing-before-something rule. The first (Julian, …/Julian Alps) probably has something to do with the comma giving priority despite the fact that it appears to violate nothing-before-something (but this is all very hazy for me now). There's lots of rules about punctuation to say nothing of numerals and those danged foreigners – "de" "D'" "von" etc. And then the question in history books of alphabetising the sub-entries or arranging them chronologically (or else in the order in which they appear in the text).

I thought about becoming an indexer in retirement. I was enchanted many years ago by this from Michael Roberts *The Early Vasas, A History of Sweden 1523-1611* (1968)

> Vasa, Cecilia: loses hair, 104; and reputation, 207; involved in plots, 248; her dowry unpaid, 309

It's the "and" that is genius. Authors and editors usually leave the index to the last minute and often don't check. Being an indexer might not have worked for me – too much opportunity for mischief.
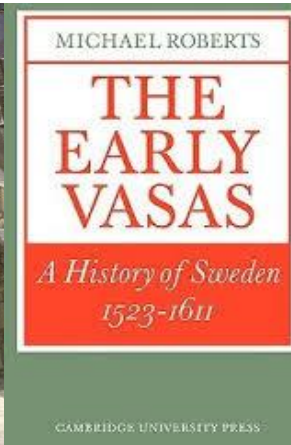
Mora, 1520   **Cecilia Vasa**

My PC returns lots of results in alphabetical order and I've sometimes wondered (not enough to find out) what rules apply. There's probably some standards (*the great thing about standards is that there's so many to choose from*). In books, most people don't seem to care whether there's a Name Index separated from a General Index (or even Place Index) or all in one.

<<<u>Kim Eberhard</u>: Thank you Chris! I chuckled out loud at the Vasa, Cecelia entry! I encourage you to take up indexing in retirement; your 'mischief' would be very welcome. I too enjoy perusing indexes (but not for the same reasons); I wonder how many of us will admit the same?>>

🏠 MUCH ADO ABOUT DESCRIPTION

2023 February 1:

<<My PC returns lots of results in alphabetical order and
I've sometimes wondered (not enough to find out) what rules apply. >>

In case anyone else is interested and knows as little as I do about it, it turns out that IT folk have their own perspective on indexing which is not just about analysing content (abstracting) and then organising data of uniform value (filing). Most of the articles I found seem to be about how search engines decide **what to display** (essentially a back-office process) rather than **how to organise** the displayed data for the user to scrutinise. The underlying purpose is to rank hits in the order **deemed most relevant** to the user's query. I suppose that's why I get a zillion results and only ever scan the top two or three pages.

Techopedia Starts off with a traditional definition but quickly moves on to "similar uses" in IT
In general, indexing refers to the organization of data according to a specific schema or plan. In IT, the term has various similar uses including, among other things, making information more presentable and accessible. One example of indexing is the legacy Microsoft Indexing Service, which maintained an index of files on a computer or in an operating system environment. Another example is database indexing, which involves creating an index for a database structure to help expedite retrieval of data. One common type of indexing in IT is called "search engine indexing." Here, IT tools aggregate and interpret search engine data, again, to streamline data retrieval. This type of indexing is also sometimes called Web indexing. IT experts explain that indexing helps to make searches less labor intensive — without an index, the search engine would have to search every document at its disposal equally, whereas with an index, much of this work is eliminated.

Search engine indexing
Search engine indexing is the collecting, parsing, and storing of data to facilitate fast and accurate information retrieval. Index design incorporates interdisciplinary concepts from linguistics, cognitive psychology, mathematics, informatics, and computer science. An alternate name for the process, in the context of search engines designed to find web pages on the Internet, is *web indexing*. Popular engines focus on the full-text indexing of online, natural language documents.[1] Media types such as pictures, video,[2] audio,[3] and graphics[4] are also searchable. Meta search engines reuse the indices of other services and do not store a local index whereas cache-based search engines permanently store the index along with the corpus. Unlike full-text indices, partial-text services restrict the depth indexed to reduce index size. Larger services typically perform indexing at a predetermined time interval due to the required time and processing costs, while agent-based search engines index in real time.

How computers link search queries to data content
Indexes are a powerful tool used in the background of a database to speed up querying. Indexes power queries by providing a method to quickly lookup the requested data. Simply put, an index is a pointer to data in a table. An index in a database is very similar to an index in the back of a book.

How search engines display websites after they have been "crawled" [basically, reviewing content and using it to rank the site when a search is made]
Indexing is where the ranking process begins after a website has been crawled … essentially … adding a webpage's content to Google to be considered for rankings. When you create a new page on your site, there are several ways it can be indexed. The simplest method of getting a page indexed is to do absolutely nothing.

And then there's automatic indexing
Automatic indexing is the computerized process of scanning large volumes of documents against a controlled vocabulary, taxonomy, thesaurus or ontology and using those controlled terms to quickly and effectively index large electronic document depositories. These keywords or language are applied by training a system on the rules that determine what words to match. There are additional parts to this such as syntax,

usage, proximity, and other algorithms based on the system and what is required for indexing.

Assigning terms (metadata) to lots of "documents"

Automatic Indexing is the process of assigning documents with search terms for search and retrieval purposes. This process in searches is widely used today to lessen the time of the search. It uses a computer to scan a large volume of documents against a dictionary, rather than manual indexing which makes use of manpower due to manual typing.

Or, indexing just one document

An index can usually be found at the end of a document, listing the key words and phrases in a document, along with the page numbers they appear on. There are two steps involved in creating an index: defining which words you want to appear in the index and then inserting the index. See also Using Content Analysis

And no doubt much else besides. Things were simpler back in 1969 but maybe not all that different. I can remember, when I worked at NLA, a large area called the National Union Catalogue of Monographs (NUCOM) where they were inter-sorting copies of 5x3 catalogue cards sent in from libraries all around the country and "normalising" them by deciding which entry to file them by because different libraries catalogued the same title under different main entries. So far as I can tell, NUCOM doesn't exist anymore but its ghost probably resides in various online initiatives.

Dennis Duncan has this to say:

> … The subject index has dominated all but the earliest chapters of this history; by contrast our twenty-first-century Age of Search is, in effect, an age of automated concordance (P.233) … Once our indexing information – heads, locators – can be 'read' by machines, then it doesn't much matter whether it is stored on punch cards, magnetic tape or integrated circuits. The indexer's job has been distilled to its analytical essence; the drudgery – the shuffling and copying – has been delegated to the machine (pp.244-245) … Ultimately … what both inclusion and exclusion methods produce is a scaled-back concordance. The terms of the index are taken directly from the text; nothing appears in the former except in the precise form in which it appears in the latter (p.247) … [But can Informatics, e.g. content analysis applied to search, overcome these limitations?] A good subject index can only be the product of a good indexer, an expert reader who knows something about the subject in hand … A specialist indexer knows that it can be helpful to tag a concept even if it is not explicitly named … they know that, thanks to metonymy, sometimes a reference to 'Number Ten' or to 'Downing Street' belongs under *Johnson, Boris,* and sometimes it doesn't … The limitations of *unimaginative indexing* … become starkly apparent if one tries to locate the parable of the prodigal son … using a Bible concordance, The parable does not contain the words *forgiveness* or *mercy*, or for that matter *prodigal* (pp.259-260) ….

<<Andrew Waugh I'm sure you … will find this interesting … It's an explanation of the recent advances in AI that led to ChatGPT and Dall-E. Not at all technical. It will amuse you, I'm sure, to learn that the first key insight was in indexing (words in text to facilitate automatic translation, then elements of pictures). The concept of auto-classification has just moved a large step nearer with this technology, though it will take a far cleverer person than me to work out how.>>

🏠 <u>MUCH ADO ABOUT DESCRIPTION</u>

2023, February 2:

Up to Newcastle this morning. This from the *Newcastle Herald* (p.12).

Brace yourself for the AI revolution

Then an article about pros and cons and at the end: EDITOR'S NOTE: This was not written by ChatGPT.

<u>Some of the pros</u>: highly cohesive, human like responses; can write an essay or a novel based on a short prompt; formulate responses for service centres; create marketing material or press releases; write editorials or news items.

<u>Some of the cons</u>: the knowledge is static and doesn't access new information (ChatGPT is stuck in November 2022); sometimes "makes up facts"; right now there's still a place for quality control, credibility, and fact-checking.

Wonder if they need to update the laws of libel and copyright?

Please yer honour. I didn't libel the gentl'man. It was me computer!

<<Richard Lehane ... and ChatGPT is already making records (Microsoft just announced a new premium version of MS Teams that uses Chat GPT to sit in the background of virtual meetings to take minutes ("intelligent recap"), to do lists, personalized highlights and various other things.)>>

2023, February 10:

Alan Kohler has this to say (inter alia)

Everybody is having a lot of fun at the moment playing with ChatGPT – asking it to explain itself and say whether it's going to take all of our jobs, as well as asking it to write columns and essays ... There's a new version of ChatGPT coming soon, expected to be vastly superior to this one ... And then on Tuesday Google launched its own conversational AGI service to **compete with ChatGPT, called Bard** ... Australia has joined something called the Global Partnership of Artificial Intelligence (GPAI), which came out of the G7 and has 15 members ... ... I don't think that artificial intelligence will be controlled or regulated and we need to prepare for the consequences ... A pioneer of the generative AI behind ChatGPT, Queensland-born Stanford University professor Christopher Manning, told the *Financial Review* this week that people will need to adapt to a world in which misinformation and false images are rife ... it should be regulated, but it won't be. It's already too late.

All this is a long way from indexing.

2023, February 11:

The Guardian is not impressed:

... in the profit-driven competition to insert artificial intelligence into our daily lives, humans are dumbing themselves down by becoming overly reliant on "intelligent" machines – and eroding the practices on which their comprehension depends. The human brain is evolving. Three thousand years ago, our ancestors had brains that were larger than our own. At least one explanation is that intelligence became increasingly collective 100 generations ago – and humans breached a population threshold that saw individuals sharing information ...

This socialisation of synaptic thought is now being tested by a different kind of information exchange: the ability of AI to answer any prompt with human-sounding language – suggesting some sort of intent, even sentience. But this is a mirage. Computers have become more accomplished but they lack genuine comprehension, nurtured in humans by evolving as autonomous individuals embedded in a web of social practices ...

... Chatbots sound more authoritative, but they are not more truthful. Prof Marcus points out their errors, or hallucinations, are in their "silicon blood", a byproduct of the way they compress their inputs ... Journalists, politicians and poets might be very concerned about the "semantic" aspects of communication, but not so much AI engineers. They look at

the information in a message as a measure of the system's disorder. That's why AI risks creating a new class of weapons in a war on truth … The danger is not machines being treated like humans, but humans being treated like machines.

The Internet could be regarded, I suppose, as the ultimate example of collectively sharing information. What the author seems to object to is the absence of cognitive intent in the process of formulation.

A disconnect of AI from intent might, however, be a good thing. We might be safer from misinformation if "untruthfulness" and "hallucinations" are an unintended consequence – a "byproduct" of systems that mimic human behaviour without cognitive ability. What I'm still unclear about when it is claimed that AI could be an assault on truth is whether there is potential for wilfully designing or using Chatbots to deceive and manipulate – a new kind of deliberate fake news. There's a whiff of that in some of the articles but I can't see anyone coming out and saying it or explaining how it would be done (although the articles that link it all to national security come closest). In terms of online opinion and incitement that we're already dealing with, it's a problem we already have – how to authenticate what you are given.

2023, February 12:

Meanwhile …

> … This week, Australia's government-owned postal service sounded an alarm for letter writing after reporting a $190m loss in its letter business over a six-month period. Every year it is costing Australia Post more to deliver fewer letters as a growing population demands more delivery points … The postal service expects the "unstoppable decline" will gather pace, making letters a peripheral form of communication by 2030 … Rewind three decades, and the letter delivery business was booming. In the 1990s, letter volumes grew in tandem with Australia's economic progress, increasing by 5% a year, according to an analysis of Australia Post financial reports.
>
> Mail volumes hit a high point of well over 5bn in 2007-08 when the basic postal rate was 50c. But the global financial crisis and the surge in popularity of text messaging and public webmail services like Hotmail prompted an irreversible change in behaviour. Australians switched to convenient and cheaper communications. Letter volumes at Australia Post have fallen ever since, diving to just 1.6bn in 2021-22 …
>
> Postal services around the world are grappling with the same problems; some have been privatised amid heated political debate, while others are reducing the number of days they deliver letters … domestic letters now contribute less than 20% of revenue to the postal service, making it more fitting to describe Australia Post as a parcel and services company that also delivers letters … Australia Post describes the current state as "unsustainable" …
>
> Peter Slattery, a research fellow at Monash University, says he sees a future role for physical letters even if more generic correspondence goes digital. "Both in the business world and the personal world, letters will be associated with high-value, selected communication and more of the mass communication will switch to digital," says Slattery, who writes on behavioural science … Letter volumes also get a boost during elections and national events like the census …
>
> Some fear the transient nature of digital communication means future generations will miss out on having documented insight into the minds of notable figures, such as the thoughts contained in archived love letters from Johnny Cash to June Carter, Napoleon to Josephine and Elizabeth Taylor to Richard Burton.

If the letters are written by AI anyway, what will it matter?

2023, February 22:

An index analyses content and assists readers by anticipating what they will look for and how they will formulate their queries. So far as I can make out (and I know next to nothing about this and would appreciate enlightening criticism from those who know more) search

engines "push" results based, inter alia, on algorithms that match your request to a "user-profile" built up over time. Conspiracy theorists posit that this is (or can be) done with evil intent from commercial, ideological, or goodness-knows-what motivation on the part of the tech giants. In the US section 230 of the weirdly named *Communications Decency Act* shields the tech giants from liability for content they host and/or recommend. Now, the US Supreme Court is deciding whether this shield should be breached on the argument that the algorithms, make the platform owners, as well as the authors, responsible for guiding users towards harmful content (however defined) and therefore liable for the consequences (whatever they may be). Do guns kill people or do people kill people?

... YouTube's parent company Google is being sued by the family of Nohemi Gonzalez, a 23-year-old US citizen who was studying in Paris in 2015 when she was killed in the coordinated attacks by the Islamic State in and around the French capital. The family seeks to appeal a ruling that maintained that section 230 protects YouTube from being held liable for recommending content that incites or calls for acts of violence. In this case, the content in question was IS recruitment videos ... In the case of Twitter v Taameneh, family members of the victim of a 2017 terrorist attack allegedly carried out by IS charged that social media firms are to blame for the rise of extremism. The case targets Google as well as Twitter and Facebook.

... The supreme court is being asked in this case to determine whether the immunity granted by section 230 also extends to platforms when they are not just hosting content but also making "targeted recommendations of information". The results of the case will be watched closely, said Paul Barrett, deputy director of the NYU Stern Center for Business and Human Rights ... "This case could help determine whether the major social media platforms continue to provide venues for free expression of all kinds, ranging from political debates to people posting their art and human rights activists telling the world about what's going wrong in their countries." ...

Holding tech companies accountable for their recommendation system has become a rallying cry for both Republican and Democratic lawmakers. Republicans claim that platforms have suppressed conservative viewpoints while Democrats say the platforms' algorithms are amplifying hate speech and other harmful content ... Evan Greer, a free speech and digital rights activist, says that holding companies accountable for their recommendation systems could "lead to widespread suppression of legitimate political, religious and other speech" ... "The truth is that Section 230 is a foundational law for human rights and free expression globally, and more or less the only reason that you can still find crucial information online about controversial topics like abortion, sexual health, military actions, police killings, public figures accused of sexual misconduct, and more."

In other words, as I read it, when the platforms ***analyse content and assist readers by anticipating what they will look for and how they will formulate their queries***, they become co-conspirators. I suppose the only real difference between an index and a search engine is the "profiling". But, while an indexer knows nothing about the reader at an individual level, culturally he makes all sorts of assumptions and shapes his work accordingly.

2023, February 24:

Everything you need to know about section 230
and
More than you probably ever wanted to know about section 230

Since we don't have a right to free speech in this country (but lots of law abridging it) it remains to be seen whether the US Supreme Court decision when it is delivered will have implications for us. The *Wikipedia* article refers to cognate Australian law:

> In *Dow Jones & Company Inc v Gutnick*, the High Court of Australia treated defamatory material on a server outside Australia as having been published in Australia when it is downloaded or read by someone in Australia.
>
> *Gorton v Australian Broadcasting Commission & Anor* (1973) 1 ACTR 6 Under the *Defamation Act 2005* (NSW),[s 32, a defence to defamation is that the defendant neither knew, nor ought reasonably to have known of the defamation, and the lack of knowledge was not due to the defendant's negligence.

Australian judges mostly like to push (or is it pour) new problems into old bottles, so goodness knows what they would make of an issue like this. This is all about liability for someone else's content on your server. Not very relevant to us? What happens if we point users of our website (deliberately or algorithmically) to relevant content on another website deemed to be harmful. Vicarious liability anyone?

PS There's a lovely journalism story from the 1960s (I think) about a fresh young reporter excited to be given the job by his editor of writing a feature article on free speech. *Oh, yeah?* says the crusty old hand in the corner. *Well, if you find any, let me know.*

<<Andrew Waugh I thought that point had been well settled in Australia. As part of the Dylan Voller defamation case, the High Court "found the media outlets could be held liable for comments by third parties on their Facebook pages" ... This horrified the media companies, and is the reason why comments were suddenly turned off in many media websites. The previous Australian government rushed to prepare the "Social Media (Anti Trolling) Bill 2022". This would have made the (Australian) person maintaining or administering a social media page *not* the publisher of other people's comments on the page. The publisher (i.e. the entity on the hook for defamation) would be the company that provided the social media page (e.g. Facebook). However, the company would always have defence - they would not be guilty of defamation if they promptly gave the defamed person the identity of the person who authored the content. This bill never became law and lapsed when Parliament was dissolved in April 2022. The current government hasn't moved in this space.>>

2023, February 25:

> **<< I thought that point had been well settled in Australia. As part of the Dylan Voller defamation case, the High Court "found the media outlets could be held liable for comments by third parties >>**

That seems to be about liability arising from merely distributing ("publishing") or contributing to the distribution of the harmful content - against which 230 protects them in the US. The attack before the US Supreme Court looks like a craftier approach to get around 230 - viz. a different kind of liability based not on "publication" but on "pushing". You may be protected there from liability arising from your passive role in providing access to forbidden fruit but when you actively intervene to (helpfully?) guide the user to it they seem to be arguing that you incur a greater liability from which 230 does not protect you. At least that is what the cases seem to me to be about.

The fact that Australian courts don't even offer protection for "publication" still leaves open the question how they would deal with additional liability arising from a successful action here based on this new argument before the American court (it seems to me).

In a funny way, it's kind of like the difference between a book with an index and a book without one. Libraries, I believe, used to put obstacles in the way of accessing their smut

collections - sometimes by cataloguing them in Latin (or is that an old wives' tale?). And the provision of Internet access through public libraries raises even thornier issues for them.

> Even though most school children are regularly blocked from browsing X-rated Web sites in class, they can access pornography on the Internet in the most public of places: the local library ... given a choice between censorship and unfettered access, most libraries have chosen not to use special software that would block unwanted material ...
>
> The American Library Association policy on Internet access, which was endorsed by many librarians in the Bay Area, leaves it up to parents to decide what materials their children can see. "In a library, our role is to make information available. But you have to choose what you want to access ..." The library association, online services like America Online and a host of civil liberties groups are trying to invalidate Internet anti- porn rules in the federal Communications Decency Act drawn up by Congress and signed by President Clinton last year ...
>
> "The San Francisco Public Library does not monitor and has no control over information accessed through the Internet. . . . As with other library materials, restriction of a child's access to the Internet is the responsibility of the parent or legal guardian." ... Pornography is available online at the San Jose Public Library ... but the staff will ask you to clear the screen of any graphic material that might interfere with "maintaining a reasonable and comfortable environment for the public," a spokesperson said ...

2023, March 4:

Perhaps I'm naïve, but I find the doom-saying about AI a bit over-the-top. What they're saying is that sources may not be what they seem and that AI makes the detection of fakery harder than before. But this is what we've always known: don't take content on face value, rely on provenance and context.

> It has taken a very short time for artificial intelligence application ChatGPT to have a disruptive effect on journalism ... What these systems are incredibly good at is emulating human prose, and predicting the "correct" words to string together ... For the purposes of journalism, they can create vast amounts of material – words, pictures, sounds and videos – very quickly. The problem is, they have absolutely no commitment to the truth. Just think how rapidly a ChatGPT user could flood the internet with fake news stories that appear to have been written by humans.

Well, humans are pretty good at doing that too.

> ... In terms of journalism, many newsrooms have been using AI for some time ... Felix Simon, a communications scholar at the Oxford Internet Institute, has interviewed more than 150 journalists and news publishers for a forthcoming study of AI in newsrooms. He says there is potential in making it much easier for journalists to transcribe interviews or quickly read datasets, but first-order problems such as accuracy, overcoming bias and the provenance of data are still overwhelmingly dependent on human judgment. "About 90% of the uses of AI [in journalism] are for comparatively tedious tasks, like personalisation or creating intelligent paywalls," says Charlie Beckett, who directs a journalism and AI programme at the LSE. Bloomberg News has been automating large parts of its financial results coverage for years, he says. However, the idea of using programs such as ChatGPT to create content is extremely worrying. "For newsrooms that consider it unethical to publish lies, it's hard to implement the use of a ChatGPT without lots of accompanying human editing and factchecking," says Beckett ...
>
> ... Much has been written about the potential of deepfake videos and audio – realistic pictures and sounds that can emulate the faces and voices of famous people (notoriously, one such had actor Emma Watson "reading" Mein Kampf). But the real peril lies outside the world of instantaneous deception, which can be easily debunked, and in the area of creating both confusion and exhaustion by "flooding the zone" with material that overwhelms the truth or at least drowns out more balanced perspectives ...

As well as worrying about the dangers and how to control the phenomenon, perhaps trying to assist ourselves to develop better "nerd immunity" would help.

PS. On the other hand, perhaps naivete lies in trusting that people want to be un-deceived about what they <u>believe</u> (or, more accurately, what they want to believe) – not to be confused with <u>self-deception</u> which is entirely different or with <u>hoaxes</u> (e.g. drop bears).

The <u>Dominion Case</u> is a current example that suggests the thirst for Truth is thwarted by confusion of mind and wishful thinking.

> Did Fox News actively promote the <u>conspiracy theory</u> that implicated Dominion Voting Systems in a "massive fraud" that supposedly denied Donald Trump a second term? Or did Fox merely report what the president and his representatives were saying? Those questions are at the heart of the <u>defamation lawsuit</u> that Dominion filed against Fox in March 2021 …
>
> … the fact that Fox News reporters were <u>appropriately skeptical</u> of those claims does not absolve Fox of liability for the <u>credulous reception</u> that Giuliani and Powell received on shows such as *Lou Dobbs Tonight* … Dominion argues that Murdoch, who was privately <u>calling</u> their story "really crazy stuff," nevertheless decided not to intervene because he was worried about alienating Trump supporters … producers and executives … knew or should have known those claims were false and had the power to stop hosts like Lou Dobbs from continuing to promote them. They chose not to do so, Dominion argues, because they were afraid of losing viewers to right-wing competitors …
>
> … Dominion tells a plausible story, backed by internal communications, that Fox continued to host "crazy" conspiracy theorists because it had a financial interest in doing so. Viewers were angry after Fox News called Arizona for Trump, and executives were alarmed by their disenchantment and their flight to Newscom and One America News Network. In short, Dominion says, Fox favored profits over truth …

Some people want to be deceived and are neither hood-winked nor led into it by others ("I am their leader, <u>I must follow them</u>"). Even with good-will, is it ever possible to find a way to <u>sort out fact from fiction</u> within the fog of perception to one's own satisfaction, let alone the satisfaction of others?

2023, April 19:

For those not on the Canadian List, Mark Hopkins has posted a link to an <u>article</u> about historical research incorporating AI.

> Historians have started using machine learning—deep neural networks in particular—to examine historical documents … [they] say the application of modern computer science to the distant past helps draw connections across a broader swath of the historical record than would otherwise be possible, correcting distortions that come from analyzing history one document at a time. But it introduces distortions of its own, including the risk that machine learning will slip bias or outright falsifications into the historical record. All this adds up to a question for historians and others who, it's often argued, understand the present by examining history: With machines set to play a greater role in the future, how much should we cede to them of the past? …

2023, March 4:

> **<< The <u>Dominion Case</u> is a current example that suggests the thirst for Truth is thwarted by confusion of mind and wishful thinking.>>**

At least we can now put <u>a price on un-truth</u> - $1.17 billion it seems.

2023, May 21:

> **<<That seems to be about liability arising from merely distributing ("publishing") or contributing to the distribution of the harmful content - against which 230 protects them in the US. The attack before the US Supreme Court looks like a craftier approach to get around 230 - viz. a different kind of liability based not on "publication" but on "pushing". You may be protected there from liability arising from your passive role in providing access to forbidden fruit but when you actively intervene to (helpfully?) guide the user to it they seem**

to be arguing that you incur a greater liability from which 230 does not protect you. At least that is what the cases seem to me to be about.>>

The Supreme Court unanimously sided with Twitter, Google, and Facebook, finding in a pair of decisions on May 18 that the Silicon Valley giants are shielded from liability for content posted by users …. Big Tech and its supporters had been deeply concerned that the court could eviscerate Section 230 of the federal Communications Decency Act of 1996, which generally prevents internet platforms and internet service providers from being held liable for what users say on them … Chief Justice John Roberts said that despite any algorithm YouTube may use to push users to view videos, the company is "still not responsible for the content of the videos … or text that is transmitted."

The Supreme Court's new 38-page decision (pdf) in Twitter Inc. v. Taamneh, court file 21-1496, was written by Justice Clarence Thomas … Thomas wrote that the plaintiffs sought to hold Twitter, Facebook, and Google "liable for the terrorist attack that allegedly injured them," but the court concluded that "plaintiffs' allegations are insufficient to establish that these defendants aided and abetted ISIS in carrying out the relevant attack." The connection between the online platforms and the nightclub attack was "far removed," he wrote …

The "plaintiffs asserted that Google had knowingly permitted ISIS to post on YouTube hundreds of radicalizing videos inciting violence and recruiting potential supporters to join the ISIS forces then terrorizing a large area of the Middle East, and to conduct terrorist attacks in their home countries," according to the family's petition. Because of the algorithm-based recommendations, users "were able to locate other videos and accounts related to ISIS even if they did not know the correct identifier or if the original YouTube account had been replaced." Google's services "played a uniquely essential role in the development of ISIS's image, its success in recruiting members from around the world, and its ability to carry out attacks." The original complaint filed in the case added that "Google officials were well aware that the company's services were assisting ISIS."

2023, May 24:

<<Perhaps I'm naïve, but I find the doom-saying about AI a bit over-the-top. What they're saying is that sources may not be what they seem and that AI makes the detection of fakery harder than before. But this is what we've always known: don't take content on face value, rely on provenance and context.>>

Reporting in the *Weekend Oz* (20-21 May, pp.1 & 7) has the Australian Curriculum Assessment & Reporting Authority (ACARA) warning that students face a "dystopian future" as a result of the transformation of knowledge by AI.

> [The] Chief Executive David de Carvalho has called for greater focus on "facts and truth" in teaching … students needed the "knowledge and wisdom" to detect lies, error, bias and deep fakes generated by AI. The role of teachers as "authoritative sources of information, knowledge and wisdom" needed urgent buttressing.

No disrespect, but I never took any of my teachers as "authoritative sources" for anything. The ones I respected helped me think and look about for myself.

> "In addition to reading, writing, numeracy and digital literacy … ethical understanding, personal and social capability, intercultural understanding and critical and creative thinking are going to be more and more important"

Crumbs! Anything else? Creative thinking. Hmmmm. Who knew?

> *UTS industry professor Leslie Loble … called for urgent controls over the use of AI and for children to be taught to question and control it. "The time is now to set standards … You cannot assume these (AI) tools are accurate" … Emeritus Professor Cheryl Praeger, one of Australia's leading mathematicians [said] students would need strong skills in critical and logical thinking to determine flaws in AI generated solutions … "Students … really need to be able to discern and critique the logic of something …"*

If education hasn't been teaching them **that** up to now, it's a bit late isn't it?

And predictably, The Australian has revealed left-wing political bias in Google's AI chatbot, Bard, which praises Labor Prime Minister Anthony Albanese as a "man of the people" while labelling Liberal leader Peter Dutton as "controversial". Strewth!

> A British teacher has been banned from teaching after complimenting his class by saying "*Well done, girls!*" when they did good. The grip that left-authoritarianism seems to have on education now is such that I doubt critical thinking is going to get much of a show. The balance between conformity to and delivery of orthodox opinions and the encouragement of inquiring minds has always been a hard one to find and the triumph of dogma has been a long time coming. It's generational, they say, and it's been developed in students at university who have themselves become today's teachers. What's encouraging is how many good teachers seem to baulk at it.

What I hate most about left-authoritarianism's subversion of the ABC and other media outlets – they're not in thrall yet (maybe the *Guardian*) but they're certainly deferential – and what I hate most about it is that they're beginning to make Sky-After-Dark look good (well, better anyway).

2023, May 25:

Another take on AI from Alan Kohler, arguing that we are in dire straits because productivity (the thing that drives prosperity and keeps the lid on inflation) is in free fall. He thinks this is because workers (us) are demoralised and that AI might save the day. Kohler says no one really knows why productivity is falling (here and elsewhere) and this is only his best guess. Seems a bit fanciful to me. I am still unclear, perhaps because everyone seems to be unclear also, whether AI is predicted to be **doing** the work or just **helping** to do the work.

> Between March and December last year, Australian productivity – GDP per hours worked – actually declined 4.1 per cent. Taken from before the pandemic, it has been flat … During the late 1980s and '90s, productivity growth peaked at 3 per cent a year. Over the past 30 years it averaged 1.6 per cent; in the past 20 years, just 1.2 per cent, and … it has been zero lately, and then negative … The reason labour productivity (GDP per hours worked) has been flatlining is because workers don't care any more. The decline in productivity growth over the past two decades has coincided with a decline in real wages and a rise in the difficulty of getting a pay rise … it is an ironic, unintended consequence of the competition reforms during the 1980s that produced such high productivity growth at the time … after a while the fun wore off, especially as it became clear they were getting nowhere against China and the hard work was not resulting in higher salaries, and with house prices rising they were falling behind …
>
> Having declined steadily for decades, productivity growth suddenly collapsed during the pandemic and then went negative last year for two extra reasons – the advent of working from home and the sudden drop in real wages as a result of the spike in inflation, followed by the rise in interest rates coupled with higher rents and house prices that are now rising again … I think the only hope for improving productivity will be generative artificial intelligence … Goldman Sachs recently predicted that AI would at least partially replace two-thirds of all jobs … If AI can be used alongside human beings and make their life easier, then human happiness and productivity could both rise. Statistical productivity, minus human happiness, will rise if another Goldman Sachs' prediction comes true – that AI will actually replace 300 million jobs worldwide. As long as the work done by AI machines is not counted in the ABS's

"hours worked", then the GDP per hour of toil by the few humans that are left will increase a lot. Those displaced will of course have to go on welfare and blow out the budget deficit again …

Geeze, this is a long way from the filing rules but, if classification lies at the heart of **our** mystery, I imagine that AI might be quite good at it -

> *"The librarian was explaining the benefits of the Dewey decimal system to her junior--benefits that extended to every area of life. It was orderly, like the universe. It had logic. It was dependable. Using it allowed a kind of moral uplift, as one's own chaos was also brought under control. 'Whenever I am troubled,' said the librarian, 'I think about the Dewey decimal system.'*
> *'Then what happens?' asked the junior, rather overawed.*
> *'Then I understand that trouble is just something that has been filed in the wrong place. That is what Jung was explaining of course--as the chaos of our unconscious contents strive to find their rightful place in the index of consciousness."*
> — **Jeanette Winterson, Why Be Happy When You Could Be Normal?**

And the work of the archivist, if Sir Hilary is to be believed, has always been (like some say AI may be) about **helping** rather than **doing**:

> *"The Archivist's career . . . is one of service. He exists in order to make other people's work possible, unknown people for the most part and working very possibly on lines equally unknown to him, some of them in the quite distant future and upon lines as yet unpredictable. His Creed, the Sanctity of Evidence; his Task, the conservation of every scrap of Evidence attaching to the Documents committed to his charge; his Aim to provide, without prejudice or thought, for all who wish to know the Means of Knowledge."* - **Sir Hilary Jenkinson in his published address "The English Archivist: A New Profession" p. 38.**

2023, May 27:

Urgent! Urgent!! Urgent!!! Are "they" coming to get us?

The *Guardian* columnist, Jonathan Freedland, who certainly thinks AI is about **doing** and not just about **helping**, seems to believe that it may be an even greater threat to human civilization than pronouns:

> … new technologies often freak people out at first … Better, surely, to focus on AI's potential to do great good … typified by this week's announcement that scientists have discovered a new antibiotic, capable of killing a lethal superbug – all thanks to AI [but] it's not just lay folk like me who are scared of AI. Those who know it best fear it most … Geoffrey Hinton, the man hailed as the godfather of AI for his trailblazing development of the algorithm that allows machines to learn [has] resigned his post at Google … confessing regret for his part in creating it … In March, more than 1,000 big players in the field … issued an open letter calling for a six-month pause in the creation of "giant" AI systems, so that the risks could be properly understood.
>
> What they're scared of is a category leap in the technology, whereby AI becomes AGI, massively powerful, *general* intelligence – one no longer reliant on specific prompts from humans, but that begins to develop its own goals, its own agency … As Yuval Noah Harari warned in a recent Economist essay, "People may wage entire wars, killing others and willing to be killed themselves, because of their belief in this or that illusion", in fears and loathings created and nurtured by machines. More directly, an AI bent on a goal to which the existence of humans had become an obstacle, or even an inconvenience, could set out to kill all by itself. It sounds a bit Hollywood, until you realise that we live in a world where you can email a DNA string consisting of a series of letters to a lab that will produce proteins on demand: it would surely not pose too steep a challenge for "an AI initially confined to the internet to build artificial life forms", as the AI pioneer Eliezer Yudkowsky puts it. A leader in the field for two decades, Yudkowksy is perhaps the severest of the Cassandras: "If somebody builds a too-powerful AI, under present conditions, I expect that every single member of the human species and all biological life on Earth dies shortly thereafter."

> ... AI is learning so fast, how on earth can mere human beings, with our antique political tools, hope to keep up? That demand for a six-month moratorium on AI development sounds simple – until you reflect that it could take that long just to organise a meeting ... This is yet another challenge to democracy as a system, a system that has been serially shaken in recent years. We're still recovering from the financial crisis of 2008; we are struggling to deal with the climate emergency. And now there is this. It is daunting, no doubt. But we are still in charge of our fate. If we want it to stay that way, we have not a moment to waste.

This all goes a good way beyond helping us to compile an index. Someone who knows more about this stuff than I do please tell me this is all nonsense.

2023, May 28:

<< **Someone who knows more about this stuff than I do please tell me this is all nonsense.** >>

Wikipedia (god bless 'em) hazards an answer, albeit a cautious one:

> Many scholars believe that advances in artificial intelligence, or AI, will eventually lead to a semi-apocalyptic post-scarcity economy where intelligent machines can outperform humans in nearly, if not every, domain. The questions of what such a world might look like, and whether specific scenarios constitute utopias or dystopias, are the subject of active debate.

High above the hushed crowd, Rex tried to remain focused.

Still, he couldn't shake one nagging thought:

He was an old dog and this was a new trick.

I would love to see **an** archives conference (and attend it for that matter) in which there was "active debate" (haven't seen that for a while in my chosen profession, possibly because I no longer mix in the right circles) on the likely/possible effects of AI on our work both conceptually and functionally. Meanwhile (nudge, nudge) I would love to have registered with **the** forthcoming archives conference in Melbourne and have received a receipt for payment of my registration fee (which I've had to ask for, so far without success - how dystopian is that?).

Naturally sceptical of either utopias and dystopias, I am glad I am retired and in any case probably won't now live long enough to have to deal with either of them r/k-wise (AI inspired ones, that is, there's still plenty of others to go around).

<< **Naturally sceptical of either utopias <u>and</u> dystopias** >>

How could I have done that? Fowler would not have approved.

🏠 MUCH ADO ABOUT DESCRIPTION

2023, May 29:

If the first AI-generated finding aid has not yet made its appearance, can it be delayed much longer? In the last little while, a deluge of on-line articles has come to the top of any Google search for the risks and benefits of AI:

- [Do the benefits of artificial intelligence outweigh the risks?](#) (*The Economist – Sep., 2018*)
- [Ethics of AI: Benefits and risks of artificial intelligence](#) (ZDNET – Apr., 2021)
- [Benefits & Risks of Artificial Intelligence](#) (*Future of Life* – Nov., 2015)
- [Pros and cons of AI: is Artificial Intelligence suitable for you?](#) (*Data Conomy* - Aug., 2022)
- [Opportunities and risks of ChatGPT in medicine, science, and academic publishing: a modern Promethean dilemma](#) (National Library of Medicine – Feb., 2023)
- [Risks and remedies for artificial intelligence in health care](#) (Brookings – Nov., 2019)
- [How do you balance the benefits and risks of AI for your personal and professional goals?](#) (Linkedin – n.d.)
- [The Risks and Benefits of Using AI to Detect Crime](#) (*Harvard Business Review* – Aug., 2018)

etc., etc., etc. And these are just from the first page of results. Note: lots of these are not part of a <u>recent</u> "deluge" but go back many months, years even, and this search has pointed to results which are almost exclusively social rather than technical.

An article in the *Weekend Oz* (Do the dangers of AI risk smothering its benefits?) has the now familiar catalogue of perils (real or imagined) including the alarming bon mot that "AI has the knowledge to pass a medical exam". This article focuses on the response of politicians and education bureaucrats which predictably involves meetings and the development of "guidelines" (god save us all). But buried in the dross are some revealing nuggets – expert opinion coming to the view that the answer may involve changing teaching methods and nurturing good teaching aspirations (how to know rather than what to know). Music to the ears of one who has always believed, ever since flirting briefly with becoming a teacher, that critical thinking lies at the basis of good pedagogical method.

It needs no politicians nor guidelines to tell us this: supervised exams, oral tasks/assessments, "a push away from recall and facts and … towards process and comprehension, towards understanding and use of information", knowing "what's true and real". How sad that these remedies are seen as innovative.

The underlying theme (relevant, I suggest, for both education and archival description) is that **we** must be in charge, not the machine. If concerns about AI (whether justified or misplaced) get educators (and archivists) to focus on truth, facts, and critical thinking and away from critical theory and social justice for a bit, that is all to the good say I. Of course, educators (and archivists) can whittle and chew gum at the same time, so they can go on doing both. It's not a social conscience that I object to; what I object to is the behaviour of some of those with one.

PS. I can't help reflecting that a good deal of this alarm echoes what I can recall of the time when the Internet became a standard educational resource. That's how old I've become.

2023, May 30:

 <<[Andrew Waugh:](#) **(Anyone reading this should take Arthur C. Clarke's strictures on elderly scientists saying things won't work…) … LLM (Large Language Models) generate text by first taking what they've already written and calculating what the next most likely word will be. Randomly (based on a preset probability) it doesn't chose the absolutely most likely word, but one of the very likely words … But very little real world non fiction writing is like that, and certainly not writing finding aids. Each finding aid is unique and the details in it depend on the context … I expect this to be a fundamental limitation of LLMs, because they have no real world knowledge. They're just stringing together words that sound right. The domains we have to watch is where following the**

facts are not important … entertainments aren't constrained by facts, and humans love to be entertained … More darker would be conspiracy texts - there are lots of examples of those on the Web and they can be completely unmoored from reality. Perfect for generation in bulk by ChatGPT.>>

2023, May 31:

<<<u>Chris Gousmett </u>Thanks Andrew, right on the button. The fears about AI waging war on humanity and wiping us out fundamentally misunderstand what AI (specifically LLMo is capable of doing … If we can't discover how and why the results are created, then what trust can we have in what is produced?>>

2023, June 3:

<u>Do we really need another threat to humanity?</u>

2023, September 28:

Merriam-Webster has identified 690 <u>new words</u>. Apparently, they do this annually. This latest batch includes a new meaning for an old word related to AI:

> <u>**hallucination**</u> *noun … ***3 :** a plausible but false or misleading response generated by an artificial intelligence algorithm

and, not unrelated to some of the commentary about AI:

> <u>**edgelord**</u> *noun, slang* **:** someone who makes wildly dark and exaggerated statements (as on an internet forum) with the intent of shocking others

You do have to wonder about the <u>shelf-life</u> of some of the new words on offer. Many of them seem (to me) to belong to one or more sub-cultures. But that may simply be because I'm old. Are "<u>square</u>" (8) and not "<u>hip</u>" (adjective) <u>passé</u> now?

2024, May 2: <u>You Can't Get to Dublin From Here</u>

<u>Well, sir, if I were you, I wouldn't start from here.</u>

I recently did a search in <u>Archway</u> (Archives NZ online search tool). I got a return of 512, 121 results. Got me thinking (again) about searching. When we began designing GLADIS (government-locator-[for]-archives-[and]-documentatiion-information-system) the underlying concept was based around the triangulation of Document-Deed-Doer to enable users (whatever their starting point) to navigate their way to an outcome (the right

outcome) by taking only three (or at most four) steps. The right outcome was conceived to be a manageable number of results that included all relevant material. It was conceded that, in some cases, deeper filtration might be needed and that no search could guarantee that all relevant material would be found. [The Minister hated the name "Gladis" so it had to be changed.]

Above all, it was the intention to avoid "flat" searching – the kind that results in 512,121 hits. An integrated taxonomy to resolve language and terminological difficulties, including archaic terms and names, would augment (but not replace) the search architecture. After Peter Orlovich's funeral last Monday, I was musing with a colleague about how conspicuously archivists have failed to realise the potential of online searching. It was a conversation Peter would have understood and approved of. The basic point is that archives/records are different and effective searching is built (or should be built) upon knowledge about them (description). This knowledge is essentially about structure, not content. Even when describing the Documents, it is the process (of which they themselves are the by-product) that should be the focus. I have alluded to all this from time to time and on Monday I said ruefully that, if I were younger, I would write more about it – but those days are now past.

I was put in mind of this when I came across an announcement that the British National Archives is "delighted" to be putting 300,000 farm records online – at a cost of £2.13 million. How often do we celebrate the digitisation of content and how often do we comment on what is being done to make it accessible? As if digitisation, in and of itself, is all that is necessary. How much of that money, I wondered, would go to making the stuff accessible rather than just digitising them? When will we (never mind anyone else) understand the term "digitisation" to include description?

It took me back many years to when I was studying in London for my archives diploma at UCL. A course requirement was doing a major project. I chose to prepare a finding aid for the Empire Marketing Board because, in part, I had been working on related records back in Canberra. Imagine my surprise and delight when I discovered that part of its function involved setting up a Film Unit that eventually became the Crown Film Unit. At that young age, I was passionate about film.

The Board's records (now at Kew) were held in Portugal St (a block away from PRO's HQ in Chancery Lane). Imagine creaking floors, unhelpful attendants in grey dust coats, and an Assistant Keeper sitting at a desk on a podium in the search room doing his own paper work and available (grudgingly) to assist if you didn't mind the icy stare. Using the online catalogue to find records **of** (as distinct from records **about**) has even now caused me no little difficulty. At one point I found them online but then I lost them again and couldn't easily find my way back. They consist of five series (CO 758, CO 759, CO 760, CO 868, CO 956). When I was there in 1975, I was permitted to call for three items at a time, making it very difficult to use the index in conjunction with minutes and correspondence.

🏠 MUCH ADO ABOUT DESCRIPTION

It was the Index (CO 759) that gave me special bother. The cards were bundled up in brown paper parcels bound with string. There were two or three bundles per box (called pieces on the quaint notion that one day they would get around to binding into a single volume the contents of each of the millions of boxes flooding into PRO as a result of the avalanche of "modern records" – bit difficult with 8x5 cards, I thought). At first, they let me have one box at a time (because each box had two/three bundles). Then they conceded that it might be more sensible to let me have three boxes at a time because when a bundle was undone in hopes that it might contain a card on "Wool", chances were there would be a card saying "See sheep". At long last, I got access to the stacks (because I was technically on secondment as part of the peculiar arrangements surrounding my stay in London – see About Me on my website) and I could rummage around for myself (taking care to rewrap each bundle after I'd used it).
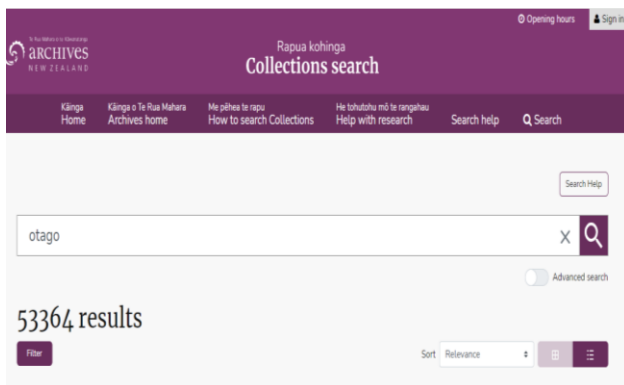
Archway
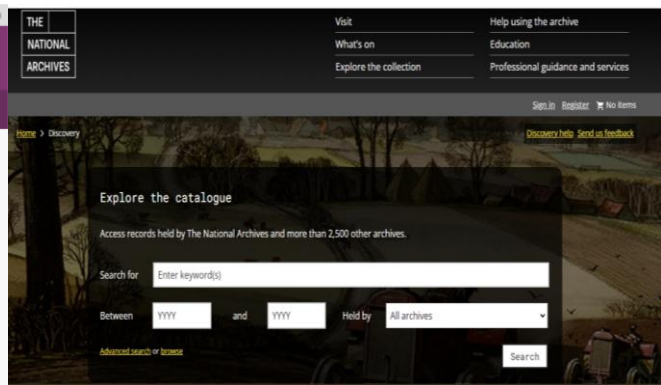
Discovery

CO 759 is still described as comprising 134 bundles. There are two sub-series (nominal and subject) but these do not appear to be physically separate. The side bar lists the contents as 759/1 (Aa-Ak) to 759/72 (Wr-Wz). The "details" for each of these entries is no more than a screen repeating what is in the side bar and how 134 bundles become 72 descriptions eludes me. Is it possible (is it really possible) that they are listing the boxes rather than the bundles? In any case, so far as I can recall, it is a lot better than I had available to me in 1975.

EMB Advertising

"Drifters"

PS. As part of my extra-curricular travels in 1974/75, I invited myself to the British Film Institute. I spent many happy hours at their Southbank screenings which I am glad to see are still going strong. They were delighted to see me because they'd been trying (they said) to make contact with their Australian counterparts for some time. I worked out afterwards that they were labouring under the delusion that I worked for the Film Section of NLA (which later became the Film and Sound Archive) and that the person they really needed to make contact with was Ray Edmondson. Anyway, as I always did on such visits, I got them as quickly as I could to show me their finding aids. We were chatting in a general way and I took the opportunity to ask if we could look something up. I suggested searching

for *Drifters* in the catalogue. In case you don't know, this pioneering documentary by John Grierson is legendary amongst film buffs. I was hugely amused when they couldn't find it.

2024, May 4:

<< **I was musing with a colleague about how conspicuously archivists have failed to realise the potential of online searching**>>

Searching the finding aids used to be simple – find a page (or card) and look it up. The description was (or purported to be) a truthful observation of the features that made the record (for us) worth keeping and (for the user) worth finding, expressed in a way that met (or was supposed to meet) the users' needs. We assumed they wanted to know what we had observed about the record (but principally about the event or circumstance that brought the record into being) and that we were correctly assuming what they would want to know and expressing it in ways most helpful to them.

A lot of assumptions in all of that. The skill lay in execution and delivery but the end-product (the finding aid itself) seemed unproblematic – words on a page or card that any one could read provided they found their way to the right place, connections with the stuff being described and with further layers of description (maybe leading onto other stuff). Being archivists, we understood that the object of a search was not the stuff for its own sake (unless it was illuminated manuscripts) but for the meaning (truth, if you like) that it embodied. The record was valuable for what it told us about life and our job was to help reveal that (and guard it from deceit).

We knew that a meeting of minds between the archivist and the user was not an uncomplicated matter and, as we moved from the pursuit of academic objectivity into more complex realms of social responsibility, we developed a sensitivity to diverse cultural ambiguities about content and the way it was represented in the finding aids. Some of us even became champions of r/keeping in the service of ideological virtue. But we did not tamper with the artefact and our core descriptions of it had to conform (in some complicated and conflicted way) to the professional dictum: *add nothing to and take nothing from*. This did not, however, prevent the construction of "ancillary" aids tailored to users' various needs.

But after all that, the instrument itself and the using of it were fairly straight-forward, essentially no different from the way our early ancestors communicated with each other using rock paintings. The mechanics were fairly simple: linguistic conventions, alphabetisation, numerical order, indexes, taxonomies, classification, categorisation of several kinds, understanding relationships and hierarchies, the way information was presented (diplomatics). Once learned, in an agreed set of cultural norms, they assisted the process of communication between the archivist and the user.

A breezy assumption that online searching simply augments and enhances our existing methodologies has led to a lazy utilisation of generic search technology, hence 512,121 results. More profoundly, it blinds us to the need to adapt our methods to what I might call the hidden benefits and hazards of search engines. One of the reasons I shouldn't be writing about this is that I have virtually no understanding of how these darn things work. I once tried to figure it out (*Strength Below and Grace Above*) but I don't think I made a very good job of it. I do understand (or think I understand) that one of their essential characteristics is that they take the user's experience of our descriptions out of our hands in ways that was never the case before and in ways we struggle to comprehend. The only writer I know who has delved into this in ways I can understand is Tim Sherratt (no doubt there are others). From our customary perspective, all this can be both good and bad – cf. *Structures, Boundaries, Contingency, and Proportion Are Good for Recordkeeping*.

🏠 MUCH ADO ABOUT DESCRIPTION

The underlying algorithms are a characteristic of online descriptions that are a determinant of the experience. They can be used to control it. Governments (e.g. China, Australia, Russia, India, etc.) use them to promote or suppress (and even block) results. Archivists could use them to "guide" the user experience in ways that might be helpful or manipulative (always with the best intentions, of course). They give new meaning to age-old issues: e.g. do we keep records of users and what they've accessed (cf. Google vs DuckDuckGo)? So, maybe things haven't changed all that much after all. We've always had to understand how our descriptive tools functioned and to make assumptions about how they would be used. The landscape has changed, but the issues haven't. Or, have they ......

<< Governments (e.g. China, Australia, Russia, India, etc.) use them [algorithms] to promote or suppress (and even block) results.>>

I have long been puzzled over how Freedom of Speech became a Right-Wing issue. The best I can come up with is this: Freedom is a threat to Power. That is why totalitarians of all stripes hate it.

Tacitus told us this

As Napoleon surely understood, Tacitus was no admirer of empire. But more importantly, what Napoleon also seemed to have understood was that in his principal writings, Tacitus aimed at making visible the soul of the tyrant ... It is by unmasking the tyrant that Tacitus brings to light the character of moderate republican life, long since lost to Rome. It is perhaps for this reason that Milton claimed Tacitus as "the greatest possible enemy to tyrants," and many of the American founders admired him greatly.
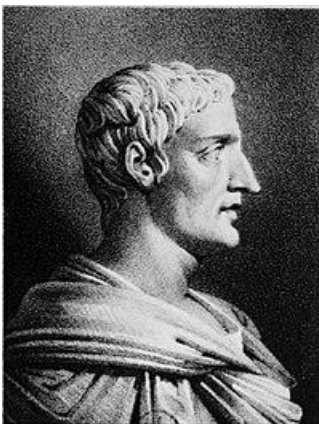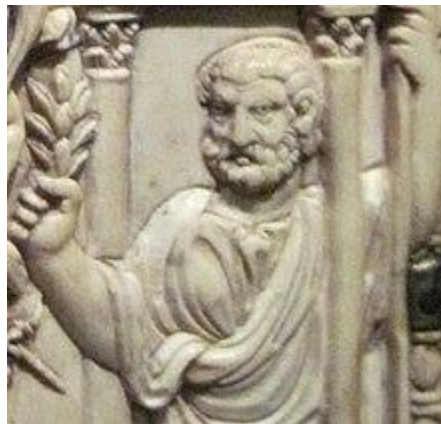
| Tacitus | Symmachus | Dickens |

Symmachus told us this

In the late fourth century, the Roman senator Quintus Aurelius Symmachus, a pagan, issued a plea for religious pluralism: "We gaze up at the same stars; the sky covers us all; the same universe encompasses us. Does it matter what practical system we adopt in our search for the Truth? The heart of so great a mystery cannot be reached by following one road only."

Dickens told us

Because Dickens nourished an uncompromising contempt for every kind of tyranny, it was inevitable that he should denounce American slavery, whose essential barbarity he observed on his first trip to the New World in 1842. From that time until the conclusion of his second visit to the United States in 1868, a period of roughly twenty-six years, he focused his attention, at intervals, on the issues which grew out of the system.

Acton told us

It is bad to be oppressed by a minority; but it is worse to be oppressed by a majority. For there is a reserve of latent power in the masses which, if it is called into play, the minority can seldom resist. But from the absolute will of an entire people there is no appeal, no redemption, no refuge but treason. The humblest and most numerous class of the Athenians united the legislative, the judicial, and in part, the executive power. The philosophy that was

then in the ascendant taught them that there is no law superior to that of the state, and that, in the state, the law-giver is above the law.

and **Orwell** told us

> Orwell was wrong about many things, but he was right to suggest that a world that turns its back upon truth also gives two fingers to freedom. But how we should legislate for this fact is no clear matter. The Saudi regime has just warned its citizens [2018] that the circulation of "fake news" is a criminal offence. With friends like this, truth no longer needs enemies.
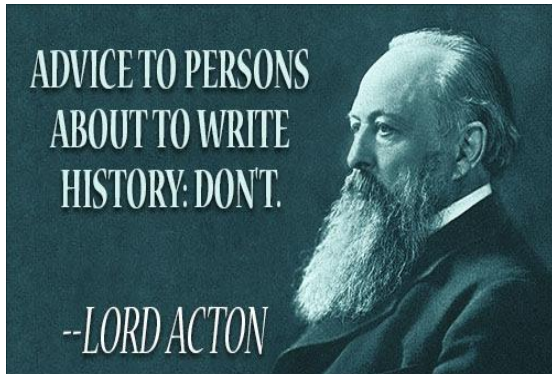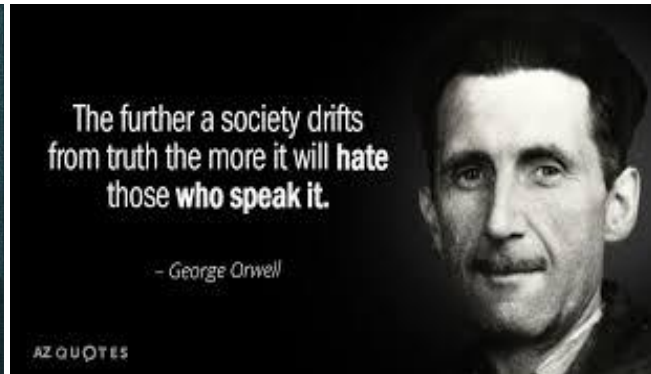
| Acton | Orwell |

How many times do we have to be told? When I was a lad, the Right was on top and Free Speech was a Left-Wing tool against Oppression. Now that I am an old man and the Left is on top, it has become a Right-Wing tool against Woke. Simplistic perhaps, but I can't think of any other explanation.

On a less highfalutin plane, however, it comforts me (metaphorically) in my belief that descriptive practice must find (and keep on re-finding) a balance between Ubiquity (Freedom) and Structure (Tyranny).

2024, May 20: Directory of Archives in Australia

When you're retired, you have time to surf the Net (does that phrase date me?) and to make a nuisance of yourself. I've stumbled across the ASA's *Directory of Archives in Australia*. I've always thought this thing had a lot of unrealised potential but it's been off my radar for a few years now. The blurb states (inter alia):

> The 2018 edition … includes a set of taxonomies to assist searching via Archive Type, Geographic Region, Locality and Thematic Area.

The Thematic Area interests me but I can't find out any more about it. Is it really a taxonomy or just a word list? How many terms does it have? What sort of control, if any, is there over (a) meaning or (b) use? Very little, I assume, if contributors decide for themselves. Is there elaboration (scope notes) anywhere? Is it maintained and updated and (if so) are revisions retro-fitted to existing entries?

If you browse "Thematic Area" in Advanced Search you get 32 results:

> Aboriginal and Torres Strait Islanders; Agriculture; Arts and Culture
> Built Environment
> Communication
> Defence
> Education; Environment
> Family/Domestic Life; Foreign Affairs
> Genealogical; Geography
> Immigration; Indigenous; Industry, Manufacturing and Commerce
> Labour; Language; Law and Justice; Local History; Local Studies
> Media; Medicine and Health
> Natural Resources
> Out of Home Care

Politics and Government; Populations
Recreation/Leisure/Sports; Religion
Science and Technology; Social Organisations and Activities
Transportation; Travel and Exploration

It's a curious list. What, for example, is the distinction between *Aboriginal and Torres Strait Islanders* and *Indigenous*? Where is *Banking and Insurance*? Where is *Welfare* (other than *Out of Home Care*)? I assume there is some intended sloppiness in the hazy boundary between intended research use (e.g. *Genealogical; Local Studies*) and functional genesis (e.g. *Agriculture*; *Local History*).

2024, October 10: Archives Canada

From the Canadian List:
(Hope they'll forgive re-posting. Enthusiasm must be my excuse.)

Why can't we do something like this on our side of the world? The *Modest Proposal* was a *cri de coeur* based on an assumption that we just needed to talk about how to do it. But forget all that. It would be enough to agree that we need to.

> **Mary Kosta (9 Oct., 2024):**
> ARCHIVESCANADA.ca Working Group Announced
> Ottawa, ON, October 8, 2024-The Canadian Council of Archives has constituted a Working Group to assess and re-envision the national finding aid network, ARCHIVESCANADA.ca. … The members are tasked with:
> - Developing a snapshot of users (researchers, archives patrons, archivists) and their needs.
> - Assessing the current status of, and problems with, ARCHIVESCANADA.ca.
> - Determining whether there are solutions that will improve the status quo.
> - Evaluating alternatives to the status quo.
>
> The Working Group will deliver a final report on ARCHIVESCANADA.ca to the Canadian Council of Archives on the completion of its two-year mandate which began in January 2024.

The terms of reference and list of Working Group members is available on the Canadian Council of Archives website. Working Group meeting minutes are open to the public and may be found here …

> **Lois K Yorke (10 Oct., 2024):**
> I well remember the excitement, challenges and high hopes back when ARCHIVESCANADA.CA was first developed and launched, all those years ago. Wonderful teamwork and collaboration! And I also remember when the federal funding for it was reviewed and audited. What a learning experience it was to serve on that committee! Both the archival landscape and access to its rich resources have changed so much since then -- but I somehow suspect many of the challenges are still the same. This is a marvelous opportunity to come together, assess and re-envision for the future. My best wishes to those of you on the Working Group -- if the past is any measure, you will each contribute much and learn a great deal in return.

PS. And you can add – admiration for what they've achieved.

2025, January 12: What, if anything, is an AI Overview?

Is it data scooped up and displayed on Google using AI? If so, what are we to make of this:

> **AI Overview**
> As of January 1, 2025, the cost of a passport in Australia depends on the age of the applicant and the validity of the passport:
> - **Adult passport (10-year validity)**: $596, including an overseas processing surcharge of $184

- **Child passport (16–17 years old, 10-year validity)**: $502, including an overseas processing surcharge of $90
- **Child passport (0–15 years old, 5-year validity)**: $298, including an overseas processing surcharge of $90
- **Senior passport (75 years or older, 5-year validity)**: $392, including an overseas processing surcharge of $184
- **Replacement passport**: $259

I was prepared to go off with a rant to National Seniors that, on the calculation that I will soon be dead, I must take a hit of two "overseas processing" surcharges @ $184ea to get the 10-years validity on a passport that other adults only have to pay once; resulting in a total cost for us oldies every ten years of $784 as against $596. So (I might have concluded) I'm being asked to pay $188 extra for my 10-years on the gamble that I may not live that long.

But, according to the [Australian Passport Office](#), fees as of 1 January, 2025 are

- 10-year validity passport (for **persons aged 16 and over**) $412
- 5-year validity passport (for **children** under 16) $208
- 5-year validity passport (**optional for persons 75 and over**) $208
- Overseas surcharge - adult applicant $184

The "Senior passport" described in the AI Overview is optional (important qualification) and the $188 difference is payable only if I choose to take the morbid gamble on my life expectancy (or, perhaps, a realistic view of my future travel opportunities). The differential actually gives me options and is not a case of age discrimination.

PS. Presentation, they say, is everything and the devil is in the detail.

<div align="center">❧❧❧❧❧❧❧</div>

2016, June 1: OAIS query

I'm trying to teach myself about SIPs, AIPs, and DIPs. Here are four questions which I hope those of you who are across Open Archival Information Systems (OAIS) may be able to answer. I trust they aren't stupid ones but, if they are, please tell me.

1. Can a DIP (Dissemination Information Package) be transformed by the recipient "user" back into a SIP (Submission Information Package) and thence stored as a second rendition of the same digital asset in the form of an AIP (Archival Information Package) in someone else's "archive"?
2. Does an AIP carry a persistent identifier assigned by the source archive that would travel with a DIP and (unless suppressed) be incorporated into the user's data management function if the user chose to preserve a second rendition of the digital asset?
3. How much of the AIP's descriptive information (used to manage the digital asset in the system that supports the source archive) travel with the DIP as metadata?
4. Can the metadata alone be harvested and transformed into a portal link back to the digital asset in the source "archive"?

Is it simply that all these things are possible (or not), irrespective of whether or not either the source or the user follows OAIS, and that the OAIS model doesn't deal with them and doesn't have to?

<<Andrew Wilson: OAIS explicitly envisages and allows for transforms of DIPs into new AIPs. The fact that that event might take place in a different repository is not specifically covered but inter-repository transfers seem to be envisaged in the discussion in section 6 on Archive Interoperability. Your other 3 questions seem to me to be essentially implementation decisions rather than being specified in OAIS. There

isn't anything in OAIS (afaik) that would preclude an archive from making any of those decisions about its implementation of OAIS.>>

2016, June 2:

<<Lise Summers: Is it OK if I send this on to the team at Artefactual, who deal with OAIS, SIPs, DIPs, etc, via the Archivematica software? I think these are really interesting questions, and I can answer from the perspective of someone who looks at the OAIS model, on an annual basis, from a fairly simplistic perspective.

OAIS is, first and foremost, just a model. How well the system based on the OAIS model will be able to determine what to keep, copy and disseminate will depend on how it is programmed (GIGO still applies).

1. yes, a DIP can be capture as a submission information package in another repository. Remember, this model was developed by the astronomical community who wanted to share their information around, yet keep it stable until someone had the time to look at it and analyse it. However, because a DIP may include both 'normalised' and access modified objects, it is not necessarily going to have the richness of the original digital object in the first system. Have a look at the digital information available on various archive and library sites, and the access or dissemination objects have been reduced or compressed to aid access via the web. I can easily see a scenario where someone captures the plans SROWA has available on their website, along with the metadata available via AtoM, into their own repository for further work, for example.

2. Yes, the SIPs, AIPs and DIPs have unique identifiers, which could be captured along with the objects and the metadata. It's not a DOI though, or a handle, so would not necessarily resolve back to the originating archive (at least, I don't think so.) Short term probably, long term no. Although you may be able to do so, with the right set up.

3. Depending on the sytem and software you use, you can configure the metadata captured in the AIP for inclusion or exclusion in the DIP. You may not want all the detail though - think accession and preservation metadata in archival description generally.

And yes, you can capture the metadata alone, without the object and make that a new submission package. Just download an xml report from the ANU report, or TROVE, or... and make that your submission information package.

Again, this is about the software you choose and the choices you make in selecting and accepting material. OAIS doesn't really do appraisal.

2016, June 3:

By all means send it on. The more information the better. Thank you, Andrew, for your reply.

<<Lise Summers: ... One thing I forgot to say is that when the DIP is ingested into a different repository it will carry it's original uuid with it, but will also get a new uuid as part of the ingestion and AIP process. This is because it is a new instance of that object - things have happened to it since it's original ingestion. If you want some sort of persistent identifier, or something that says this object is like this object, then Peter van Garderen's post on distributed archival collections seems to me to be discussing how that might be achieved.>>

2016, June 22:

<<Sara Allain (Systems Archivist, Artefactual Systems) I think that Lise answered most of your questions, but I can speak to how Archivematica (a popular open-source tool that creates OAIS-compatible digital preservation packages for long-term storage) handles the four issues that you brought up. Most of them, as Lise mentioned, are not part of the OAIS functional model, which envisions a direct line from submission to access - no detours allowed. However, in practice, we've learned that detours are necessary to ensure that Archivematica users are able to do what they need to do with

their data, and that means that we've built in capacity for some of the workflows that you've asked about. Answers about how Archivematica handles these workflows are in purple.

1.   Can a DIP (Dissemination Information Package) be transformed by the recipient "user" back into a SIP (Submission Information Package) and thence stored as a second rendition of the same digital asset in the form of an AIP (Archival Information Package) in someone else's "archive"?

Yes - I think Lise does a great job of explaining why this is possible (and why it's so important that it *be* possible) above. Any DIP that is created in Archivematica, which will include things like access copies and image thumbnails, for example, can be download and turned into a new SIP at a later date - the DIP wouldn't necessarily be treated any differently than a non-DIP bundle of content. Archivematica also lets you transfer items that have existing derivatives, like access or service copies, and it's possible to include previously-created digital preservation mechanisms, like checksums, with the transfer to ensure that the chain of preservation fixity is not broken.

However, it's not currently possible to automatically/systematically reingest a DIP to create a new AIP - it has to be downloaded and then transferred back into the system. In Archivematica 1.5, we've added functionality to reingest AIPs for DIP generation and metadata updates (sponsored by the Zuse Institute Berlin) - the reingest process is documented in the relevant METS.xml, creating a link between the original and the reingested content. Future releases will see this feature expand greatly - and it's certianly possible that the same could be done for DIPs in the future, if there's a community need (message for everyone: get in touch if this is an area where you'd be interested in sponsoring development!).

2.   Does an AIP carry a persistent identifier assigned by the source archive that would travel with a DIP and (unless suppressed) be incorporated into the user's data management function if the user chose to preserve a second rendition of the digital asset?

Archivematica always creates unique universal identifiers (UUIDs) for every piece of content that is transferred into the system. This is required for Archivematica to carry out its many preservation actions, and for future fixity checking. As Lise pointed out, though, the UUID isn't meaningful - it's just a random, long alphanumeric string intended to ensure that each item is identifiable in the system. Archivematica users usually assign more meaningful identifiers as filenames within a SIP - for example, for archival content, they might use the accession number + any digital object identifier that is used locally.

In the scenario above, where a DIP was downloaded and reingested into Archivematica, the system would create new UUIDs for the content. This is desirable functionality because the new DIP items are different than the older items from whence they came - they've been normalized. Stuff has been done to them. Archivematica sees them as completely separate items with no semantic link. This could be changed if there was an automatic way of reingesting a DIP, as it is for the AIP.

That said, the filename that used more meaningful numbers and letters could certainly be consistent across iterations of the DIP. The filename is always preserved, but with the UUID appended to it. I suppose you could have successive UUIDs appended to the filename, but I think you'd max out the filename limit pretty quickly!

3.   How much of the AIP's descriptive information (used to manage the digital asset in the system that supports the source archive) travel with the DIP as metadata?

Archivematica supports Dublin Core metadata. If you use a description management system that allows you to export descriptions as Dublin Core, this can be added to the transfer as a CSV and ingested into Archivematica. This is optional, though - a lot of users don't include metadata at all, or they just include metadata about the transfer as a whole rather than metadata for each individual object. Use cases differ by institution, but the functionality is there.

If metadata is included in a transfer, it is also included in the DIP. In Archivematica, it's possible to send a DIP directly to an access system like AtoM, and the metadata would go with it. In this case, the user would probably create the Dublin Core CSV file first and then upload the item to AtoM along with its metadata.

4.   Can the metadata alone be harvested and transformed into a portal link back to the digital asset in the source "archive"?

41

🏠 MUCH ADO ABOUT DESCRIPTION

Archivematica is content-agnostic, so having a single XML file in a transfer is totally valid. The workflow that Lise described would work!

You can read more about Archivematica's approach to implementing the OAIS functional model on our wiki, here: https://wiki.archivematica.org/Overview. I won't summarize that now, but if there are any further questions about how the Archivematica project approaches and implements OAIS, I'd be happy to answer them!>>

Egad! Records in Context

2016, September 4: ... Records in Context draft standard ...

ICA Experts Group on Archival Description:
Records in Contexts draft standard released for public comment

<<Adrian Cunningham: Sent from my iPhone

The ICA Experts Group on Archival Description (EGAD) is pleased to announce the release for public comment the initial draft of the first part of a two-part standard for archival description named Records in Contexts (RiC). When completed, the standard will include a conceptual model (RiC-CM), and a formal ontology (RiC-O) ... At the ICA Congress in Seoul, Korea in September 2016, three representatives from the EGAD will provide an overview and introduction to RiC ...

Since 2012, with members from thirteen countries, the EGAD has been developing the new standard for the description of records based on archival principles. In the course of its work it has taken into consideration critiques of current practices, established and emerging national conceptual models as well as the models of allied professional communities, and the opportunities presented by new and emerging communication technologies. The objective of the standard is to reconcile, integrate, and build on the four existing standards: General International Standard Archival Description (ISAD(G)); International Standard Archival Authority Records—Corporate Bodies, Persons, and Families (ISAAR(CPF)); International Standard Description of Functions (ISDF); and International Standard Description of Institutions with Archival Holdings (ISDIAH)...>>

2016, September 7:

<<Evelyn McLellan <eve...@artefactual.com> ... I wonder if there isn't some better way to solicit and share feedback? Sending email is a one-way street that doesn't facilitate community discussion in the same way that, say, sending feedback to this listserv (or some other public forum) would. Another option might be to run a survey and make the results available. I think it would be very useful for to see the comments that come from other members of the profession and to be able to respond to those comments publicly.>>

2016, September 11:

Just returned from Seoul where I listened to the presentations but I haven't reviewed the document yet. I agree with Evelyn McLellan and, so far as Oz is concerned, I would like to see ASA develop a coordinated response as we did in the early 1990s when ISAD was first mooted. I have begun to wonder whether such standards are useful at all but I suppose they are helpful as an expression of common purpose, in teaching, and in software design. My initial thoughts on RIC:

1. In 1992 Australia asked for a reconsideration of the *Principles* ... on which ISAD was based but it was declared to be an "historical document" and debate on it closed. Accordingly, the last 25 years has been about implementing a conceptual model which was (initially) incompatible with series-based description. Over the years, bits of series-based practice have been grafted into the ISAD suite but without a unifying theoretical underpinning. For this and other reasons the existing ISAD suite

became incoherent (which was why they had a review, I thought). Gratifying then that RIC-CM re-opens debate on the conceptual model and from what I heard in Seoul the intention is to keep it open. Very good.

Principles of Description*

- "Multilevel Description"
  - Proceed from general to specific
  - Provide information relevent to the level of description
  - Link each level of description to next higher unit of description
  - Do not repeat information, provide it only at highest appropriate level

2. I thought I heard it said that a multi-dimensional (entity-relationship) model will effectively replace the multi-level approach which, though more honored in the breach than in the observance, remains ostensibly the basis for ISAD. This is something else Australia asked for way back when because while multi-level description (fonds based) can fit within multi-dimensional description (series based) it can't work the other way round. Multi-level description would still be possible, for those who wish to do it that way, but merely as a particular, albeit limited, implementation of the multi-dimensional approach.

3. For that reason, ISAD, ISAAR, and ISDF and the existing descriptions based on them remain valid although I thought they were saying they will no longer be maintained. As far as I can see, however, the ill-conceived ISDIAH must go (and good riddance). Those still describing fonds may have to re-imagine "authority records" as entities but that shouldn't be a leap.

4. The aspiration for ISAD to be a vehicle for sharing and integrating archival descriptions in networked space continues but it's not clear that RIC will take us any closer than the ISAD suite did. It depends a bit on how they regard ambient authority and on their entity-type definitions. I heard mixed messages on whether or not RIC is supposed to be restricted only to records in archival care.

5. The long-neglected area of relationships has been addressed and nearly 100 relationship-types are enumerated - with the expectation of more to come! Presumably a relationship or two ("in custody of", etc.) will replace ISDIAH as it should always have done. The defining characteristics of a relationship (reciprocal : showing when related and how related) are followed but it wasn't clear from the presentations that they are being consistently crafted according to rule and the same relationship looked to be duplicated, in some cases, for past and present links (it would be odd if that were the case). I wonder if the notion of treating relationships as entities will get another run?

6. RIC hasn't been adopted yet but if it is Artefactual will have some work to do in the further development of AtoM.

2016, September 12:

CORRECTION:

I've now had a chance to glance through RiC-CM. Yesterday, I said **<< nearly 100 relationship-types are enumerated >>**. I don't know where I got that figure from. There are 792 of the blessed things. Blimey!

Most of them are not reciprocal, as I had supposed. In fact, most of them are :

- duplicated by first listing a relationship (between A and B) and then listing it again separately as an "inverse" relationship (between B and A);

- then duplicated further by first listing a relationship as an "is" relationship and then listing it again separately as a "was" relationship.

This probably means you could divide 792 by 4 if you made them reciprocal and dated them. I haven't checked yet to see if each of the inverted relations is also represented by "is" and "was" but then you could reduce the number even further.

Even more disappointing, the relationships are simply enumerated (e.g. "was created by/created") with nothing, seemingly, to indicate how "create", for example, is defined and nothing much to tease out the many proposed meanings for creation and similar difficult concepts. Yes, it's a lot of work but that's what needs to be done.

2016, September 14:

<<Adrian Cunningham: … The following response/clarification has been developed with the help of my EGAD colleagues.

You are quite right, there still is a lot of work to do. That said, I would point to the editor's notes that explain some of the choices made (p. 39) and emphasises all the limitations of the current version of the list, which reflect the lack of time the group had to progress this work in the lead-up to the Seoul Congress-scheduled release. The draft that has been released is work in progress.

Of course there are far fewer actual relations than the list suggests. The list has been built in order to present a full list of the relations that have already been identified for each entity, thus it duplicates a lot of relations. The first column of the table thus does not provide an identifier for relations, simply a number. Moreover, the list does not yet cover all the discussions and work done about relations within EGAD. We regret any confusion that this might cause, but we felt it was better to get the current work in progress draft out to the professional community now so that people could see where we are heading and provide comments/ask questions/etc. There may well be better ways of presenting these relationships within RiC – and that is one of the things we would be glad to receive comments and suggestions about.

One view is that the next version of RiC should present the relations like the entities: by themselves, independently from any domain or range entity - which does not mean that domains and ranges should not appear. Before that, we need to discuss which relations are in fact the same (associated with for example is one unique relation ; others, like results in, should be considered one too) and which are generic (again, for instance, associated with), which are specific (for example: maybe we could introduce an Agent 'is family linked to' Agent as a generic relation, having as specific is parent of, is sibling of, etc.); thus design hierarchies of relations whenever they can be applied. In any case, this work must be done before building the ontology – RiC-O.

In the end, we will hopefully have a nice tidy list of relations, hierarchically arranged (general to specific) and probably presented in at least two ways: 1) an alphabetical list with definitions …; and 2) A to B, B to A, … as they now are. The latter makes the number seem much larger than it is. Nevertheless, quantity is not the issue; addressing what needs to be addressed is. And so we will have "as many relations as it takes to get the job done." Another factor that will be employed is to base the relations on rules. At the moment these are implicit, but need to be made explicit and formal. For example, "one and done relations," that is something that happened the is enduring, e.g., an Agent authored a Record; or recurring, e.g., Agent holds Record. Such a relation is transferrable, as we well know. As for past-future, well, we should have a discussion on this. There are implementation arguments for having past-present, namely to make it possible to have simple binary relations versus more complex n-ary relations.

Finally, taking up Evelyn McLellan's suggestion on the ICA-list that we set up a dedicated RiC/EGAD open listserve to facilitate discussions on the draft, EGAD has

established just such a dedicated listserv. Anyone who has in interest in this is encouraged to sign up. Please visit this page and subscribe: >>

2016, October 6: ICA's Records in Context – Conceptual Model

<<Cassie Findlay: … Comments are sought on the conceptual model by 31 December 2016 …>>

… The attachment to her posting gave the link to the EGAD Listserv and I would urge those interested to join it. I have made three postings and probably won't be making any more until it becomes more chatty. At least, I have tried to make three postings but I keep being blocked – so far only the first (on relationships) and its postscript have made it through. Here, for those may be interested, are my postings :

RELATIONSHIPS IN RiC

Someone has certainly been busy - 792 relationships and still counting. Phew! I read somewhere that a diligent German historian was only able to find 210 reasons for the fall of the Roman Empire. We certainly got that beat. This is a list of implementation options rather than a conceptual model – some of the logical possibilities when designing and implementing an application. To explore the full range of possibilities, two things are needed :

   1. the underlying relationship-types must be identified;
   2. the terms must be defined (cf. p.39) so that we all interpret the words in the same way.

Then we can pay more attention to refining or expanding those concepts that are currently being most contested (e.g. "create") and to discovering additional instances (e.g. "received by" under Transmission, "involved party" under Formation, "adopted (by)" under Existential Features, etc.). But it is more important to conceptualise than to itemise, therefore (by way of example):

One could begin with a thesis (inviting the antithesis) that provenance is to be found in Relationship-Type : Formation (see below). This could be tested by examining whether the 63 instances listed so far are, in fact, acceptable statements of provenance and whether any other ideas about provenance, of the kind that have been put forward lately in the literature, can fit within the instances listed or require additional instances to accommodate them. Is provenance only to be found within Formation? Are there formative relationships that are not allowable statements of provenance? Can provenance be found in other Relationship-Types? Does a formative relationship between Agents ("establish", for example) confer ambient provenance vicariously on a document-type? If so, how would that differ from "uses [agent-delegate]" which I have nominated as Existential? Alternatively, should ambience and provenance be kept conceptually separate? Does the Relationship-Type framework assist or hinder in (re)defining or (re)imagining our core concepts such as provenance.

I have trouble with two of the proposed entity-types (viz. Date and Place) of which more anon, so I can't yet come to terms with those proposed relationships involving one or other or both of those (204 out of the total). Interestingly, I singled these two out as problems long before I reached p.91 where Date and Place are also nominated as "properties" of relationships so maybe I'm not alone in needing to think some more about them. And I don't think it's worth dwelling long over the relationship-type "associated with" (292 out of the total). We've used that for years as a cop out for making links where we are too lazy or too uncertain to be specific. Anything can be associated with anything and, once you've said that, there's not much more to say and little benefit from saying it 292 times. Of the remainder, here is my first attempt at a categorisation into relationship-types (without the benefit of certainty as to what any of the terms mean):

- Relationship-Type : Formation (63 instances) viz. "create/created by"; "authored"; "collect(ed); "wrote/written"; results from/in"; "accumulate"; "assemble"; "arrange"; "establish".
- Relationship-Type : Governance (42 instances) viz. "owns/owned by"; "rights held"; "controls"; "directs"; "manages"; "superior/subordinate".
- Relationship-Type : Succession (22 instances) viz. "successor/predecessor"; "parent/child".
- Relationship-Type : Belonging (30 instances) viz. "part/part of"; "member of"; "is/has example".
- Relationship-Type : Possession (12 instances) viz. "held/holder".
- Relationship-Type : Transmission (4 instances) viz. "sent by".
- Relationship-Type : Documentary Features (73 instances) viz. "copy of"; "draft/original of"; "subject of"; "addressee"; "documentary form"; "evidence of".
- Relationship-Type : Existential Features (57 instances) viz. "has/had functional relation"; "assumed identity"; "sibling/spouse"; "uses [agent-delegate]"; "pursues/occupies [position or occupation]"; "fulfils [function]"; "performs [activity]: "authorize(d)"; "required competency"; "defined/revised [by mandate]".

There is, of course, much room for debate (e.g. is "authorize" an instance of the Governance or Existential type?). Nevertheless, I would find discussion at that level more rewarding than simply multiplying instances before something like that has been done.

Postscript

I have no problem with a long list that illustrates a concept. The RiC 1.0 list could easily stretch from 792 instances to 7,920 and beyond. Thinking up new instances could become a parlour game for archivists. My interest is in what principle(s) the instances illustrate. My suggested categorisation was derived from what is there is RiC 1.0 and is not what I would have come up with if I'd started with a blank page, so "something to live with" would indeed be most welcome. What I mean by implementation is that, w/o further explanation, one has to infer what the terms mean and how they might be used. Taking "creates", for example, and ignoring for the moment its diverse and often contested meanings (simply taking it as an unproblematic idea) it can be applied as a relationship thus:

> [ACTOR A]<creates>[RECORD X]

and this seems to be the how RiC 1.0 means it to be understood.

But all recordkeeping is based on describing action and circumstance and "creates" is an action which can, therefore, be rendered as a FUNCTION rather than a relationship (as well as, not instead of). The descriptive statement "A creates X" can then be rendered differently within the RiC 1.0 framework, where FUNCTION M = creates, as:

> [ACTOR A]<performs>[FUNCTION M]<to produce>[RECORD X].

It may be that somewhere in the list of possible relations in RiC 1.0 the option of using this second formulation is already provided for but, if so, only the diligent will find it and, absent more explanation, some of them may not understand that these are two allowable ways of achieving the same result. I agree, therefore, with those who have argued that it is important to draw out statements about how relationships are formed from the list of enumerated possibilities.

In the first formulation, according to RiC 1.0, Date & Place could be formulated as properties of a relation and also as instances of Entity-Types in their own right (instead of rather than as well as in any particular instance, I suppose). In the second formulation, it would be easy to link an instance of a Date-Entity and an instance of a Place-Entity to an instance of Function M. For those working a formed archive, the second formulation may seem

unnecessarily complex but those involved in active record-making may encounter '000s of create transactions every day and a developer might find it a more effective way of reaching the same outcome (viz. a statement to the effect that "A creates X"). Developers are clever people and could, no doubt, come up with lots more ways of achieving the same outcome for every rule, taking account of the differing needs of their client populations, so long as we provide them with a robust conceptual framework.

<div align="center">

ENTITIES IN RiC

</div>

Confusion between Recordkeeping Entities and Authority Records began with ISAAR. This seems an apposite moment to correct the misunderstanding. Four of the 14 proposed Entities (Documentary Form, Date, Place, Concept/Thing) could be represented as properties of the ten remaining. There is no harm in having those four as entities if that is useful (though the utility eludes me) and many more besides. In some metadata schemas, Relationships are nominated as entities, for example. But, if you're going to name four, you should make it clear that many other kinds of entity are possible and, if you're going to name those four, you should make it clear that they can (optionally) be treated as properties.

Alternatively, true Authority Records, like EAC-CPF and SNAC, could be built for Documentary Form, Date, Place, Concept/Thing, etc., etc. to control data content of the properties of Recordkeeping Entities. This leads on to the question whether we need to stipulate the properties of Authority Records used in recordkeeping. The other ten Entities proposed in RiC 1.0 are true Recordkeeping Entities whose properties can be controlled by Authority Records of one kind or another (or not, as the user decides). These ten entities can be conceptualised as instances (not the only possible ones) of three basic Entity-Types that are particular to recordkeeping :

- **DEEDS**: events or circumstances that give rise to recordkeeping – e.g. functions, functions (abstract), activities, mandates, processes, responsibilities, products, etc.;
- **DOERS**: actors who undertake the Deeds - e.g. agents, occupations, positions, corporations, agencies, processes, persons, families, etc.;
- **DOCUMENTS**: memories of Deeds undertaken - e.g. records, record components, record sets, series, *fonds*, documentary objects, processes, artefacts, legends, myths, etc.

I deliberately include "process" under all three types to illustrate the point that the same thing can be described in more than one way, using different Entity-Types as appropriate. I have already suggested the use of Relationship Types and I think using Entity Types is a better way also.

Four properties are common to all Recordkeeping Entity-Types in RiC 1.0 (Global Persistent Id, Local Id, Name, and General Note) and to those I would wish to add Date (either as a relationship or a property). Within the framework of an entity-relationship model, that would satisfy what I see as the mandatory requirements for all Recordkeeping Entities - viz. that they possess :

- **IDENTITY**: because every record is unique;
- **DATES**: because every record is time-bound;
- **RELATIONSHIPS**: because no record stands alone.

Other common properties, such as name, are useful but not essential in recordkeeping. If I were modelling RiC, I would represent the common properties as belonging to a Super-Type of the kind I have sometimes called the URO (Universal Recordkeeping Object), and more facetiously the HERO (Hurley's Enduring Recordkeeping Object). I think a good many more properties (e.g. Description) could be remodelled as common to all Recordkeeping Entities and brought into the URO either because they are already common to all Recordkeeping Entities in RiC or should be.

Other properties might be better handled in other ways, at least as alternative options. Some of these are trifling but "Accruals" (P24 & P25) should be given further thought. Accruals are part of a Process (viz. accessioning) and some people might want to document accessions as Record Sets (or Sub-Sets for incorporation into Sets). I would. That suggests that an option needs to be provided allowing accruals to be treated as Record Sub-Sets with relationships to Record Sets as part of the history of the formation of the Set and not merely as a property forecasting future possibilities. In the physical world, it was sometimes necessary to manage Transfers or Deposits as entities (Record Sub-Sets) separately from Accessions because they comprised one of more Accessions, formed before, during, or after relocation, and I imagine that similar entities might be useful during data migrations.

2017, January 17: RiC – Quo Vadis?

Note: When someone comes to write a history of RiC, they'll have trouble finding material using embedded links in various ICA and scholarly presentations online, at any rate I do. So, apologies, but a lot of the links in what follows lead nowhere.

Just before the deadline for comment closes on RiC, here are some clumsy existential questions. They're not just questions for RiC, of course, or what EGAD's next steps should be – but they may apply to the future direction of description overall (RiC or no RiC) – at least I hope it may be so. Fingers crossed that the formatting comes out. In case it doesn't, I append the diagram as an attachment.

Visual Diagram of How the Records in Context Conceptual Model works
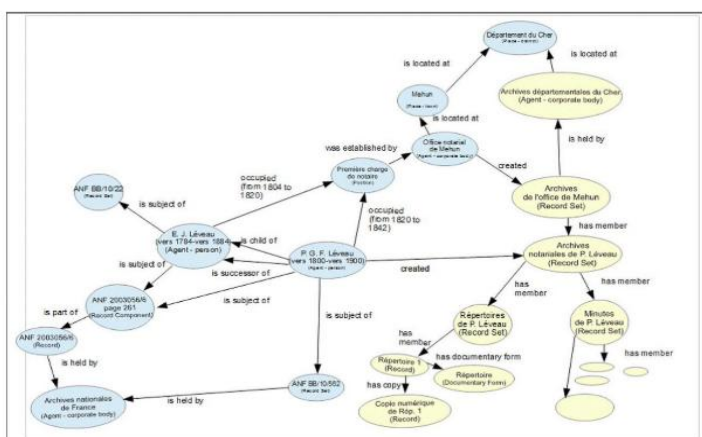
"This example shows some information about P.G.F. Leveau a French public notary in the 19th century including:

Query 1 (Structure):

Can we define an Entity/Relationship type as one containing instances that all operate according to the same recordkeeping requirements (allowing for extensions by sub-types that augment but do not conflict with the common requirements)? Can they be managed, in other words, using identical rules or practices (with extensions) that are set at the level of

48

each type rather than each instance? I once theorised that an ownership relationship is a succession relationship in disguise – easily demonstrated (see below), but not so easily proven. Can we separate conceptualisation and implementation so that a proliferation of instances within each type would not matter. You could have 8, or 800, or 8000 instances of any type; the same standardised practice would govern all. Implementers could then select those instances that are useful to them, ignore the rest, and then apply the rules (or not) as appropriate. Could that approach be taken within an infrastructure (policies, procedures, roles, etc.) that is not particular to any one descriptive programme, jurisdiction, or prejudice?

Query 2 (Identity):

How should we think of the nexus between the description of an entity/relationship and the entity/relationship itself? How does an instance-in-action being described differ from the description of it? Is description simply a parallel universe, laying down a descriptive world alongside an actual world? Does a recordkeeping (descriptive) system operate in a descriptive universe or in an actual universe or does it straddle the two? Where does our understanding of a corporation, for example, "exist" – in the actual world or within a descriptive (registration) system, or both? Can a description of an instance-in-action in the actual world (physical or virtual) be turned into a kind of avatar so that it can operate in a recordkeeping system as if it were the thing itself, not just a description of it? What is the difference (if any) between action in the virtual world and action in the physical world? How can two different descriptions of the same instance (in the descriptive world) be reconciled? Is there ever a case of a graphical representation for which no extant personality or actuality exists?

Query 3 (Validation):

How can authenticity be conferred on descriptions that operate outside of the source or native system? Could they be trust-worthily registered or validated using PKI and/or blockchain? What kind of recordkeeping system would be needed to validate them (viz. descriptions of description) and could that system be a source for persistent identification? To what extent would that require re-contextualisation? I once asked my friend Terry Cook when he was in full flight about top-down appraisal: How do you know when you're at the top? Reminds me of a great story I once heard about Hilary Jenkinson when he was interviewing a nervous young Oxbridge graduate for a job. Asked what had been his special field of study, the youngster replied, "The end of the 17th century, sir". Jenkinson growled, "Which end?" An archivist's question.

PS. Demonstration of a succession relationship disguised as an ownership relationship:

Consider a simple succession relationship:

AGENCY B------<**succeeds**>-----AGENCY A

Now, consider two ownership relationships:

FUNCTION G	FUNCTION G
<**exercised/owned by**>	<**exercised/owned by**>
(**from 1901-1925**)	(**from 1925-1980**)
AGENCY A	AGENCY B

The ownership relationships can be described in a table:

FUNCTION G **exercised by** ….	Dates
AGENCY A	**1901-1925**
AGENCY B	**1925-1980**
AGENCY C	1980-1995
AGENCY D	1995-ct
etc.	-

The ownership data captured in the Table is sufficient, without any need for further data input or description in the form of a succession statement, to generate a succession relationship:

> AGENCY B--------<**1925: succeeds in exercise of FUNCTION G>**-----AGENCY A

Not only has an ownership relationship metamorphosed into a succession relationship, there is added value from depicting how and when the succession arises. The data table can, in fact, generate the following descriptive statements:

> AGENCY B--<**succeeds>**--AGENCY A in 1925 in exercise of FUNCTION G
> AGENCY A--<**succeeded by>**---AGENCY B in 1925 in exercise of FUNCTION G
> AGENCY A--<**exercised/owns>**FUNCTION G from 1901 to 1925
> AGENCY B--<**exercised/owns>**---FUNCTION G from 1925 to 1980
> FUNCTION G--<**was exercised/owned by>**---AGENCY A from 1901 to 1925
> FUNCTION G--<**was exercised/owned by>**---AGENCY B from 1925 to 1980

PPS. The query "Does a recordkeeping (descriptive) system operate in a descriptive universe or in an actual universe or does it straddle the two? " was posed way back in the **SPIRT Project** (Business Recordkeeping entity class posited as a sub-set of the Business entity class). I never thought the answer was entirely satisfactory, but it was the right question to ask.

2017, February 1: Submissions to EGAD's Records in Context

From Barbara Reed to EGAD –

> The submission period for comments on the Records in Contexts document – a really important restatement about archival description – finished yesterday. There hasn't been much discussion, despite Chris's best efforts and the material forwarded by Lise.

> In the spirit of fostering discussion, Chris has encouraged me to put this on the discussion list. So here is what I responded to the RiC in a personal capacity (minus the attachment) and all last minute as is my wont:

Comments on EGAD - Expert Group on Archival Description, Records in Context - Conceptual Model http://www.ica.org/en/call-comments-release-records-contexts-egad

Introduction

> Thank you for the opportunity to comment on the evolving RIC (Records in Contexts). I commend the work of the EGAD group in this complex and demanding work. Exposure to the archival community for comment is appreciated. The introduction of a multi-entity relational model which enables recursive relationships within entities, and extensive relationships between entities is totally supported.

> However, as indicated in more detailed comments below, the definition of the entities is problematic, as is the management of relationships. Relationships and their management become critical in this type of model. Modelling relationships has always been difficult and problematic. Other disciplines do not seem to have the same requirements for persistence and management of relationships over time. Indeed it can be said to be one of the key features of the Australian Series System, an intellectual basis from which my practice evolved. As a community I would assert that we haven't cracked the expression of relationships yet, and neither has the RIC, with problems in the models proposed here.

> It is with great pleasure that I read this document, and applaud its aspirational stance. The networked, flexible model for archival description at the basis of RiC-CM will serve the archives profession well into the digital future. The alignment with recordkeeping metadata approaches, which can be seen in the multi-entity and

relationship definition, will serve the broader recordkeeping community well. Compatible models for records regardless of the domain they are managed in (current workplaces, archives or in the 'wild') will enable much greater interoperability, inheritance opportunities and enhancement rather than replacement of metadata sourced from various processes over time.

Qualification for comments

I have had some involvement in definition of metadata for archives/records through work on the original SPIRT metadata project at Monash University, the development of the ISO Metadata for Records standard suite (ISO 23081) and the AS 5478 Australian Metadata Reference Set. I chaired the Australian Society of Archivists Archival Descriptive Committee for a time, and contributed to the codification of the Australian Series System. I have also had development and implementation experience with jurisdictional metadata standards such as those published by National Archives of Australia, Archives NZ, HK SARs Government Records Service.

Comments on Entities

Archives and records work deal fundamentally with three types of entities in relationship. This has been the bedrock of Australian archival practice for over 50 years. These entities are Records, Agents and Functions. These are the entities that we are professionally responsible for. Other entities introduced into the RIC may well be useful for description (eg Concept/Thing) in the world of semantic web construction. But they are generic and not our core business. While supportive of their inclusion and their potential to link to other professional and cultural domains, the RiC-CM should prioritise those entities which are our core business.

I commend the introduction of limited set aggregations. The inclusion of multiple recursive relationships for Records Set/Agents in particular opens up the data model to be responsive to multiple archival descriptive traditions (although the examples used in the property descriptions could more actively embrace this by using not only ISAD G representations, but also explicitly acknowledging that these can be inherited/used to support other descriptive traditions). The recursive nature also increases the relational power of the model, something endorsed totally.

In the articulation of the entities provided, there seems to be a confusion between conceptually defining our core professional entities (which doesn't exclude inheriting the expression from others), and the practical demands of constructing system entities or more general descriptive entities. I cannot find the conceptual rationale or logic of separating out some of the entities, rather than making them properties. At minimum an articulation of the logic is required. But I would argue:

- Occupation, Position are attributes of Agents.
- Documentary Form is an attribute of Record Set/Record
- Date and its expression is an attribute of every entity, at every layer of aggregation, and also an attribute of relationships which need to be timebound. An alternative construct is to move further towards making relationships central. If every action is expressed as a dated relationship (for example existence, extent, actions on a record) then all critical statements are made as relationship statements which are dated. This might mean that date is no longer needed as an entity, but becomes critical in expression of relationship – some of this can actually be seen in the graph diagram included in Appendix 1.
- Relationship must be an entity if we are to coherently express and manage relationships over time (see below).

- Place and Concept/Thing is a subject based attribute that can be added to anything and while it links us to the Linked Open Data community and is desirable, it is a set of add ons, nice to have, not critical to our archival practice. In that spirit of great, but not exclusively recordkeeping relationships, I would think that Events could also be added, but recognising a potential overlap with what is currently inadequately expressed as 'history' needs distinguishing from recordkeeping events (see further below).
- Place is confusing – it seems to be both location (has holding location etc), as well as physical positioning, such as geographical coordinates. I suspect these two quite different notions of Place is made to enable inclusion of ISDIAH, which itself was always out of step with the ICA descriptive entities. Archival repositories are a type of agent, as has been argued before.

<div align="center">

Relationships

</div>

The relationship based approach is central to our Australian descriptive practice and I support it wholeheartedly. However, it has also been notoriously difficult to achieve an adequate expression of relationships. In traditional Australian practice relationships have been a component of each entity's description – so a record description would include its relational link to its creating agency, controlling agency, related records etc. Using these relationship traces makes the tracing of relationship networks flexible, time bound and complex.

However we recognise that this de-emphasises the nature of relationship. In other work, we have expressed relationship as an entity itself, allowing relationships to have identifiers, dates and to contain persistent links to the related entities. This has been OK but subject to problems even at a conceptual level in expressing reciprocal relationships and ensuring persistence of relationships. It is the best we have done to date.

In RIC, the adoption of semantic data models, and graph technology construction is highly commendable and I endorse the notion that these are likely to be the technologies of the future. Relationships are central to these technologies and the approach is completely consistent with our Australian archival practice. However it is unclear to me how the technology or the data models document these relationships.

While I do not pretend to understand graph databases, the graph models such as the one attached here indicate that relationships can be managed as entities effectively in graph databases. Surely modelling relationships in this way is more sustainable for archival description. The relationship notion is central, not peripheral.
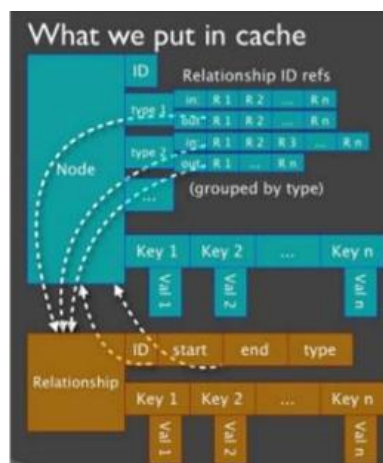
Managing the relationships which are critical as only statements in RDF triples only provides persistence to the nodes that are linked, rather than attributing persistence to the relationships themselves. While the EGAD committee acknowledge that the expression of relationships is still a work in progress, I would suggest that confirming the data model around relationships is essential.

The extensive list of relationships provided is acknowledged as not comprehensive and in need of further work. Given that it is a central component of the RIC model, further development must take place before the model can be endorsed.

In Australian practice we have traditionally identified types of relationship. These are Provenance, Succession, Containment and Associative relationships. The recordkeeping metadata standard introduced a further relationship type of Events or Actions which allowed description of things done to or on records as a further relationship type (see below). Using this type of characterisation of relationships might assist in creating more clarity about what types of relationships are appropriate to each entity. Chris Hurley has proposed a further categorisation of relationships. Developing this thinking would be beneficial in RiC-CM.

Recordkeeping Events

Noting that the introduction to the draft RIC stated 'RiC-CM also does not yet offer a model of the role of the archivist and the activities he or she performs in the formulation and ongoing maintenance of description' this points to a conceptual gap in the articulation at present. The property 'Authenticity and Integrity' in Records Set/Record (RIC P5 and P22) inadequately encompass the requirements to enable assertions of authenticity and integrity. I would argue that the material not well encompassed in 'History' documents the events and actions that are taken on a record and this is the information needed to assert authenticity and integrity over time. For digital records integrity is also the digital checksum or hash of the specific record element. I find the conceptual thinking to be unclear about this.

Within our Australian practice we have identified 'recordkeeping events' to document these actions. This may, or may not, be the answer, it does propose a different way of thinking about these actions. Further, such recordkeeping events are expressed as relationships – themselves an expression of something done, by someone, on something at a particular date. Some of the relationships identified in the listing are of this nature (eg was written by, was collected by). This could be extended and, depending on the nature of relationships as they evolve, might prove a mechanism to clarify notions of authenticity and actions.

Particularly when encouraging inheritance of metadata from current recordkeeping systems, attention to recordkeeping events for digital records particularly, is essential.

While not proposed as necessarily authoritative, the work done in AS 5478 Australian Metadata Reference Set for event relationships may be useful. This is attached to this comment for reference.

Specific comments

Granularity: One of the things we know about translating archival descriptive systems to digital records is that increased layers of granularity of description are required. Thus what may be expressed as a 'Record' in the paper world, may in fact be composed of many digital components. Thinking of a file, traditionally considered an item in archival descriptive systems (a 'Record' in RiC-CM) because it is a complete, and 'issuable' thing, can disaggregate into many many specific images of individual pages, each one of which can be considered an 'issuable' thing. Is it the intention to manage images of an individual page as a 'Records component'? I suspect

that this will not work particularly well, and certainly considerations of sequence must be addressed (noting that this too is explicitly noted as requiring further development p39). Alternative renditions or formats (microfilm, pdf, jpeg etc) may also exist for each page. Perhaps a better expression is to allow 'Record' too to become recursive.

Parallel provenance or multiple simultaneous provenance: Increasingly archival practice is allowing alternative models of description to co-exist with 'official' interpretations. This allows alternative versions of context to be constructed, and to have equal validity with 'official' expressions. How will RiC-CM enable these alternative expressions?

User contributions: Linked to the need to enable multiple provenance expressions but not the same, is the increasing prevalence of user contribution to archival descriptive systems. This might be through alternative expressions in the metadata of an archival descriptive system, or alternatively, the contribution of other items to an archival system. These need to be managed and attributed appropriately to the contributor. At present such notions do not appear to be enabled in RiC-CM and they are already a requirement of practice.

Digital records: presumably as a requirement to encompass the existing ICA descriptive standards, the examples and properties fields seem to reflect a paper based paradigm. Better examples and further thinking on characteristics of digital records would enhance the RiC-CM model.

I welcome the advent of RiC-CM and encourage the EGAD Group to continue development of this exciting initiative.

2017, August 14: New to my website

Records in Context (RiC) 1.0 – Comments on First Draft (2016)

In September 2016, the ICA Experts Group on Archival Description (EGAD) released Records in Context: A Conceptual Model for Archival Description – consultation draft v01. The draft was open for comment until early 2017. A listserv was set up and I made four contributions, reproduced here-

- Relationships in RiC (1) 17 Sep 2016
- Relationships in RiC (2) 24 Sep 2016
- Entities in RiC 11 Oct 2016
- RiC: Quo Vadis? 30 Jan 2017.

RiC is described as a "comprehensive descriptive standard that reconciles, integrates, and builds on the four existing [ICA] standards". It is a puzzling document. The authors state that it covers - "…all of the essential content of the four existing ICA description standards, except "control." It thus includes the core descriptive entities, the properties or attributes of these entities, and essential relations among them. Further specifying and defining the relations among the entities remains an outstanding task, as the relations in the current draft are intended to be suggestive and not complete or normative. RiC-CM also does not yet offer a model of the role of the archivist and the activities he or she performs in the formulation and ongoing maintenance of description. EGAD will extend the model in this respect as a next step…"

This draft is not, therefore, a conceptual model at all but a set of elements that might be utilised in any one of the many possible conceptual frameworks that could answer the fundamental questions: viz. what are we trying to describe, why, and how are we intending to go about it. The purposes of description are said to be management, preservation, and use but we don't yet know how these entities and relationships are intended to be used to support

those purposes. That will all be added in later (see p.2). If the present draft is "suggestive and not complete", as stated, we need to guess what it is suggestive of –

- augmentation of the descriptive philosophy originally underlying the ISAD suite ("fonds-down") or
- its repudiation and replacement by something else ("multi-dimensional"; entity/relationship) or
- what is more likely, continuation of conceptual ambiguity pragmatically supporting either approach?

But surely it is not too early to state whether or not it will be normative. Time will tell, I suppose.

Maybe we have been told. They say that "efforts to realize ... integrated access [that] have focused on developing a shared standard [by] reducing the different descriptive practices to one is intellectually and politically challenging." You bet! But it is not acknowledged that shared standards can operate (differently) at different strata – from the highest level of principle and concept down to the most basic level of implementation. David Bearman taught us many years ago that diversity of method can operate within a harmonisation of functional requirements. They continue "[t]his objective though, does not require such a reduction, as the communities need only identify and cooperate where there are shared (or largely shared) concepts and practices." Hmmmm. So, what is to be standardised: what we agree on or what is important? This anti-reductionist formulation begs the question. If we don't need to standardise what we can't agree about, why do we need to standardise anything at all? I could understand "we only need to agree about and standardise what's important" or "we must agree the requirements, not the methods" but that is not what has been said. What then do we agree upon? What are our shared (or largely shared) concepts and practices? Do we even know what's important and what's not?

The draft speaks of "established" archival principles and practices but we don't have a statement of them. That which is claimed, without specifics, to be established, agreed upon, or acknowledged seldom is in my experience. In the early 1990s a Statement of Principles upon which the first draft of ISAD(G) was based drew forth such ferocious dissent in Beijing that I was added to the drafting Committee to argue for alterations to it that might meet some of those objections. But that document was shelved at the very first meeting I attended in Stockholm on the grounds that it was now "an historic document" – pretty obviously a ploy to avoid further discussion of concepts and practices – and I found myself the odd-man-out in hassles over drafting that could never be resolved because there was no common ground. That was good politics but bad policy. The uncorrected Statement could not be appealed to as an authority for "established" concepts and practices nor used as a point of reference for further development of the standards, even though that is how it was regarded by everyone in the room but me. I have since argued that we would be in better shape today if that Statement of Principles had been dealt with properly before rushing ahead to develop the four standards that came to be in the ISAD suite. I hope history is not about to repeat itself.

Note: I have recovered my Comments on RiC 1.0. They are [appended](#).

2017, September 17: RiC at Riga

New on my website.

RiC at Riga – ICA-SUV Conference held at Riga 21-24 August 2017

I attended this three-day conference on Cultural Heritage Materials – University, Research & Folklore Archives in the 21st Century. On the last day, there were two sessions that paid special attention to RiC. I was asked to chair one and Stefano Vitali the other. We were each

asked to make some remarks of our own and these are mine, worked up from notes I used on the day.

> **Ric at Riga**
> Drawn from Notes for Remarks at the Closing Session ICA-SUV Conference
> Cultural Heritage Materials – University, Research & Folklore Archives in the 21st Century
> Riga, 21-24 August 2017
>
> RiC is a conceptual model in search of a concept. But it's early days yet; there's time to get it right.
>
> Many speaking at this Conference have made the point that the ISAD suite poorly supports their part of the "archival multiverse" referred to by Maryna Chernyavska, including (inter alia) archiving of traditional knowledge, folklore archiving, community archives, oral traditions, indigenous knowledge, audio-visual materials – and, I might add, geospatial archiving, digital archiving, and the archiving of maps/plans and images. Some of those speakers have welcomed RiC because of its "flexibility". The argument, as I understand it, is that too many assumptions that limit its flexibility were built into ISAD(G) about the kinds of material to be described, the methods to be used, and the purpose description is meant to serve. RiC's flexibility, these speakers have argued, will permit them to better adapt it to their special needs.
>
> But flexibility is not the same as lack of direction. It is sometimes forgotten that ISAD(G) was meant to be the first in a series: ISAD(H), (I), (J), (K), and so on, I suppose. At least, that is how it was explained to me when I joined the drafting Commission in 1993:
>> The areas of description covered by these general rules are those thought to have the widest applicability in an international archival context. This is only the beginning of a standardization effort. Further specific rules should be formulated to guide the description of special categories of material (such as cartographic materials, motion pictures, electronic files, or charters, notarial deeds, property titles).[1]

2019, December 13: Press Release ICA-EGAD

<<Adrian Cunningham: FWD … Public release of Records in Contexts - Ontology v0.1 and Records in Contexts – Conceptual Model v0.2 preview … The released version, v0.1, is the first public release of RiC-O. It is compliant with RiC-CM v0.2 preview, released by EGAD at the same date …
The RiC-O URI is: https://www.ica.org/standards/RiC/ontology
Access the HTML view (documentation) of RiC-O v0.1:
     https://www.ica.org/standards/RiC/RiC-O_v0-1.html
Access the OWL/RDF ontology file: https://www.ica.org/standards/RiC/RiC-O_v0-1.rdf
EGAD calls for comments on RiC-O v0.1>>

Here's the comment I've made on RiC 0.3

Since I'm now in my dotage, I no longer have the attention span to do this thing justice. I'll limit myself, therefore, to a few cursory comments (based on an equally cursory reading). If I get it wrong please chastise me. The definition of THING is bewildering.

E01 THING	E01 THING cont'd	E01 THING cont'd	E01 THING cont'd
--:E02 Record Resource	--:E07 Agent	--: --: E12 Position	--:E18 Date
--: --: E03 Record Set	--: --: E08 Person	--: --: E13 Mechanism	--: --: E19 Single date
--: --: E04 Record	--: --: E09 Group	--:E14 Event	--: --: E20 Date range
--: --: E05 Record Part	--: --: --: E10 Family	--: --: E15 Activity	--: --: E21 Date Set
--:E06 Instantiation	--: --: --: E11 Corporate Body	--:E16 Rule	--:E22 Place
		--: --: E17 Mandate	

Is it
   a) "all possible concepts, material things, or events within the realm of shared human experience and discourse" OR

b) something exclusive "of entities that are not explicitly identified and described in RiC [which] are commonly the responsibility of allied cultural heritage communities, academic and research communities, or specialized or expert communities." OR

c) "all of the entities that are of primary interest to records managers and archivists, as well as other entities **used** (my emphasis) in the description of the primary interest entities [including] all other possible entities that are not explicitly identified in RiC as entities [that is] the subject of a Record Resource, or associated with an Activity"?

Why I think this is important will become clearer below (I hope).

a) **All possible concepts, things, or events** … i.e. a universality, everything, whatever is or might be. Like the Universe, it is meaningless to ask what lies beyond. If anything lies beyond, it is (by definition) part of the Universe and therefore it does not lie beyond. It follows that there'll only ever be one THING and that it will be everyTHING. It is like Le Grand K, one of the seven fundamental measures, the ultimate recordkeeping datum, a reference point for everything descriptive - until it is redefined in RiC 0.3 of course. Not unlike what I once described as the BIG 1 – a universal ambient entity conferring context on everything else. If BIG 1 were ever actually implemented it would, naturally, annihilate parallel provenance.

b) **Cosa Nostra (our THING) but not Theirs**. Entities which are (or could be) described in RiC but not those which are the responsibility of "allied" communities. To say nothing, I suppose, of entities for which non-allied communities are responsible. This limits the concept considerably to a boundary that comprehends recordkeeping purposes but no others. Our THING exists in a world of many THINGS but the others are of no use or interest to us. The question remains how we identify our THING. Once, when Terry Cook was expounding top-down-appraisal, I asked "How do you know when you're at the top?" He didn't answer but he thought it was a good question. Still is. (Query: Perhaps we are meant understand that our THING is one of many *hypostases* consubstantial in one universal entity?)

c) **That THING we do:** The Cosa Nostra might be understood as a pure concept, a Platonic Ideal – something that encapsulates "the true and essential nature of things, in a way that the physical form cannot". Under interpretation (c), the THING is an applied concept, one which is to be "used in the description of the primary interest entities". An entity definition that is to be employed in any application such as AtoM. An architectural definition of actual descriptions used in portraying THINGS that are observed. A utilitarian device derived from Aristotelian Empiricism – with each real-world implementation of the RiC THING being another instantiation of it.

It's still a hierarchy of sorts harking back to the dreadful Multi-Level Rule, perhaps to reassure the unenlightened. The entity-types are represented as "levels". Goodness me, aren't we past that yet? The hierarchy is starting to break apart (Person is not the child of Family thankfully) but it all needs to go much further, Otherwise structuration built into the conceptualisation queers the Relationships. It's axiomatic. If you place one thing in subordination to another then you are forming a relationship in advance. Entity-Types need to be conceptually autonomous, not bound to each other in any defined way – e.g. Document, Deed, and Doer. They need to be connected purely by means of Relationships - contingently. Binding them conceptually or ontologically usurps the work of Relationships. The recordkeeping character of any Entity-Type derives not from how it is defined but from how it is used.

Some of the Entity-Type definitions give rise to similar problems – e.g. E 18 Role. Role definition is usually (not always) best expressed as a Relationship. A loan, for example, involves a lender, a borrower, and a guarantor which is a situation best understood as a set

of Relationships subsisting between three Doers, not as characteristics of the Entities involved (i.e. their respective Roles). The author, the recipient, and the custodian of a letter (or chain of correspondence) are best represented as Doers standing in three different kinds of Relationship to the same Document.

I won't comment on the list of Relationship-Types. There's fewer of them and that's good and, besides, I've already had my say. It's a start but we've still got a long way to go. It's the conceptual imprecision around THING (see above) that bothers me most. Most of the Relationship-Types in this model will operate OK (sort of) within the boundary of a THING, but the model depicts all other Entity-Types already standing in a definitional relationship with the THING and this confuses matters. And, unless the THING is a universality - option (a) – trying to craft a Relationship with an Entity outside the boundary of the THING will run foul of the predetermined Relationships already hard-wired into the hierarchy. There are two reasons you don't want that to happen –

- It cripples your ability to deal with parallel provenance, and
- It limits your scope to provide for scaleability.
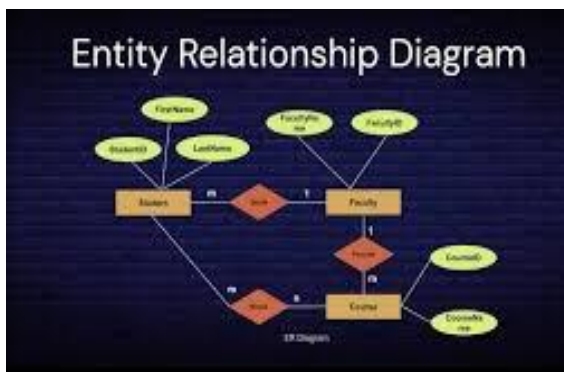
Scaleability allows you to show a Series, for example, as an Ambient entity in one kind of relationship (to a letter in a file, for example, where the letter's Author - its Creator - is shown as its Provenance and the Author is shown within another Ambience altogether while the Series is simultaneously represented as a creation of the Doer to whose fonds it belongs. In other words, the Role of any Entity depends upon the Relationships crafted for it with other Entities and instances of the same Entity-Type can perform almost any role that can be depicted.

2022, October 8: RiC 0.2

Those not on the ICA List who have 45 mins to spare should check out the video recently posted there by EGAD. It outlines the latest stage in the development of the international descriptive standards. https://www.youtube.com/watch?v=oHG_pupre8w

They are following the pathway indicated in earlier iterations but this time they are setting out more clearly (I think) and at long last a clear conceptual understanding, including –

- **Status Quo [from which they are departing]**: ISAD/ISAAR/ISDF are mono-hierarchical, unintegrated, and outdated [all bad things]
- **Core Concepts**: Entities, Attributes, Relationships + Extensibility, also Scaleability, but they don't call it that [all good things]
- Distinguishes between Physical and Intellectual characteristics
- **Relationships**: Down in number to 78 [!] but now "categorised" into 13 groups. [The analysis of relationships is still polluted by alignment of some relationship-types with entity-types but they're learning]

They seem (to me) to be well on the way to coherence and flexibility (two good things). Conceptually, this is where we should have been twenty-five years ago. Alas, that only gets

us up to the starting gate. As EGAD seem to be well aware. Issues of implementation and integration of archival description with the larger digital universe are yet to be solved.

RiC still has the smack of custodialism about it (probably unavoidable at this stage). As it develops, the standard may work for established archival institutions (I hope it does) but it also has to provide for recordkeeping outside of the archives (or, to put it another way, for the time when archives becomes part of the recordkeeping process, if it ever does). And, what's more, I can't see much joy here for the barefoot archivist.

Still, a lot to like in this.

❧❧❧❧❧❧❧

2017, June 14: ATTN AtoM users – help wanted

On the Artefactual demo site, the instruction for Description Id is (from ISAD) to –

> Provide a specific local reference code, control number, or other unique identifier. The country and repository code will be automatically added from the linked repository record to form a full reference code.

The instruction for the Archival Institution Id on the "linked repository record" is (from ISDIAH) to –

> Record the numeric or alpha-numeric code identifying the institution in accordance with the relevant international and national standards.

Can someone who is using AtoM please advise which "relevant" standards Australian AtoM users are applying to formulate country and repository codes in their Repository Identifiers? Specifically, have agreements been reached amongst you, of which we haven't heard, regarding –

1. The source or standard Australian AtoM users are employing for country code? There are numerous standards from which country codes can be derived. The country code "AU" appears to be in use. Have Australian archival/recordkeeping programmes agreed on which one to use and if so which one? Is there an international agreement amongst archives/recordkeeping programmes as to the source of country codes to be used? Is ICA doing this?

2. The standard or register of Australian repository/programme codes being employed? Repository codes are being used (e.g. "NBAC', "WA", etc.) but where do they come from? Is there a national agreement amongst Australian archives/recordkeeping programmes on a system for assigning repository codes? Who maintains the register, manages changes, and assigns codes to new participants? To whom do we apply for ours?

2017, June 26:

<<**Joanna Sassoon: The silence is resounding on your question, but it is one that students doing archival description units raise every semester. Is this yet another sign of the lack of an Australian archival system? …>>**

After the silence had resounded for a while, we posted the same question about country and repository codes to the AtoM User Group and received the following replies:

> From Dan Gillean (AtoM Program Manager)
> In AtoM, the country code is automatically assigned…As for repository codes…Ultimately this varies…Sometimes there is a national body providing control, or an aggregator (like Trove) takes on the role; sometimes repository codes are self-assigned.

We're testing AtoM at the moment for item level control and it doesn't seem to quite work in the way Dan describes. I have just created a test repository description and populated the contact area with Australia from the pick list as directed. "AU" was <u>not</u> then automatically assigned in the Id area as expected. A devil then prompted me to split an infinitive and manually assign a country code "NZ" and it was accepted.

> From Maggie Shapley (ANU)
> The repository code is up to you here in Australia. It's important to note that it will appear between the country code AU and your series numbers so you need something fairly obvious and not too long. We went with NBAC for Noel Butlin Archives Centre which is followed by various alphanumeric and numeric identifiers for those collections. For the University Archives we use ANUA (Australian National University Archives) as that was already the series prefix so that's now followed by the single number that is the series number. I guess you'd want to avoid anything that might get you confused with any other archives in Australia. To set it up all you do is enter it as the Identifier in the repository record.

It appears, then, that there is no national registration control and you just make it up. If we proceed, we'll try to grab CBA before we lose it to the Centre for Backyard Astrophysics.

2018, January 10: AtoM Repositories
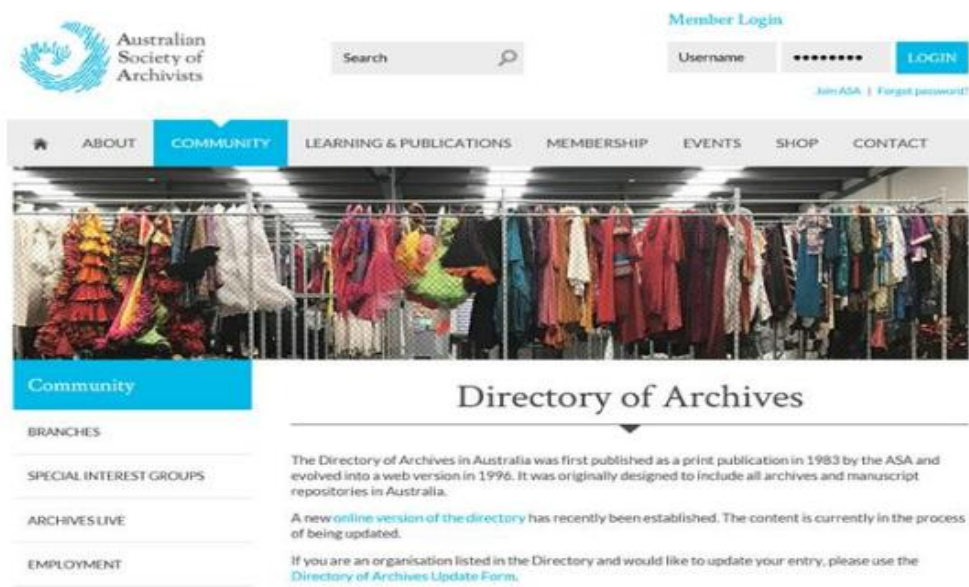
Some months ago, I asked if AtoM users in Australia had agreed a registration system for repository codes. The answer, I believe, was "No, we're individually making them up." Now, the ASA is updating its (once defunct?) Directory. The link on the ASA site takes you to an AtoM record which I assume ASA is using as the platform for the updated Directory (there's nothing said but that is what I surmise). The entry for Commonwealth Bank is based on information submitted some years ago and has an "identifier" (daa/62) which could

conveniently serve as an AtoM repository code if Australian AtoM users agreed amongst themselves to make use of the Directory in this way. Might a poor warmint make so bold as to ask if Australian AtoM users have given any consideration to this possibility (may already have decided to do so)? After many years condemn'd to hope's delusive mine in the field of standardisation, my mind recoils from even contemplating international co-operation in the matter of repository registration. You will observe that I am using my semi-retirement to reacquaint myself, inter alia, with Dickens and Johnson.
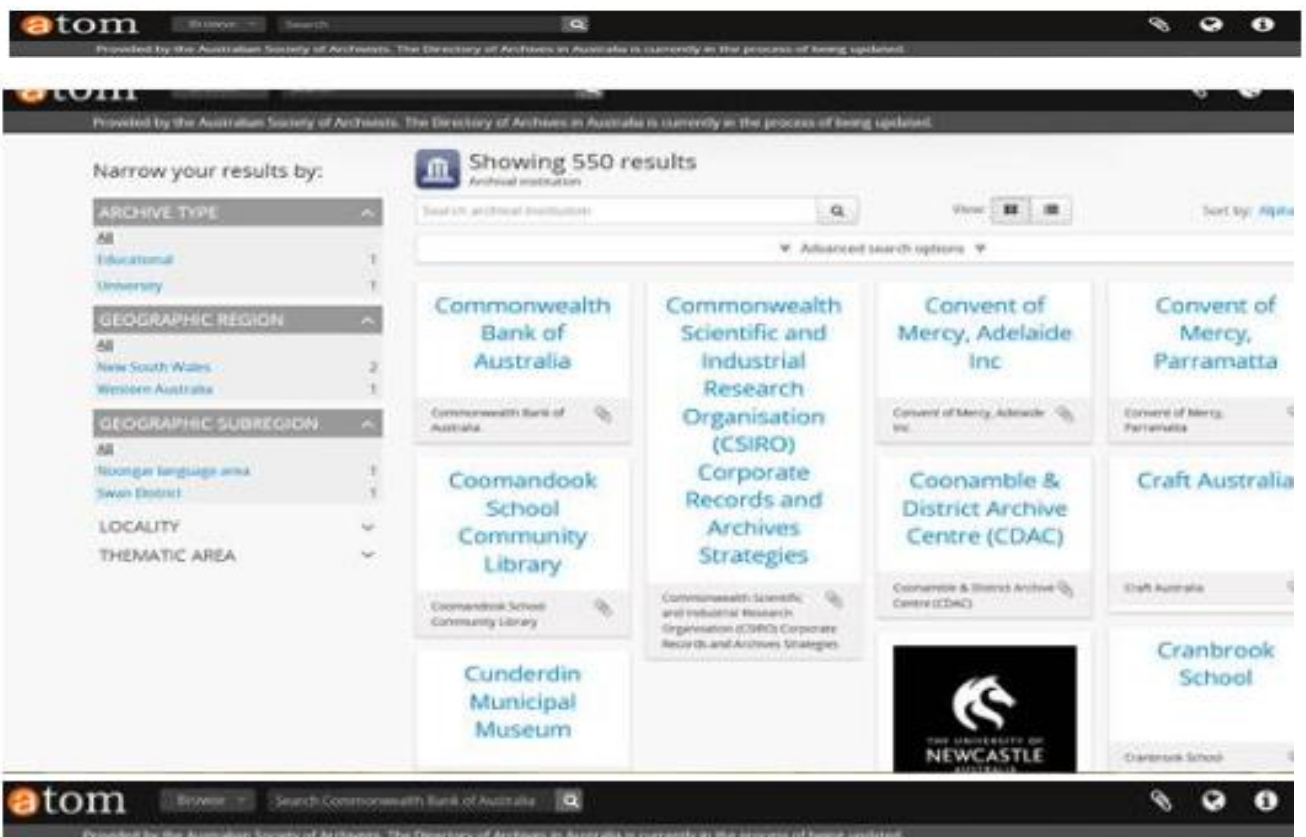
2018, January 13:

<<Lise Summers:... With respect to the repository identifier, how it is used is really a question for all archives, not just those using AtoM. The repository identifier comes from ISDIAH, which states institutions with archival holdings should use a nationally

recognised identifier, if one exists. So, if we all agreed to use the identifier, it could be used in any system, including TROVE, or ANDS, or Wkipedia, where the standard might apply. I think that the code, its implications and use, could be part of a broader discussion ...>>

2018, January 14:

So, the answers to my questions would appear to be:

1. Q. Might a poor warmint make so bold as to ask if Australian AtoM users have given any consideration to this possibility (may already have decided to do so)?
   A. Nope. Not yet at any rate (though there is nothing stopping them/us from doing so).

2. Q. After many years condemn'd to hope's delusive mine in the field of standardisation, my mind recoils from even contemplating international co-operation in the matter of repository registration.
   A. And it still recoils from doing do. The form of the ASA's Directory identifier code "if we all agreed to use [it]" would conform to internationally approved practice. But we haven't agreed (yet). If we did, it would still only be a national registration system based on international rules.

The biggest issue that I see is the need for change management rules. The work in progress is now picking up old data that is (in some cases) long out of date. For example, CBA Archives needs an identifier to show the location of Trust Bank records it has on permanent loan with TAHO, but there is no entry for TAHO (that I could find), only an old one for AOT. Who has (or will have) permission to add, modify, and delete entries? What protocols will there be? Will superseded entries be maintained so that programmes that have used superseded identifiers won't be cursed with broken links? Will there be an editorial/review process to iron out duplication, ambiguity, error, etc. from contributions made to the Directory?

2018, January 16:

<<Katie Bird: ASA Council ... have recently embarked on a project to update it ...and I would encourage anyone who represents an archival institution currently listed in the Directory to update their entry via this process. This page also includes a set of instructions on mandatory fields and other Directory editorial information.>>

2018, January 18:

Leaving aside the AtoM tie-in, my questions are about management protocols for the Directory (as a possible source for AtoM repository codes). The biggest issue that I see is the need for change management rules. The work in progress is now picking up old data that is (in some cases) long out of date. For example, CBA Archives needs an identifier to show the location of Trust Bank records it has on permanent loan with TAHO, but there is no entry for TAHO (that I could find), only an old one for AOT. Who has (or will have) permission to add, modify, and delete entries? What protocols will there be? Will superseded entries be maintained so that programmes that have used superseded identifiers won't be cursed with broken links? Will there be an editorial/review process to iron out duplication, ambiguity, error, etc. from contributions made to the Directory?

It appears that contributors to the Directory will go on managing their own stuff (as before) without the editorial control and management I am suggesting.

1. What guidelines are being issued to potential contributors as to <u>what should (can) and what should not be put in</u>? Was it appropriate, for example, for each and every branch of NAA to be listed in the old Directory? I don't say it was, I don't say it

wasn't, but there should be a rule about such things. The new guidelines suggest that there is no rule, contributors decide for themselves. Is this a good thing?

2. What arrangements will be in place (if any) to <u>monitor the quality</u> of data going in and to ensure that contributions conform to the guidelines? What <u>permissions</u> are need to make, modify, or delete an entry? If I submit an entry for AONSW, for example, what is to stop me?

3. What are the arrangements for managing <u>currency</u>? Will contributions be dated? Will this date be refreshed when changes are made? Will there be a requirement for contributors to refresh their data periodically or else have it marked moribund?

4. What steps will be taken to monitor and manage <u>legacy data</u> when institutions fold or, as happens increasingly nowadays, they are merged (appropriately or not) with some random GLAM body?

5. What will be the protocol for withdrawn, superseded or moribund entries? Will they be <u>versioned</u> and maintained for the benefit of users who have made hyperlinks? Will superseded entries be marked with links to later versions? The Guidelines seem to indicate that contributors may withdraw entries. Should this be allowed or should entries once included be kept forever?

The Guidelines appear to be about updating existing entries and submitting new ones. They relate to mandatory/optional fields and technical requirements – not quite what I had in mind. I don't think they deal with any of the questions raised above.   If this is an ongoing process, why is there a deadline? It seems to be a project with a beginning and an end. It has the smell of a self-managed process and you know what they say: when everyone is responsible, no one is responsible. The entries will be a hodge-podge of identity data, descriptive data, very volatile data about services and contacts, and searchable metadata of various kinds. Because volatile data is being included. Attention needs to be given to currency. We don't want another one-off that gradually becomes dated and then has to be overhauled in another 20 years, do we?

<p align="center">ॐॐॐॐॐॐॐ</p>

Quo Vadis?

2018, September 27: From Perth 2018 9c Directory of Archives Project

The most depressing and muddle-headed session I attended. I questioned and received no coherent answer on what the functional requirements for this Project are. It is a cardinal error to supply answers to questions that haven't been formulated. Like putting the cart before the horse. Enthusiasm is no substitute for reflection.

Enduring Identity

It appears that entries for defunct repositories are to be deleted. That is not right. Once registered they must remain. Put dates around them and run a redacted version if you wish (but why would you want to?) but a register system endures. God forbid they start re-using the numbers. It never occurred to me to make enduring identity a requirement in the *Modest Proposal*. I just took it as read. An illustration that you can't assume **anything**. I'm not even sure if ISDIAH allows for time-bound relationships (previous/subsequent) and it would be too painful to look it up. If it does someone will tell me.

Including the Ungathered

They are using AtoM/ISDIAH and only wishing to notice "collections" or repositories. Requirement 4 calls for inclusion of the ungathered as well as the gathered. Examples of ungathered resources include land data, life data, geospacial data, statistical data, meteorological data, research data sets. They're unlikely ever to go to an archives repository. The objection to ISDIAH (in case you've forgotten) is that institutions

registered under ISDIAH are, in fact, corporations that should be registered under ISAAR.

... And then some

These were the only two issues I got to raise (out of a very long list). But I had already concluded they weren't going to get to Dublin from here. There was talk of expanding its role to something like a directory of archival resources. Ye gods – revenge of the flatlanders!

2018, September 27: From Perth 10 Replicating the Replicants ...

Replicating the Replicants, or, do archivists dream of downloadable sheep?

This was almost a re-run of a similar session at Parramatta (under the heading of ubiquity) involving some of the same speakers. It's about "liberating" archives from structures and hierarchy and enabling enriched and more useful access via connections (relationships) that are random and unbounded.

REPLICANT \rep'-li-cant \n. See also ROBOT (*antique*): ANDROID (*obsolete*): NEXUS (*generic*): Synthetic human with paraphysical capabilities, having skin/flesh culture. Also: Rep, skin job (*slang*): Off-world uses: Combat, high risk industrial, deep-space probe. On-world use prohibited. Specifications and quantities—information classified.

NEW AMERICAN DICTIONARY
Copyright © 2016

My response to the Parramatta presentation was along these lines

> It would have been possible to conclude that ubiquity, in and of itself, is an absolute good, virtuous in its own right ... My response is [that] ubiquity is neither good nor bad in itself but only contingently – by reference to how it is used, what it is for, what purposes it serves, and what requirements it fulfils (or fails to fulfil) ... contingency, limits, boundaries, and frameworks are virtuous in recordkeeping. It is how we include and exclude things and how we position them in relation to each other that supports (or imperils) evidence ... Both approaches are based on relationships. There is potential, therefore, for congruence rather than contrast. Alongside limitless and random conjunction ... defined relationships that 64ealized6464at recordkeeping (if only our descriptive practices were better employed) might thus powerfully deliver results that illuminate the character and meaning of the record. Meaning comes from a statement of what is and what is not; it comes from providing the user with information that this is important and that is not. Evidential relationships are relevant precisely because they are preferred over others, because they affirm the pertinence of one relationship over another – at source and prior to use. When all relationships have equal value, their meaning as testimony to the circumstances of their creation and use is vitiated: when everyone is some-body, then no one's any-body (W S Gilbert).

The problem with both sessions (today and two years ago) apart from the misleading impression that archival description has not abandoned hierarchy and singularity and has failed to embrace multiplicity, is that they do not contextualise themselves. Ubiquity enhances the role of the archivist as a custodian of heritage assets, but that is not the only role archivists fulfil. Making evidential records cannot be based on random and boundless associations – quite the opposite. And it is the work of the records maker that the records keeper inherits and must preserve. By a curious juxtaposition the concept of recordkeeping

by design came up in the next session (13a: Thinking Machines, smart applications and recordkeeping innovations)

Ubiquity is fine within the role of archivist-as-heritage-curator but inimical to the archivist-as-recordkeeper role. There are yet more roles in the space in between – such as those that support communities and special needs (e.g. Find and Connect) where a blend of structure to support evidence and ubiquity to support discovery is required. Failure to clarify that different approaches are appropriate to different roles simply confuses things.

2018, September 29: Keynote 4: … Archives should love Wikipedia

Keynote 4: Wikipedia loves Archives: Archives should love Wikipedia

The most enchanting session I attended – not least because it resonates so well with my *Modest Proposal* for a wiki approach to federated access. I went into this session feeling old and tired and I left (I swear to God) with a spring in my step – not because I thought Australian archivists would do something about it but because of the possibilities it disclosed.

It was all about how organised data can be incorporated into Wikidata and used not just by us but by everyone. That would include all of the higher-level structures I imagined would be needed for my wiki to provide a framework for contributions. My goodness! They may already be there. So far as I could tell, it would also accommodate the higher level contextual data from archival programmes themselves and provide protocols for resolving duplication (e.g. NAA's registration of the State of Victoria alongside PROV's registration of same).

With this resource available, how can archives programmes with budgetary restraints justify systems of their own to support such data? They would, of course, still have to maintain and update the data itself but they wouldn't have to maintain the systems. (If you look up the Wikipedia entries for kingdoms, principalities, countries, and political offices you will find an entity/relationship approach which accommodates ours). For that matter why isn't SNAC in Wikidata? I managed to ask a question verifying that externally maintained taxonomies of the kind I proposed we would need to make use of, such as the ABS work classifications, could also be incorporated into WikiData.

And, of course, it would be ideal for hosting the upgraded *Directory*. With the advantage, supposinglad can bring itself to concede the requirements set out in the *Modest Proposal*, that programmes for both gathered and the ungathered records could be accommodated without any further nonsense derived from AtoM/ISDIAH and the GLAM-orous flatlander wheeze.

<<Lise Summers: … I'm glad you enjoyed the Wikemedia presentation, and hope that you will get a chance to engage with Wikimedia Australia or other wikimedians to look at what can and cannot not be done.>>

2018, September 29: Perth 2018 Description -Quo Vadis? Part 1

Quo vadis? Can be rendered, inter alia, as "whither are we drifting?" or "and where are you off to?" This will be the first of a series of posts about it over the next week or so, deriving from sessions set up at the Perth Conference and at the Melbourne Conference a year ago – in particular from three sessions:

a) At Melbourne, a workshop that attempted to launch a collective approach by "small" archives to shaping AtoM into a series-friendly adaptation;
b) At Perth, a demonstration of Morty (Session 13b) purporting to be a proof-of-concept implementation of the Records-in-Context (RiC) conceptual model;
c) Also at Perth, a presentation (Session 17) by a panel of consultants of approaches at differing levels of conceptual and technical complexity and cost that might be used by

"small" archives (but no reason why it couldn't be used regardless of size) to develop series-friendly projects.

To begin, I shall try to lay the groundwork for what is to follow.

Background

What are the features of the descriptive landscape (for the moment, in a fast-moving world)?

a) *Describing Archives in Context (DAIC)*: in the absence of anything else, it is believed by some to be an authoritative statement of Australian theory and practice. It emphasises (wrongly) that "separation" of agencies and series is the defining characteristic but contains (muted) reference to the true essence of our approach – viz. entity/relationships.

b) *ISAD Suite (ISAD+)*: the four currently approved international descriptive standards that incorporate the hierarchical multi-level rule although the implications of that have been softened and have become almost irrelevant through successive editions that have incorporated more series-friendly adaptations. Comprising ISAD(G); ISAAR (CPF); ISDF; and ISDIAH.

c) Software: e.g. *AtoM* developed by Artefactual and purporting to implement the ISAD+ Suite but very forgiving of breaches of the ISAD+ rules. Other Software, such as *ArchivesSpace*, etc. Some of these provide a measure of digital asset control but other dedicated software – e.g. *Preservica*, *Archivematica*, *MirrorWeb*, etc. – exist that may be integrated with descriptive software.

d) *Records-in-Contexts (RiC)*: a purported normalisation of the four ISAD+ standards that has become something else altogether and could be the basis for a revolution in international descriptive thinking by replacing multi-level description with entity/relationship approaches.

e) Local Standards in Other Countries: e.g *US-DACS* which is currently being redrafted to abandon multi-level description in favour of an entity-relationship approach recommended by RiC (to say nothing of Australians who have advocated this approach over the last 50 years – and, believe me, they do say nothing; it's like we have been living on a different planet). Hold your breath – establishing this idea in North America would be a game changer.

f) *Morty*: a proof-of-concept implementation project revealed in Perth, the "concept" purporting to be RiC.

g) Emergence of ancillary descriptive support tools that can be used to enhance our endeavours (e.g. *SNAC* ; *ADB*; and *Wikidata*) and of federated access platforms (such as *Europeana*) that could be emulated here – some will think of *TROVE* in this regard but that is problematic for us in ways so complex that discussion needs to be held over. The rest of the archival world is getting a lot better than we are in deploying the results of description in imaginative and useful ways – honourable mention though for *AWM*.

Above all, as I alluded to in one my posts from Glasgow, the ground is moving beneath our feet. Just as it is no longer possible to think about stand-alone EDRM systems, it is no longer possible to think in terms of stand-alone descriptive systems. Description must be integrated conceptually and, so far as practicable, architecturally with the whole archival/r-keeping process (viva, Ian Maclean!). Even small archives that cannot do so in practical terms should learn to think conceptually of descriptive systems as merely one component of their architecture.

Disclaimer: Where Do I Stand?

My "friend" Barbara Reed stated in the Morty Session that I am an opponent of RiC. This is untrue. I am used to being misunderstood and misrepresented (poor me!) but, if portraying me thus can assist Barbara in whatever it is she is trying to do, I am happy to help in the role

of a straw man and as a stock figure of the unworldly theoretician out of touch with "practical" matters. But for the record and for the purposes of the postings to follow I should give a more accurate account of my own bias:

a) Conceptually, I stand by my understanding of the so-called Series System.

b) That is no longer the same as the articulation given to it by Peter Scott, although it is true (I believe) to the underlying concepts deriving from his work and that of Ian Maclean.

c) I think I am not alone in this role as a *continuator* of "series" thinking, but I would not claim that others of that ilk are in agreement with me or with each other.

d) No archives in Australia (including NAA) still practices the "Series System" as articulated by Peter. There is no uniformity in Australian descriptive thinking or practice that can be embodied in any single implementation or proof-of-concept. Worse, there is no proper understanding or acknowledgement that this is the case.

e) I believe RiC is a great break-through but I am wary of becoming over-enthusiastic about the prospects of its being adopted internationally. We must wait and see. Its impact is equally uncertain: what effect would the replacement of ISAD+ by RiC have on AtoM which is ISAD+ compliant, for example? Will archives around the world reconfigure their data to comply with standards based on a new conceptual model? How about a proof-of-concept that such transitions can be facilitated? If posing difficult questions be opposition, so be it.

f) No software package can (alone) provide any archives with what is needed. Descriptive software can only ever be part of the design solution underlying archival processes. Accessioning, processing, repository control, preservation, lending/issue, and (lord-a-mercy) the whole of r/keeping back out into creation-space, through the processes of migration, normalisation, replication, rendition, and digitisation, and then forward into secondary-user-space (including rights management and redaction) must now be aligned.

g) I agree, therefore, with Piers Higgs in the Consultants' Session that there is no magic bullet and that a variety of "solutions" (scaled up or scaled down depending on circumstance) rather than a single implementation model is likely needed – utilising a smorgasbord of software offerings and approaches to implementation design.

2018, December 12:

Just to underscore the point that AtoM is "very forgiving of breaches of the ISAD+ rules" here, for those not on the AtoM Listserv, is a response to a recent question from the Artefactual guru Dan Gillean:

> AtoM is purposefully designed to be very flexible and permissive, to allow for infinite customization and different uses depending on local conventions. I have seen users create collections that include fonds-level descriptions below them; I've seen users use repository records as collection-level records; Australians use the series as their top-level; and we have to account for users creating new custom level such as sub-sub-sub-fonds. Because we can't predict how users might want to implement AtoM based on their local conventions or how they might customize the levels of description, we don't organize the levels of description in any hierarchical manner that would restrict users from adding lower levels however they choose.

2018, September 30: Quo Vadis? Part 2 – Why Relationships Matter

The original impetus for *ISAD(G)* was to standardise archival descriptions to facilitate federated searching (it wasn't called that back then). I know, I was there. If all descriptions were alike they could be inter-sorted much like the old-fashioned union catalogues as described in the *Modest Proposal*. Unfortunately, this vision was driven by library-type thinking, Flatland thinking. It was Flatland thinking mixed up in an

unholy gallimaufry along with recordkeeping thinking in the standards they drafted. In Flatland, each asset is singly described and carries almost all of the metadata needed for its discovery. Federated discovery requires an alignment of the metadata and the technology has become very much better at doing that without the need for as much standardisation as was once thought necessary.

In Flatland, aligning metadata was achieved using "authority files" (e.g. *LC Headings*). An authority file controls the value of the characteristics of an asset. Following the library model, the ICA Descriptive Standards folks, when they began to accept a degree of separation between Doers and Documents, entitled the standard for describing Doers – *ISAAR(CPF)* – an authority record. To some degree they continued to think of *ISAAR* like that. I know, I was there. The emphasis was on bringing descriptions into alignment *a posteriori* by (essentially) building relationships as you search – just like ubiquity does in Flatland. This is what happens when archivists dream only of downloadable sheep (Session10) and it can produce some amazing and valuable results.

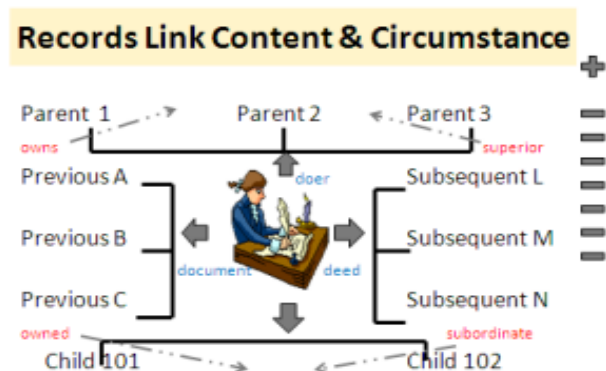

Figure Three – Access via TROVE

Figure Five : Records in Context

But evidential recordkeeping relies on *a priori* relationships, established at creation or identified as part of archival description, and preserved (in the service of evidence) thereafter. These relationships support different kinds of alignments – particular, contingent, and structured, not ubiquitous. Customarily, this was achieved by grouping assets together when they had a common relationship:

a) Items were grouped by reference to the Series to which they belong;
b) Series were grouped by reference to the Agency that creates them;
c) Documents were grouped by reference to the Activity that generates them;
d) etc. etc. etc.

The common device used was the dreaded hierarchical list which gave rise to:

- entrapment of entities within a singularity (you had to choose the group an entity belonged to);
- interdependence of descriptions (an Item description didn't make sense unless read in conjunction with the description of the Series to which it belonged).

The ICA folks sought to overcome these defects by borrowing the idea of authority records from Flatland. I know, I was there. In my time with them, the notion that such authorities could be used for the double purpose of separating descriptions of different entity-types, thereby placating the pesky Australians, had just begun to glimmer. They were genuinely puzzled why it didn't placate me.

> Disclaimer: This is not (or ought not to be) a binary analysis – on/off, good/bad, yes/no. Ubiquity and structure should augment each other. They both work and uphold different aspects of the recordkeeping enterprise.

MUCH ADO ABOUT DESCRIPTION

Relationships lie at the heart of the "Series System". The singularity is broken when we allow Series to belong (descriptively) to more than one Fonds. Initially, as formulated by Peter, that (along with the separation of Provenance and Ambience) was the only point of separation. He and I had ferocious arguments over whether or not one Item could belong to two Series. He was concerned, I think, to preserve as much as possible from traditional thinking so as to avoid further charges of heresy. But the logic of the S/system requires that we break-up singularity at all points – everything can be related to anything else – giving rise not only to multiple provenance but to simultaneous multiple provenance and parallel provenance as well (at all "levels").

Now, this may sound very much like ubiquity but it isn't because our kind of multiplicity supports a recordkeeping view derived from the actual, observed circumstances of creation and use (and preserved thereafter) and not from a Flatland view in which anything can be related to anything. The challenge for us, so far unmet as far as I can see, is how to

- incorporate r/keeping relationships into federated searching;
- support alignment of entities across the boundaries of Ambience (at all "levels").

Maybe that's one problem, not two. <u>Note</u>: we must stop thinking of Ambience, Provenance, etc. as defining characteristics of an entity-type. It is relationships that determine what role an instant entity plays. The same entity may confer Ambience or Creation depending on the relationships forged, not upon its essence.

All of which makes relationships so important. *RiC* has 800 or so of them, piled on, one after the other, in a riotous display of intellectual virtuosity (see how clever I can be!). If ever there was a case of the wood being lost in the trees, this is it. In the Morty Session, I ventured to suggest a way of dealing with this. They used the term relationship types to characterise the *RiC*list. If you read *Documenting for Dummies*, you will see that I made the point there that entities must be grouped into entity-types in support of scaleability. Thus Items, Sub-Series, Series, Super-Series, Sous-Fonds, and Fonds are all instances of the Document Type. The (as yet unproven) hypothesis is that we can construct rules around entity-types that apply more or less equally to all instances belonging to that type. Anybody: know about a proof-of-concept for that?

I asked the Morty folk to consider replacing their use of the term relationship type with the idea of instances of a relationship-type, but they didn't get it. They thought than when I said I could reduce the *RiC* list to about 15 I meant eliminating 785 relationships from the list. Quite the contrary. No one outside of Bedlam would dream of implementing 800 relationships. But it doesn't really matter if they're only instances of 15 types – the number could grow to 8,000 and it still wouldn't matter so long as we are clear about how many relationship-types they fit into. An implementation will choose which instances best fit the circumstances and then apply rules based on the relationship-type to which the chosen instances belong. That's the theory, anyway.

Take outs:
- r/keeping relationships differ from ubiquitous relationships – everything can be structurally related to anything but anything cannot be ubiquitously related to anything;
- we still have a long way to go in understanding how r/keeping relationships work and how they can best be used.

2018, October 1: Quo Vadis? Part 3 - Whither RiC?

RiC is a conceptual model in search of a concept – cf. *RiC at Riga*. The drafting Committee (EGAD) was asked, in effect, to normalise *ISAD+* along the lines I suggested in *Documenting for Dummies*, saying the same thing about dating entities in all four

standards, that sort of thing. They've produced something quite different, but hardly a conceptual model – yet. Apart from a few conceptual ideas (approving the entity/relationship approach, for example) it consists mostly of tables of possible entities and relationships about which I have already had my say – cf. *Records in Context (RiC) 1.0 – Comments on First Draft (2016)*. As for what kind of conceptual model might emerge over time, we must wait and see.

Now, for a happy thought. When draft *ISAD(G)* was launched we made a lot of fuss and I was made a member of the drafting Committee. At that time, *ISAD(G)* was accompanied by a *Statement of Principles*. At my first meeting, I was told the *Statement* was no longer open for discussion and they moved straight on to drafting *ISAD+*. I have told this story in *RiC at Riga* and elsewhere. Suppose, just suppose, that *RiC* is not intended to replace *ISAD+* but, instead, to subsume it.

Aims of ISAD(G)

- To define and control the **structure** of archive finding aids
- To define and control the **content** of archive finding aids

- Note: This standard provides general guidance for the preparation of archival descriptions. It is to be used in conjunction with existing national standards or as the basis for the development of national standards.

Hierarchy: Archival arrangement often follows a hierarchical structure, with series being further divided into subseries, files, and items. This hierarchical arrangement allows for the efficient organization and identification of records at different levels of detail.

Twenty-five years ago, I struggled to convince the internationals that it was not our intention to displace multi-level description in favour of a series-friendly approach. It was always my contention that an agreed set of descriptive principles could support alternative approaches that could be written into the standards. I'm not sure they believed me. After a lifetime of crushing disillusionment, I have come to believe that it is next to impossible to persuade someone of the correct answer to a question they haven't yet formulated for themselves. Could it be that what will emerge is a *RiC* that fills the gap left by the abandoned *Statement*? That after twenty-five years (!) EGAD stands on the cusp of repairing the damage done when the *ISAD+* path was taken without principles – not by abandoning *ISAD+* but by redirecting it?

It would not then be necessary to urge that *RiC* be made series-friendly. The international descriptive discourse would become (as I have always believed it should be) a song made harmonious with many melodies. Should this occur (God, I'm being optimistic here) the concept that would have to be proved is that they have succeeded in fashioning a conceptual model that can, in fact, accommodate series-friendly description (along with others). But for that we must wait and see.

Whither the Series System?

Peter Scott published what were effectively implementation guidelines for the S/system as it developed under his hand at NAA. The files there groan with many other unpublished memoranda. The whole experience came to be embodied in the *CRS Manual*, which I am told is still accessible on the web. Whether the *Manual* is still maintained I know not. I am certain, however, that neither its most recent version nor NAA's current practice still follow Peter's precepts. Nor do any other applications currently being maintained by *inheritors* of the system.

As for a coherent conceptual statement, where are we to go?

- There is the work of the *continuators* – those academics and writers (including myself) who have extended Peter's thinking and expressed themselves in articles, research studies, and metadata models. It's all good stuff (most of it) but homogeneous it is not.
- There is *DAIC* – which purports to be a conceptual presentation of the S/system but which can only be marked as C+. It was written in a muddle (and rewritten, and rewritten, and rewritten) and then issued in desperation to be rid of it. I know, I was there.

It follows that exercises in implementation now (such as *Morty*) must be guided by one person's interpretation of the S/system since no coherent agreed interpretation is available. We are all more-or-less operating in the same space and facing in the same direction but not with precision – no, not with that. The question is, would it be worth revisiting *DAIC* and coming up with a better conceptual model of the S/system? Would it, indeed, be possible?

2018, October 2: Quo Vadis? Part 4 – What really matters?

Alice comes to a fork in the road and is puzzled about what to do. Sitting in a nearby tree is a Cheshire Cat. She asks the cat which road she should take. "*Where do you want to go?*" asks the cat. "*I don't know*," Alice says. "*Then*," replies the cat, "*it doesn't matter*."
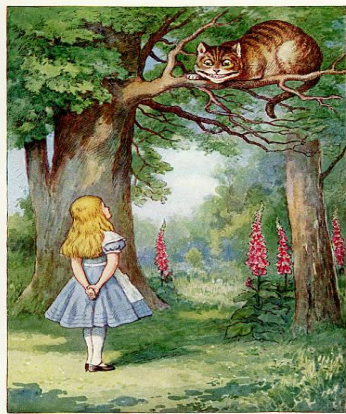
"*I GROWL WHEN I'M PLEASED, AND WAG MY TAIL WHEN I'M ANGRY*"

One talks of separation and of models, and of entity/relationship vs multi-level approaches, and of structure and ubiquity – but all of that is about technique. Technique is worthless gymnastics until it is deployed for a purpose.

What is our purpose? What is description for?
- ⬥ To make good records. Yes.
- ⬥ To preserve them. Yes.
- ⬥ To make them available. Yes.
- ⬥ To make them usable. Yes.
- ⬥ To make them interesting and more serviceable. Yes (but more dangerous, that).

But all of these are objectives which the techniques are employed to achieve. They still don't tell us what direction to take (where we want to go). They don't tell us what matters.

Above and beyond the technical features of the S/system (whatever they may be) lies a view about what matters. A philosophy of archiving, if you like. It is seldom spoken of (alas) and is (I fear) often lost sight of by some of its practitioners. But the ghost of Ian Maclean speaks of it loud and clear. What matters is recordkeeping. If we are anything, we are recordkeepers – first, last, and always. What does it mean to be a recordkeeper? Well, thanks be to Jenkinson, a lot has been said and written about that and it's a discourse that is still alive and well. Isn't it? (he added nervously).

Nothing in the contemporary descriptive discourse is more distressing than the all but ubiquitous use of the word "collection". That others use it constantly is not surprising because their use of it is supposed to be the thing above all others that marks us and our way of doing things off from them and their way of doing things. Not using it is supposed to signify what it means to be one of us – to be on the side of the S/system. Nowadays I find that we are using it too – all the time. Understanding that collecting doesn't matter is supposed to be the hallmark of our understanding of what does matter. Terminological carelessness is a little thing but, in this case I fear, it goes deeper - that it reflects an ignorance of what matters, that we have forgotten.

Ubiquity also abhors collections. Collections are boundaries that ubiquity seeks to traverse. How odd that recordkeepers seem to have forgotten that. And do not subscribe to the great untruth that collecting can be enforced upon you by circumstances. It doesn't matter what

kind of materials you deal with – personal or institutional, private or official, current or non-current. It doesn't matter what kind of employment you are in – library, archives, museum, gallery, or office. Wherever you are and whatever you do, you can still have a recordkeeping frame of mind.

So, what matters is the focus of our descriptive efforts. That, unsurprisingly, is the connections (the relationships) we employ and an understanding of the different purposes we use them for. Clarification of the different purposes of structure, ubiquity, and collection would be a first step to understanding but certainly not the last.

2018, November 27:

<<Debra Leigo: Are you suggesting that the purpose of description is to "make good records" through the inclusion of contextual metadata? Otherwise (she asks hesitantly) shouldn't the task of "making good records" be the responsibility of the creators? May I suggest an expansion to one of your statements: "Wherever you are and whatever you do, you can [and should] still have a recordkeeping frame of mind." Interestingly Kate Cumming suggests that "[Records Managers] need to be more archival in our thinking"
http://archiveslive.ning.com/video/video/show?id=6334236%3AVideo%3A133377&xgs=1&xg_source=msg_share_video
OR
https://www.youtube.com/watch?time_continue=658&v=6QVUpEyRSCE
It seems that Jenkinson is, perhaps, not as alive as could (and even should) be, given that we seem so keen on collection rather than natural accumulation (she adds tentatively). In addition to preservation and access - making records available and usable into the future - what is our purpose - do we need a greater purpose - and where would you like us to go, Chris? >>

It is emphatically not my view that our role is limited to preservation and access. We are recordkeepers and that role extends to both records-making and records-keeping. I don't want us to go anywhere we shouldn't already be and should always have been. Descriptive skills are needed to make records and to keep them. That being so, anyone involved in the making and keeping of records, not just archivists but recordkeepers however designated, must be masters of description.

> **<<Are you suggesting that the purpose of description is to "make good records" through the inclusion of contextual metadata?>>**

Yes, but not only (or even mainly) through the inclusion of contextual metadata "in the future". I take an expansive view of "description" and see it as a core recordkeeping activity. In traditional practice, we had the content of the artefact, the conjoined "metadata" (we didn't call it that back then), and the related metadata management systems (registers, indexes, movement) governing the artefact. Conjoined metadata is the file no,. initials, etc. embedded on the face of correspondence, for example, or the numbering/titling/dating inscribed on a container (file or docket). When we showed David Bearman a standard issue file cover he was enchanted because he'd never seen one and the protocols for handling the metadata were all laid out. I would include all this, handling the data about the content, under the term "description" and we haven't even got to contextual metadata yet. Sometimes, not always, the name of a department or business unit would be inscribed on a file title (never, in my experience, on a docket) but even this was not full contextualisation (it lacks information about purpose, process, and function, for example) and files often move on from their native creation-source (multiple provenance),

In the old times, contextual metadata came later when the artefact moved out of an environment where knowledge of origin, purpose, and use could be assumed, while it was still in the care of registry clerks, for example - the living finding aids as I once called them.

🏠 MUCH ADO ABOUT DESCRIPTION

This looks like a life cycle, but conceptually it isn't. In an archives they joined artefacts with other origins. Then an archivist stepped in and set down knowledge of the "missing" contextual data from the source system(s) and other metadata needed to distinguish artefacts from one source from those coming from another OR added descriptive metadata concerning post-formation activity that had never before been captured as part of the continuing process of records-making. On this view, the archivist is as much a records-maker as the file clerk who opens the file. I will let others, who are more worried about such matters, speculate on the authenticity of post-formation description added to an artefact (in my view, so long as the r/keeper's hand is transparent, anything that aids understanding and use is better than nothing).
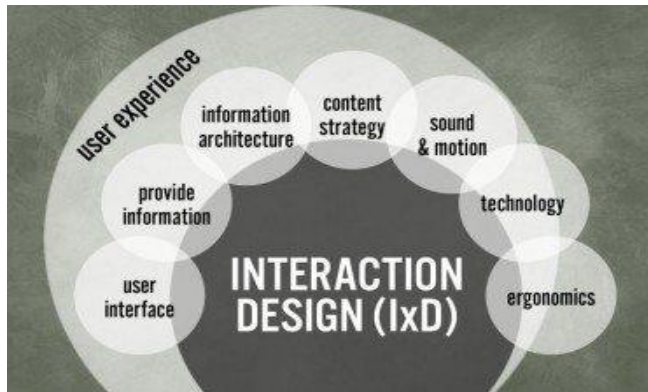
In the digital world, the boundaries between separate r/keeping processes are much more porous; records enter a larger context from the word go. Digital r/keeping, therefore, requires "us" to pay more attention to description up front. It now lies at the heart of what we do from the outset More descriptive metadata must be inscribed onto the record than we were formerly wont to do. The living finding aids have all been made redundant. System administrators move on fast to seek new fields of endeavour. Records are duplicated and moved about within an inter-operative native environment. Systems themselves lack stability in a technical environment of change, updates, upgrades, and migration. To survive, the record itself must now carry with it much more of the descriptive load and that includes a lot of contextual information that we formerly left to the archival stage. In fact, in the digital world, archiving begins the moment a record is formed and you stay our hand over the delete button (and maybe also if you don't stay your hand supposing that you want a trace of the undocumented transaction to survive).

<<Andrew Waugh: Actually, I like the word collection for what's in an archive. For me, it reminds me that one of the things archivists do is select what is to be preserved as 'the record'. It highlights the role of the archivist as a gatekeeper. Thinking of an archive as a 'collection' means that you are thinking about why the archive exists; what you have, what you don't have but should, and what could be disposed of. ... In short, the word collection highlights that archivists are building the archive for a purpose; archiving is not a neutral act.>>

Well, Andrew, it sounds to me like you'd be right at home at the Documenting Australian Society Summit

<<Andrew Waugh: While I wish them well, it's actually the reverse of my view. The summit starts from the question of what information should be kept, and then considers who should be responsible for collecting it. My view is that an archival institution should consider what its mission is, and then work to ensure that its archive is the best possible match to that mission. (Although I think that the summit is addressing an interesting issue; it's essentially a sanity check to make sure that important aspects of Australian society aren't falling through the cracks between institutions.) >>

MUCH ADO ABOUT DESCRIPTION

<<**Joanna Sassoon: To me, this is all about how we think … and the difference between how we transform information into evidence - whether of 'me', 'them' or 'us'. Very unGLAMorous.>>**

Quite so. Collecting is an archival method, it should not become a mind-set. It is one amongst many different methods towards an abiding r/keeping purpose. That purpose can be achieved in a variety of different ways. It becomes dangerous when it forms the basis for the way we think about r/keeping. because it blocks our minds from other (possibly better) ways. This was the essence of the message in Bearman's *Archival Methods*. If you approach r/keeping with the mind-set of a collector, you <u>will</u> go astray.

I'm reminded of Lincoln's answer to Horace Greeley who accused him of a lack of focus.

> As to the policy I "seem to be pursuing," as you say, I have not meant to leave any one in doubt. I would save the Union. I would save it the shortest way under the Constitution … If I could save the Union without freeing any slave, I would do it, and if I could save it by freeing all the slaves, I would do it, and if I could save it by freeing some and leaving others alone, I would also do that. What I do about Slavery and the colored race, I do because I believe it helps to save this Union, and what I forbear, I forbear because I do not believe it would help to save the Union …

That is my case. If I could make and keep good records by collecting them, I would and if I could do so by dismantling all collections I would do that, and if I could do so by dismantling some and keeping others that is what I would do. Mistaking method for purpose is the evil thing. But paradoxically (perhaps deviously) he abolished slavery in the end.

<<**Adrian Cunningham: "… I have never seen recordkeeping as an abiding purpose. Just as collecting can be a legitimate means to achieving one or more higher purposes, so is recordkeeping just a means to one or more worthy higher objectives … While defending records may be a worthwhile and important thing to do ,,, it s not in and of itself something that would ever motivate me to get out of bed in the morning. So what are the higher purposes that recordkeeping and collecting might help us advance? Social justice. Good, transparent and accountable governance. A healthy democracy. Human rights. A society that can understand, explain and account for itself over time through its documentation. They are all things that get me out of bed in the morning. And if recordkeeping and collecting help us achieve those things then I am all for both of them.>>**

I am all for a long life, a healthy diet, and an ample daily portion of red meat. Ooops! An ample daily portion of red meat and two vegs was once thought to provide a healthy diet but no more (the idea has been sourced to the Depression when it wasn't possible to get much red meat and to a time, my parents' time, when pricing ensured that the portions were smaller anyway). It was once possible to say "I am for both of them" but ideas about nutrition have changed.

> Napoleon's plan for 1812 -
> 3.Conquer Russia
> 2.Capture Smolensk
> 1.Get out of bed

Whether or not r/keeping itself has a "higher" purpose isn't really material. It certainly has uses other than those worthy ones that Adrian nominates and which I imagine wouldn't get him out of bed in the morning - totalitarian oppression and the denial of human rights to name only two. The point is this: are collecting and r/keeping simply two means leading by different paths to the same end (like rent subsidies and public housing are two <u>alternative</u> ways of providing low-rent accommodation for the poor) or is one a means to achieving the other (like getting out of bed is a <u>necessary step </u>towards capturing

Smolensk)? You could say that getting out of bed is also a necessary step towards conquering Russia but that hardly puts it on a level with capturing Smolensk.

Of course, a larger view can always be taken (of almost everything). Napoleon got good advice that trying to conquer Russia was a bad idea and if he'd taken it the stars may have aligned differently for him. If SOCIAL JUSTICE or some other high purpose is the reason for r/keeping we still must ask how does collecting fit? If it is a co-equal and equally satisfactory means to the same end (as Adrian suggests) then they are alternatives and you could achieve SOCIAL JUSTICE through collecting without r/keeping (and, please take note Andrew, vice versa). That may be the way that the GLAMorous strive for SOCIAL JUSTICE, but it's not our way.

If we employ collecting, it is because it serves the ends of r/keeping not because it is a satisfactory alternative. So, the undistributed middle term in this syllogism is the answer to a question not yet asked or answered: how do we employ r/keeping to achieve SOCIAL JUSTICE (or whatever higher purpose we have in view) and is that materially different to the ways that others strive for SOCIAL JUSTICE? Even if the GLAMorous are also striving for the same larger purpose through collecting is that path an equally satisfactory one for us or do we follow a different star?

2018, November 28:

<<Adrian Cunningham: Chris says 'Whether or not r/keeping itself has a "higher" purpose isn't really material.' I disagree - it is the most material thing we can grapple with in our working lives >>

I disagree too – with your interpretation of my words. They were clearly intended (I thought) to say that the "higher" purpose is immaterial to this discussion not that it is immaterial in an existential sense. Why? Because whatever the higher purpose(s) maybe you don't need to stipulate them in what is essentially a discussion about contrasting methods. The distinction between the "lower" purposes of r/keeping and collecting can be debated without reference to the "higher" purposes which I take to be broadly similar for the practitioners of both methods. I think I can stand on over 40 years of my writings as sufficient testimony to my commitment to the proposition that r/keeping must have a purpose and that my devotion to that purpose should not occasion anyone to quit their job (devoted collectors apart, maybe). What I thought this discussion was about is how we (as distinct from others) achieve that purpose - a discussion about r/keeping vs collecting as means to an end.
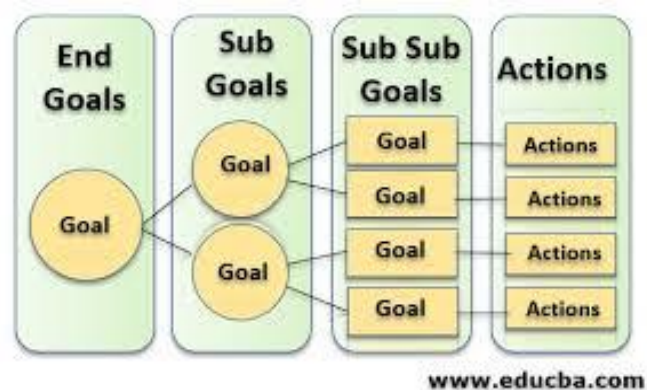
<< **I was not suggesting that recordkeeping and collecting are co-equal and equally satisfactory means to the same end**.>>

Good-oh. I would go further and say that collecting is an "equally satisfactory means to the same end" if we are talking about a "higher purpose" that r/keepers and collectors share. My argument is that how r/keepers approach that shared end is different from the way collectors

do. Perhaps "equally satisfactory" would be over-stating it, though. We might want to explore and contrast different ways in which the two approaches support a similar "higher purpose". I don't want to launch a further debate but maybe look at evidence vs poetry and different meanings for evidence found in archives and collections. This is more than a cultural debate (in the narrow sense) it also relates to the quality of those 200,000 documents being used so entertainingly in the Banking RC.

> << **They are different (and in some cases related) activities which may or may not contribute in different ways to achieving the same end... collecting is only one of many ways in which recordkeeping outcomes may be pursued...** >>

I agree. Collecting is a perfectly acceptable method for archivists, r/keepers, and others to employ. But when archivists collect they do so with r/keeping goals in mind, essentially evidentiary goals - a term which has a quite specific meaning in our world and relates closely to the methods we employ. Referencing the 200,000 documents, we also need to be aware that our use of "evidence" is not co-extensive with that of lawyers (let alone lay persons and journalists). So ...

> << **it is important to grapple with how the two activities inter-relate with and differ from each other** >>

Yes. Where I come down is this: as Joanna said it is about "how we think" rather than about what we do or where we work. You can work in a collection and think like a recordkeeper and (regrettably and all too frequently) you can work in an archives and think like a collector. The forthcoming Summit (if its propaganda is to be believed) is focussed on collections but it is tendentiously titled *Documenting Australian Society*. As I said in an earlier post, r/keepers should approach that topic from a completely different frame of reference.

<<Michael Piggott: My thoughts:
i) surely support for record making and record keeping, incl. funding, is strongly tied to the purposes of record making and record keeping in the relevant societal context, and the more convincingly and cunningly those purposes are articulated, the more one's changes of support improve?
ii) surely there's a continuum of means/ends/means/ends, which gets you out of bed to .. to ... to beat the Russians ... ; and closer to home for example, the "means" of good records regimes and practices promote the "ends" of a well run Red Cross blood service and box flight recorders promote the "ends" of a safer aviation system. And yes (sigh), one could still ask, aha!, but WHY do you want to beat the Russians or have safe blood fusions .... The continuum has a long tail.>>

> << **one could still ask, aha!, but WHY do you want to beat the Russians or have safe blood fusions** >>

Isn't the answer to that : 42?
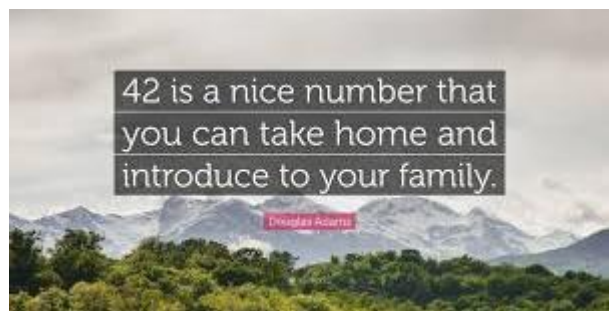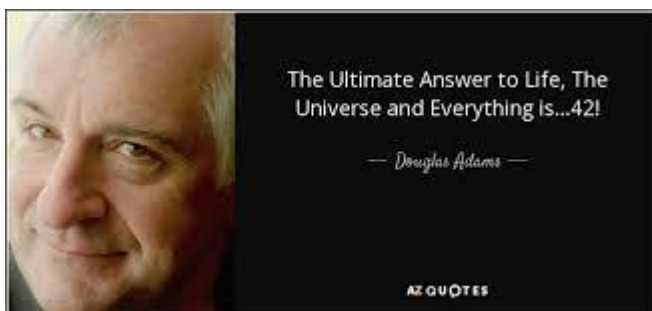
2018, October 3: Quo Vadis? Part 5 (final)

A year ago, in Melbourne, a workshop looked at possibilities for enhancing AtoM as a more series-friendly implementation in "small archives". It was announced in the Consultants' Session (Session 17) that this didn't really go anywhere. No more one size fits all. Instead,

they took a new tack in this Session by providing guidance and examples of tools to assist small archives in going about the business.

Stand-Alone Descriptive Software

Proprietary products, even if they are Open Source, aren't cost-free. They take up time and resources to deploy and to maintain. They are technologically dependent and I have yet to hear of an instance in which the archivists don't need IT support of some kind. AtoM seems (to me) to suffer from being developed on the run and there are other problems – e.g. download issues related to Linux (good for universities, bad for banks). I am a lurker on the AtoM User List and a day does not go by when there aren't a dozen or more postings about bugs, problems, and queries. Artefactual are very good at maintaining this List and offering free advice but it suggests to me that small and sole archivists would need support. That is available from Artefactual itself at a cost or from someone else who would also need to be paid or budgeted for. I should also add that many archives (including ANU) seem to have installed AtoM without (apparent) difficulty but that may prove the point because I suspect they had in-house support.

Integrated Descriptive Solutions

As alluded to in earlier posts (and in the Consultants' Session), stand-alone descriptive systems are becoming a thing of the past. Even small archives need to look at the possibility of integrated approaches and the Consultants' Session was basically about how to go about this. I approve the approach taken and have nothing to add. But there is another aspect to integration. As alluded to in my session at Parramatta two years ago, the focus on digital (and digitized) assets runs the risk of sidelining the as-yet-un-digitized physical assets, for the foreseeable future the overwhelming proportion of the assets we manage. Our descriptive efforts must continue to ensure that their existence is noticed and that access is given to as much information about them as possible in ways that are integrated with, not separated from, access to digital/digitized content.

Meanwhile, What About Standards?

I am still not clear about what Morty is about, but anything that makes standards more series-friendly is a good thing because it means standards-compliant software (like AtoM) will be more useful for us.

And, What About Really, Really Small Archives

These are the ones I call the barefoot archivists: tin shed, uncertain electricity, and Internet only three days a week. Community archives, historical societies, back office on Thursdays, that sort of thing. Requirement 3 is for them to be included somehow. They may still need a product that truly is simple and cost-free like *Tabularium* was. Tabularium wasn't actually free, of course, it was just done as a labour of love by David Roberts.

Federation?

The default position is that our descriptions are essentially available in one of three ways

- seriatim, knowing where to look and searching the on-line tools provided by each archives;
- looking in TROVE to see if the assets have been harvested (or contributed);
- doing a Google search and hoping something shows up.

The last two are searches in Flatland. Ubiquity, if I understand it correctly, points towards clever, purposeful, targeted harvesting as an alternative route to the same end.

In the *Modest Proposal* I tried to get a conversation started on what the requirements are for federated searching in our world. I suggested a wiki approach to illustrate one possible implementation model (not the only possible one) but my first concern is clarifying what we want and need, not how to get it (that comes later). There has been very little interest shown.

Do we think it's an issue? Do we care? The Directory Session threw out a possibility that the *Directory* might be used to prompt searchers (somehow) towards institutions – for a seriatim search (I suppose). In Hobart, Michael Piggott et all threw out the possibility of a resources-assessment approach that I thought could be integrated into the wiki proposal. But I have no sense that federated searching is a live issue amongst us at the moment. Am I wrong?

2019, March 6: "… Wikipedia/Wikimedia in Australia in 2019" …
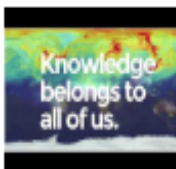
"Check out "Pru Mitchell - Wikipedia / Wikimedia in Australia in 2019" on Archives Live

Archives Live
Australian Society of Archivists Inc. - Keeping Archives, Keeping Records

Archives Live…
  Archives Live
  Administrator-

Check out the video 'Pru Mitchell - Wikipedia / Wikimedia in Australia in 2019'

Knowledge belongs to all of us.

Pru Mitchell - Wikipedia / Wikimedia in Australia in 2019
Using Wikipedia & Wikidata Sydney, 4 March 2019
https://www.meetup.com/en-AU/Information-Innovation-UTS/events/256103126/

Video link:

Pru Mitchell - Wikipedia / Wikimedia in Australia in 2019

I attended the Wiki-World session in Sydney last Monday. Lots of stimulating stuff and lots of unanswered questions. The three presentations are now viewable via Archives Live:

- Pru Mitchell – overview
- Tony Naar – Wikipedia and the Australian Paralympic History Project
- Toby Hudson – Using Wikidata for Chemistry, Education, Australia, and #FakeNews

Two lines of thought suggested themselves to me.

Using Wikipedia to do the job for us in the management of our content

- The Directory Project. Why not use Wikipedia as our platform the way the Tony Naar describes? The good news is that they know about persistent identifiers the way some archivists seem to struggle with. Many of our archival programmes are already there – e.g. NAA, State Records NSW, AWM, Geelong, etc. etc. and Wikipedia already contains lists of archives in the UK and in Canada.

- The Modest Proposal etc. I've already suggested a wiki approach to establishing a contextual gateway to resources across institutional boundaries. Most of the content needed to populate such a project would already be readily available – buried in the host systems and easily extracted. The main problem would a protocol for keeping it up to date. And volunteers (I'd volunteer for it in my retirement in a flash). PS Wikipedia loves digitised content.

Using Wikidata in the management of our data – e.g. for online "finding aids"

Wikidata holds data sets and uses techniques such as linked data to extract value as described by Toby Hudson. Most archival "finding aids" produce reports and lists (very 20th century) and these are, in effect, data sets. We are also using generic search-and-display

protocols borrowed from libraries, et al. to respond to user queries. But what if we think outside the box for a moment and approach it from an enabler point of view? I suggested in an earlier post that access enablers may be the future model for reference archivists. Our structured search fields (place, name, topic, etc.) are all taxonomically controlled elements that could be linked to data set(s) to establish patterns that are effectively search results (I think – I'm very new to this). That would result in a completely different experience for the user than simply display-narrow-filter and one in which he/she would be more in control (I think – I'm very new to this). Global searching would present different challenges, but I'd like to hear from someone who is more across data sets and linked data to shoot this idea down. PS This idea does not preclude the development of archival systems; it is about use, reuse, and deployment of the data we manage BUT for those archival programmes that struggle with developing and funding archival systems their descriptive data at item and series level (generated in simple database or spreadsheet formats) could be exported as data sets and Wikidata used as their front-end (?maybe).

<<Lise Summers: ... I have long thought that Wikidata might be a useful tool for small archives, which is why we had the Wikiworkshop and presentation at the 2018 conference. Wikisource and Wikicommons may also provide ways of making content available. There is a wikipedia and libraries facebook page - Wikipedia + Libraries. (Maybe we need a group for Wikimedia+Archives ?>>

2019, March 7:

Count me in.

2019, July 3: Wikidata for archivists

Some months ago, both Lisa Summers and I speculated on this list about the potential of Wikidata for Archivists – especially the littlies and barefoot archivists. Developing SNAC– like shared contextual data is one possibility. Well, now there is a group for just that very thing.

2019, October 26: Adelaide 2019 – Session 4.3 AtoM

Situation not much changed from last year. Some penetration in Australia but not yet ubiquitous amongst medium-to-small. SROWA the only government archives wholly committed but limited use and dabbling by some of the others. Realistic assessments of pros and cons. Continuing lament about lack of fit with Series System (whatever that means). No one present able to predict how RiC might affect ISAD-based AtoM. There is now a Foundation formed (or forming) independent of Artefactual (apparently) to consider suggestions for improvement.

Fit with Series System:

This always saddens me. When I joined the ICA Commission developing the ISAD suite (at ASA's expense and to represent the Australian view), there was a widespread misconception amongst the internationals and amongst some in Australia that I was trying to get them to replace Fonds-based description with Series-based description in the Rules. This was not my brief and it would have been ludicrous to try but I gave up trying to persuade people of this. My task (as I saw it) was to re-jig the draft Rules so they were equally and simultaneously supportive of both approaches. Perhaps naively, I have always thought this was (or would have been) easily done. After all, there aren't two different ways of dating things (tho' there may be debate over which things to put dates on). What I foresaw then was that, without such fusion, software development would lead to just such a situation as we have now – confliction and confusion in a space where we should (and could have been) working together. Some convergence there was over the years and RiC might take that even further but it has all been ... well, untidy. And unnecessary. A lost opportunity if ever I saw one. Sigh!

User Front End

The realisation is growing that this needs attention. I don't disagree. But this can be said of the whole archival enterprise. It is our most neglected area and our biggest failure. And you can't just single out AtoM for blame. Everyone is at fault. More of this anon.

Relationships

This is the area where the disjunct between Fonds- and Series- based description is (still) most striking. Well, I would say that wouldn't I? But it is far from unresolvable. A simple understanding of the difference between reciprocal relationships and non-reciprocal connections would do the trick. Meanwhile, there is insufficient understanding that as well as formal, documented relationships between entities, we have hidden relationships in the narratives and lists. Once they are identified and unpacked, we may see the last of those dreadful lists.

2019, October 26: Adelaide 2019 – Session 1.2 The descriptive tradition

> *We shall not cease from exploration / And the end of all our exploring /*
> *Will be to arrive where we started / And know the place for the first time.*
>
> T S Eliot "Little Gidding" (*Four Quartets*)
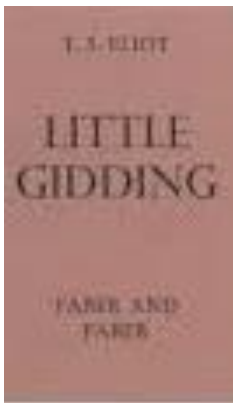
One of the less pleasing things about conferences is pronunciamientos from the platform of insights that give us new tools to work with and enrich our users' experience but are presented as breaks with our tradition, repudiations of past practice, sometimes as denunciations of our predecessors' obtuseness and wrong-headedness. Indeed, as I age, I feel more and more like I'm becoming a predecessor and sometimes a target. Presentations of new ideas is what you come to conferences for. But the accompanying razzmatazz can be irksome. There have been truly revolutionary moments in our world and I have no doubt there will be again. But they are few and far between. The over-hype is irritating and the lack of modesty is embarrassing.

There is a tradition (my particular interest being our descriptive tradition) and things must change if they are to stay the same. Tradition is the glue that holds us together and the foundation for moving on. I am conscious that in much of my own work I could be accused of trashing it (I don't think that's true but I am blushing a little as I write this). I hesitated, therefore, about raising it because I know I am vulnerable. The over-blown claims to which I take exception are graceless and untrue and if that were all I would keep my grumpy self to myself. But larger issues are in play.

Tradition is a long, winding path made up of stepping stones, used to build on previous insights when taking the next step. These new steps may involve insights of which our predecessors were ignorant or to which they were blind or (increasingly) they may involve technological developments not available to them. The temptation for young innovators in search of fame and admiration is to ignore the tradition in which they work and repudiate

those who pioneered it. Their excitement in discovery should not be curbed but they also need to be counselled in the true nature of progress.

Repudiation and disrespect of tradition denies to innovators themselves the value of earlier work as a foundation and direction-finder for what they are doing (or even just a point of departure conceptually) distorting their understanding of their own innovatory ideas and inviting their audience to launch off along a fork in the road rather than recognise how a professional discipline should (and for the most part does) develop step by weary step within a constantly evolving mainstream. To the extent they are believed, we are all losers because we come to devalue (at worst, to disrespect) tradition. The next generation of innovators, when today's innovators have become tomorrow's fuddy-duddies, will lose the strength that tradition bestows, will be more likely to roam far and wide outside the direction it sets, will find it harder to get a secure footing on the next step (standing on the shoulders of giants). Far-roaming can liberate but it can also stifle.

A case could be made for trashing tradition and roaming far and wide outside it – particularly in times of dramatic technological developments like now. I think this case stands or falls on whether enough tradition-trashers are in fact coming up with insights that are truly revolutionary. On that basis, I don't think the argument can be sustained. I will use Session 1.2, not because the presentations were especially egregious but because I was there and they help to make my point.

Ubiquity and Structure

This reprised presentations at other conferences. I won't repeat my earlier analysis except to say that my plea to understand ubiquity and structure as complementary rather than alternative methods of description makes my point. Ubiquity can be understood (I think without violence to its conceptual integrity) as a new approach to reference guides. What is novel about it is that it multiplies the opportunities, puts the users in charge instead the archivist, and permits a new approach to discovery that is essentially iterative. This is all truly new, truly valuable, and truly exciting. But revolutionary?

City of Sydney User-Based Finding Aids

They did surveys and found their user base comprised skimmers, delvers, and deep divers. We used to call them serious researchers and genealogists (how dumb was that). They claim to have discovered that most of our users want item-level access. Well, yeah. We've kind of known that for quite a while now. They've figured out that the structure of the descriptive data doesn't have to limit the design of the user interface. If they'd asked their colleagues there are plenty around who could have told them that but I suppose validation from the surveys does no actual harm. Some of us would have confirmed their conclusion that poor user interface design is the profession's biggest failure and some robust tradition busting on that front is badly needed. They want user feedback to guide future development. Not a bad idea (or a new one, for that matter) but …

- User surveys are good at discovering what they think but bad at discovering what they might think if they knew more.
- There is no need to be binary. Clever design can yield an interface that satisfies all, not just the largest user group.
- The distinction between user wants and user needs remains valid (they want McDonald's but they need vegetables).
- You may be trapped when your user base changes and lock out potential users who haven't yet voiced their wants and needs.

<div align="center">❧❧❧❧❧❧</div>

Maps

2018, October 11: Fun with maps

The Guardian has a story showing a gallery of maps that display errors, falsifications and other infelicities. I like –

> **North America, 1783**
> This map was used by the British at the 1783 conference that established the independent boundaries of the US. One drawn red line shows that the British were amenable to relinquishing even parts of Canada. When they got off more lightly, the map was embargoed for more than a century to prevent anyone finding out.

But my favourite is –

> **West Indies, 1506**
> Since the earliest printed map to show the Americas was produced a mere 14 years after Christopher Columbus made American landfall in 1492, it was still believed that he had reached Asia. Accordingly, this map shows North America welded onto eastern Siberia, with Cuba and Hispaniola floating in the same sea as Japan.

because I often use this case as an example in my presentations and now I have a new slide.

2019, October 20: … mourning the end of paper maps

"The perfect combination of art & science" mourning the end of paper maps

<<Joanna Sassoon: A **reminder** of the joys and practicalities of the material world.>>

2019, October 21:

As a way of understanding and representing the material world (materially or digitally) map-making resonates with our work of understanding and representing events and circumstances (descriptively or curatorially). I once used an episode of *The West Wing* to illustrate this when explaining parallel provenance.

2020, June 25:

> <<Digital maps might be more practical in the 21st century,
> but the long tradition of cartography is magical>>

And, like the best traditions, it keeps on evolving (magically) -

> Earth's mysterious eighth continent doesn't appear on most conventional maps. That's because almost 95 percent of its land mass is submerged thousands of feet beneath the Pacific Ocean. Zealandia — or Te Riu-a-Māui, as it's referred to in the indigenous Māori language — is a 2 million-square-mile (5 million square kilometers) continent east of Australia, beneath modern-day New Zealand …

Now, GNS Science — a geohazards research and consultancy organization owned by the government of New Zealand — hopes to raise Zealandia (in public awareness, at least) with a suite of new maps and interactive tools that capture the lost continent in unprecedented detail … The new maps reveal Zealandia's bathymetry (the shape of the ocean floor) as well as its tectonic history, showing how volcanism and tectonic motion have shaped the continent over millions of years …

The team also released interactive versions of both maps on a new Zealandia webpage. Spend a few minutes clicking around the hyper-detailed images — and, when someone asks what you're doing, simply tell them you're "discovering Earth's lost continent."

2023, September 25:

There's life in the old dog yet

For 400 years British hydrographers have made paper charts of the world's seas and oceans. Each one captures the detail of coastlines, bays, straits, or channels. A document like this brims with information, noting the sea's depth at various locations, the position of rocks, or places where vessels can't drop anchor … Every day, staff at the UKHO [UK Hydrographic Office] make corrections or improvements to some of the **3,500 charts** they maintain, such as adding the location of hazardous new wrecks and submarine cables or even changes to coastlines. A weekly bulletin communicates adjustments to shipping vessels worldwide and crew members must then get out a pen and manually correct any outdated paper copies. Yet the last of these weekly updates is on the horizon. The UKHO is gradually **preparing to drop its paper chart service** and switch to digital-only versions, which would be accessed via Electronic Chart Display Systems on ships … The digital transition will bring to an end the tradition of hand-drawing hydrographic charts, then designing them on computer software before being printed. More recently, the UKHO started sending electronic copies to customers who could print the charts themselves.

It turns out that quite a lot of ships still need paper charts. Due to maritime regulations, vessels must carry some form of chart and, despite the availability of electronic versions – which don't have to be manually updated every week – paper charts continue to be used as backups, or, in some cases, the only such resource on board. The Royal Yachting Association has also said that despite the withdrawal of the UKHO paper charts, it will continue to **teach navigation techniques that use them**. Paper, it seems, still rules the waves.

… A 2,000-year-old tradition, real paper made from trees is still considered crucial to countless businesses and government systems globally, despite **the environmental impact of producing it** … **A study published in 2021** indicated increased brain activity is associated with remembering information once it has been written down by hand, as opposed to recording it on a smartphone or tablet … "The mind better grasps elaborate, complex, deep arguments that run over several pages of paper," says [Richard Harper, an expert on human computer interactions at the University of Lancaster], noting that when you have something particularly nuanced and elaborate to say, putting it down on paper may be a good idea … Much like the UKHO, lots of organisations attempt to go largely or exclusively digital only to encounter hurdles. The US government **is due to go paperless but it is taking longer than expected**. Last year, the National Archives and Records

Administration found a third of the sprawling federal government had still not adopted e-records and the Administration was forced to extend the deadline for this by 18 months, to 30 June 2024.

Paper also plays a role in more clandestine sectors of government. In the UK, for example, the Government Communications Headquarters (GCHQ) **keeps thousands of secret paper files on its personnel in a vault in its basement** while the MI5 security service **states on its website**: "Paper files remain important to MI5." Russia's Federal Guard Service (FSO), which is responsible for security at the Kremlin, reverted to using typewriters in 2013, **reportedly to prevent computer leaks**. Separately, the logistics industry has long relied on paperwork for documenting the transit of goods, leading to hefty paper trails and sometimes inefficient processing. **Although that is starting to change**, it's notoriously difficult to do away with paper records in this sector, **say industry insiders**. The healthcare industry has a historical reliance on paper too. From prescriptions to **hospital documentation**, paper has persevered well into the 21st Century. To take one example, the majority of care homes in south-east Scotland still use paper-based management systems, **according to a study published last year**. Even when hospitals switch to digital, they may be faced with the burden of **storing historical paper documents relating to patient care off-site**. Within the European Union there are **11 countries that still use paper for medical prescriptions** rather than digital systems. In the US, paper stubbornly remains in use in some parts of the healthcare system despite attempts to modernise – **96% of hospitals and 78% of physicians** were found to use electronic health records in 2021.

Paper is still considered the backup medium whenever electronic systems fail – which, naturally, they do. In the aftermath of **a cyber-attack on a small Alaskan community** in 2018, municipal staff quickly switched to paper forms and typewriters when their computers went offline. Even Wikipedia, a gigantic online resource continually updated and edited by people all around the world, has **an emergency plan called the "Terminal Event Management Policy"**. During some potential future apocalyptic turn of events such as "imminent societal collapse" or "an imminent extinction level event", Wikipedia's millions of editors would be tasked with printing out various pages of the online encyclopaedia for posterity – because paper, ultimately, is considered reliable …

2023, October 3:

How records are made

Can a map of the ocean floor be crowd sourced?

… Cloaked in ocean, the seafloor has resisted human exploration for centuries … We have only just begun to map, much less explore, this enormous subsea world … In 2023, Seabed 2030 announced that its latest map of the entire seafloor is **nearly 25% complete**. The data to make the world's first publicly available map is stored at the International Hydrography Organization (IHO)'s **Data Centre for Digital Bathymetry (DCDB)** in a government building in Boulder, Colorado.

So far, the DCDB holds over 40 compressed terabytes of seafloor data. The biggest contributor is the US academic fleet: **17 research vessels** owned by American universities which constantly circle the globe studying the deep ocean. Other contributors include the National Oceanic and Atmospheric Administration (NOAA) fleet, the Geological Survey of

Ireland, and Germany's Federal Maritime and Hydrographic Agency. The biggest users are scientists all over the world who rely on the data to conduct research.

Seabed 2030 has made extraordinary progress by asking countries and corporations to share maps with the DCDB. But unfortunately, the map is not growing quickly enough. Between 2016 and 2021, the map leapfrogged from 6% to 20%. Since then, the pace has slowed. In 2022, it reached just 23.3% complete; in 2023, 24.9%. The ocean mappers came up with a new plan: crowdsourcing. "Crowdsourced bathymetry came about a few years ago when the IHO was saying: 'At this rate, we're never going to map the whole darn ocean; we need to start looking outside the box,'" says **Jennifer Jencks**, the director of the DCDB and the chair of a crowdsourced working group at the IHO …

2021, February 5: Record vs reality

There's a nice piece in the *Daily Mail* about disputed/uncertain borders involving SA, Vic, and NSW.

> … NSW and Victoria are divided by the winding route of the Murray River with the South Australia border running perpendicular. That imaginary line was supposed to cut straight through, but instead ended up sliding west and north [beginning] with a simple '141 degrees mistake' in 1847 … The boundaries were drawn 3.35km north and 2.96km south, incorrectly giving Victoria and extra 1,420sqkm of land … {A} 64-year-old conflict came to a head, when the SA Government announced they would send in their own surveyors to subdivide the land. Victoria saw this as an act of war and said they would be treated as trespassers and arrested. So the states took their battle to the High Court, which ruled in favour of Victoria …
>
> … The borders of NSW, SA and VIC now connect via the 11km of the Murray River, but still aren't crystal clear. Victoria agreed that the entirety of the river was in NSW, defining the border with NSW from the river's southern bank … still it's ambiguous which state has sole control over the Murray River beyond the southern border, making it a legal grey area for authorities.

If you're tickled by this sort of thing, I recommend *Off the Map - Lost Spaces, Invisible Cities, Forgotten Islands, Feral Places, and What They Tell Us About the World* by Alastair Bonnett. The sub-title (Oxford comma and all) tells you everything you need to know.

Amongst the almost fifty oddities collected in this book are –

- **Sandy Island**, 700 miles off the coast of Qld - on the maps but not actually there. How many other places shown on maps aren't there? How many places not shown on maps are there?
- **New Moore**, an island that came and went in the Bay of Bengal. Similar to the Pumice and Trash Islands that drift about. And ice islands that come and go.
- The **Aralqum Desert**, that used to be the Aral Sea - 'nuf said.
- **Zheleznogorsk**, a "closed city" missing from many maps where the Soviets produced weapons-grade uranium. Bit like Pripyat, abandoned after the melt-down in nearby Chernobyl.
- **Bir Tawil**, 795 sq. miles of desert bordered by Egypt and Sudan, neither of which, for bizarre and convoluted reasons, wants to claim it (accepting sovereignty would undermine their claim for more valuable, oil-producing land elsewhere).
- **Twayil Abu Jarwal**, a Bedouin village in the Negev Desert controlled by Israel but treated by the Israelis, like dozens of others, as a non-place. Likely to share the fate of Aghdam, obliterated in the war between Azerbaijan and Armenia.
- **Wittenoom**, the WA asbestos mining town – not just abandoned but expunged. A similar fate, but for different reasons, to Kangbashi, a Chinese ghost town (a "conjured landscape") – built but never occupied. Just like Kijong-dong, the fake North Korean city (the "Peace Village") near the border with South Korea.
- **Camp Zeist**, a Dutch military facility that was brought briefly under Scottish nationality to enable the trial to take place of the Lockerbie accused under Scottish law.

🏠 <u>MUCH ADO ABOUT DESCRIPTION</u>

- The **Geneva Freeport**, where high-value goods are kept and traded free of customs duties of any kind. It is a mediaeval commercial concept alive and well in the modern world.
- **Baarle**, a town on the Belgium/Netherlands border where you can never be quite sure which country you're in. Like the Chitmahals, neglected enclaves on the India/Bangladesh border.
- The **LAX parking lot** where air crew lay over between flights in trailers.

Aralqum Desert

Bir Tawil

Wittenoom

Camp Zeist

To say nothing of places from which persecution refugees have become disconnected
e.g. Uighurs, Rohingya, Palestinians
just three contemporary horrors to add to a long, sad list. If you want your mind blown a little, read the chapters on international airspace and gutter spaces. And then, of course, there's the Spratlys.

<p align="center">☙☙☙☙☙☙☙</p>

The DAD: Daddy-of-All-Descriptions

2018, November 15: Mother of all kilograms

Interesting take on accuracy and authenticity:

> ... Scientists from more than 60 countries will vote on Friday on whether a lump of metal held in a Parisian vault should continue to be the definition of a kilogram. Le Grand K, a small cylinder of titanium alloy, has set the standard since 1889. All the scales in the world are ultimately calibrated against it ... The problem, though, is that while the mother of all kilograms has only been taken out of its protective case four times in the last century, it has lost atoms and therefore mass. It amounts to just 20 billionths of a gram, about the weight of an eyelash, but in a world that needs to weigh objects with ever greater accuracy, that's a big deal ... Scientists ... are now part of the global effort to devise a more accurate, immutable
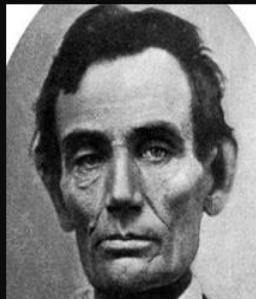
definition of a kilogram that is no longer dependent on a physical object. They are using what is known as a Kibble Balance … to express the mass of a kilogram in terms of the amount of upward electromagnetic force is needed to balance the downward drag of gravity. Then with some heavy-duty maths, they relate that to a fundamental physical law of nature.

2018, November 16:

<< Andrew Waugh: It's the last of the seven fundamental measures to be redefined in terms of absolute characteristics. So it is a really big deal. (The other six are length (metre), time (second), temperature (Kelvin), amount of substance (mole), electric current (Ampere), and luminos intensity (Candela).) Incidentally, Le Grand K cannot change its weight. The weight of Le Grand K *is* one kilogram by definition. If it loses a couple of atoms, then everything else gets slightly heavier.>>

An interesting inter-play between absolute truth (the definition of a kilogram) and qualified truth (the actual mass of Le Grand K). Puts me in mind of an anecdote told by Abraham Lincoln in a speech to the Wisconsin Agricultural Society (1859):

> It is said an Eastern monarch once charged his wise men to invent him a sentence, to be ever in view, and which should be true and appropriate in all times and situations. They presented him the words: "And this, too, shall pass away." How much it expresses! How chastening in the hour of pride! -- how consoling in the depths of affliction!

But context, as we r/keepers know, is all. 1+1=2. But one pile of sludge added to another pile of different volume just produces one big pile of sludge that is not even twice the volume of either of the two you started with. The joke here is that Lincoln's speech proceeds to express the hope that the American Union will not succumb to this universal truth:

> And yet let us hope it is not quite true. Let us hope, rather, that by the best cultivation of the physical world, beneath and around us; and the intellectual and moral world within us, we shall secure an individual, social, and political prosperity and happiness, whose course shall be onward and upward, and which, while the earth endures, shall not pass away.

- foreshadowing the words of a later (and greater) speech: "...shall not perish from the earth...".

2020, November 18:

Still on the theme of accuracy and authenticity, I have long wanted to write something about the way we describe records (not about the more obvious theme of how we perceive and understand them but about the prose we use). Controlled vocabularies used in taxonomies are more forgiving of weasel words (though the paucity and poverty of scope notes is still a worry) but we continue to employ joined up prose to describe records and recordkeeping. So, maybe there's still time for a useful debate about prose standards. An article in the *Guardian* dealing with word-extinction has some relevant things to say.

> … when I read last week of Edward Allhusen, a writer who has gathered together in a book called *Betrumped* (*v* to cheat or deceive) some 600 English words he fears are shortly to become extinct, something about it spoke to me – and not only because "lickspittle" is a term

I use quite often … Perhaps it was the politicians with their reliance on such emptinesses as "going forward" and "a deal that delivers" … Perhaps it was the highly intelligent people I heard last week fall back on the dreaded "journey" to describe something that was not even close to being one … I experienced a sudden and powerful yearning for the novel and the lively, the particular and the pungent. All around us is an enervating linguistic blandness. It comprises an awful lot of repetition, a certain amount of misuse, and a drastic devaluation when it comes to words that should have huge and grave import. These politicians who talked of "catastrophe" long before any iceberg was in sight. Why didn't they save it up? … People don't listen to politicians not only because they're inclined to dissemble. It's because we can no longer *hear* them; their vague, insipid words, as pale and lifeless as primroses kept too long out of water, simply don't catch on the ear … When language is precise, chosen with care and used with relish, it rings like a bell. We all know this, and many of us have a great hunger for it – an appetite that only swells as the word pail grows ever dustier …

Many years ago, I made somewhat similar remarks on the decline that had taken place in the prose being employed in source documents we use in preparing descriptions:

The first problem we face is that the object of our study is not scientific phenomena operating according to the "laws" of nature but products of the human mind and the political process. It may be pleasing to reduce human endeavour to the scale of plants and microbes but we are warned against carrying the analogy too far. On the other hand, we may take comfort from the observation that in nature's complexity, too, unqualified predictive laws can rarely be applied without allowing for numerous exceptions. Any methodology of function analysis must similarly allow for the illogicality, confusion, and obfuscations in human thought and behaviour. Even the documentary evidence of these obscure truths is now debased. Compare the noble clarity of nineteenth century administrative prose with this recent example from which only three identifying phrases (8 words) have been omitted –

> *The Department of …… has the objective of developing as a responsive, responsible, effective and efficient organisation implementing Government policy for maximising long-term economic development for Victoria through …… in ways that are consistent with sustainable and efficient use of resources and equitably meeting the priority needs of Victorians in …….*

Deriving any useful idea of purpose, leave alone discrete function, from this verbal sludge is quite impossible. "What, if anything, is a function" p.4.

PS. The Department of … in the quotation above was, in fact, the Victorian D. of Agriculture.

2024, May 6: More Books

And so, despite the rain, to Sydney once more. This time to sign the latest (and possibly last) version of my Will. Spent happy hours at the SMSA Library, averting my eyes from the cupcake shop across the street. Then I dropped $360 at Abbey's and that was **after** they took $40 off the top (frequent reader points from my last visit). Haven't looked at them all yet, but two that I did scan on the train home may be of interest to others.

Geoffrey Robertson (KC, he would want me to say that) *Lawfare: How Russians, the Rich and Government Try to Prevent Free Speech* (2023) remaindered at $12.00.

For centuries the law of defamation has worked to cover up misbehaviour by the rich and powerful. Now, through misguided judicial development of the laws of privacy, breach of confidence and data protection, a new terror has been added to suppress the reporting of truths of public importance … ["Lawfare"] came into vogue in Britain in 2022 as a description of the work of 'reputation lawyers' who had been issuing threats and writs against authors and publishers of books about Russian oligarchs … The most noticeable victim was … Catherine Belton, author of *Putin's People: How the KGB Took Back Russia and Then Took on the West*, which attracted a sudden blizzard of legal actions … There had been preliminary skirmishes before the case was settled, at a cost to HarperCollins of £1.5million … and a cost to Belton of a year of stress and exhaustion … over the previous few years books had gone unwritten, or had been censored or simply not published , for fear of defamation actions about statements reasonably believed to be true …

🏠 MUCH ADO ABOUT DESCRIPTION

The parallel with the intimidatory tactics of ACMA and the eSafety Commissioner are obvious and, if you want topicality, I refer you to the misguided (and ultimately unsuccessful) attempt by ASA to suppress a conference paper of mine on the spurious claim it was defamatory. When I went "up" to Sydney University in 1964, the SU Union was trying to emulate the Oxford Union with stars like Michael Kirby, Geoff Robertson, and Richard Walsh dazzling their peers in the debates. For a gauche, working-class lad like me, burdened with a mediocre Catholic high school education and the urbanity of an Echinoderm, it was awesome. I couldn't work out how or from where these guys got their sophistication. I concluded they were so far ahead of me that I would never catch up. I was in my forties before I became that glib.

| Michael Kirby | Geoff Robertson | Richard Walsh |

John Elledge *A History of the World in 47 Borders: The Stories Behind the Lines on Our Maps* (2024). This book sparkles with glittering inanities on every page. I have long urged the development of a "descriptive datum", in the service of federated searching of archival resources, at least for Australasia. These suggestions have been met with bone-jarring indifference from the profession. None the less, I continue to nourish my passion for primes.

In 2021, pollster YouGov conducted a survey to find out how famous the forty-five men who had preceded Joe Biden as president of the United States currently were. Chester Arthur came forty-fourth … The only *less* famous was Franklin Pierce … This is a shame, really, because President Arthur shaped the modern world far more … than any of those other forgotten men … In the autumn of 1884, his government invited representatives of all nations … to Washington, 'for the purpose of fixing upon a meridian … to be employed as a common zero of longitude' … In other words, Chester Arthur invented the prime meridian and everything that flows from it. Did Millard Fillmore do that? No, he did not.

Although *the* prime meridian was not agreed upon until 1884, the concept of *a* prime meridian has been around for rather longer. The story begins in Egypt in the third century BCE with one of those infuriatingly impressive polymaths that the ancient world seemed to specialise in. In his eight and a bit decades on the planet, Eratosthenes of Cyrene managed to: become one of the greatest poets of his age; get a job running the library at Alexandria; found the scientific discipline of chronology … create a system for finding prime numbers; calculate the diameter of the sun; *and* examine the mathematical basis of music … At some point, you're surely just showing off … he also [devised] the system of grid lines we know as longitude and latitude … making it possible to determine where distant places were in reference to each other … by comparing noon shadows cast in two places on the same longitude and doing some clever things with trigonometry, he could get agonisingly close to calculating the size of the earth [a feat which eluded Columbus nearly two millennia later].

The reason this was so clever, in a 'Well done, Eratosthenes, you've changed the world *again*' kind of a way, is because … the two halves of the system actually had to work in subtly different ways … lines of latitude run parallel, those of longitude … converge at the poles … latitude has a clear starting point in the equator … There's no naturally existing line that divides the Earth into eastern and western hemispheres … Eratosthenes got around this

in the same way everyone who came after him would: by simply picking somewhere convenient to him and decreeing that the prime meridian …

... in October 1884 … forty-one delegates from twenty-six countries … picked a line that ran through the capital city of just one of their number … to ensure the accuracy of the required measurements, the international scientific community had already agreed that the new meridian should 'pass through an astronomical observatory of the first order'. That effectively limited the options to … Paris, Berlin, Washington and Greenwich. The dingy London suburb won … To mark the meridian which it gave to the world, you'll find a brass strip set in the ground; the museum's website and social media encourage visitors to take a selfie with one foot on either side … By night, a green laser fires out towards the Thames to mark the path of the famous line.

These are all complete and total lies. The prime meridian isn't there at all and providing you have a smartphone you can prove it. Check your maps app … and you'll find yourself at 0.0015 degrees west … approximately 102 metres into the western hemisphere. To find the *actual* prime meridian, you need to move slightly to the east [where] you'll find a bin, suggesting itself as a repository for any dog faeces you happen to have about your person … All this has been clear since the 1970s … Faced with the choice between redoing the entire global longitude system … or just redefining the meridian, the global scientific community opted to do the latter …
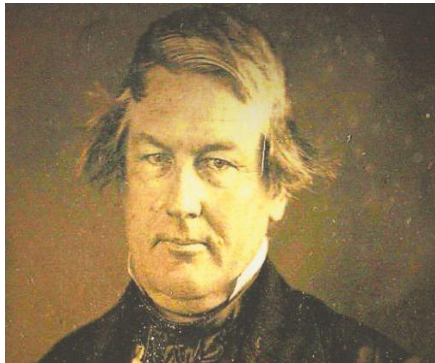
Chester Arthur   **Millard Fillmore**   **Eratosthenes**
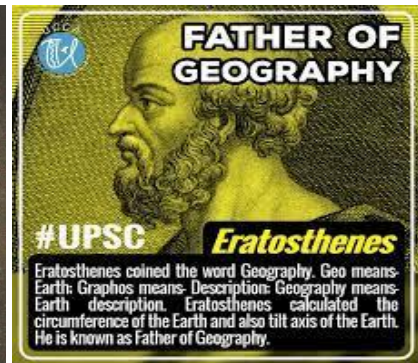
<< Andrew Waugh:

These are all complete and total lies. The prime meridian isn't there at all and providing you have a smartphone you can prove it. Check your maps app … and you'll find yourself at 0.0015 degrees west … approximately 102 metres into the western hemisphere. To find the *actual* prime meridian, you need to move slightly to the east [where] you'll find a bin, suggesting itself as a repository for any dog faeces you happen to have about your person … All this has been clear since the 1970s … Faced with the choice between redoing the entire global longitude system … or just redefining the meridian, the global scientific community opted to do the latter …

From memory, the centre of mass of the earth is slightly in the wrong place, meaning the vertical marking the meridian comes out slightly out slightly off where they thought it did… The history of Datums (how they assign latitude and longitude values to particular points on the earth) and projections is endlessly fascinating as they get more and more accurate in measuring the lumpy shape that is the earth. And then you mix in the fact that the continents are moving fast enough to throw off coordinate systems - Australia moves about 7 cm a year. Here's an XKCD… >>

<<And then you mix in the fact that the continents
are moving fast enough to throw off coordinate systems >>

No faster than the vagaries that impede archival descriptions from different hands aligning with each other, he could not resist saying.

As I've remarked in another place, my very favourite datum is in Macquarie Place Park, just south of Circular Quay. I visit it often. Not when the weather's like this, though. The

small stone obelisk has black letters that someone cares enough about to keep freshly painted to mark the place from which all public roads in the Colony were to be measured:

> The Obelisk operated as the "zero point" for measuring the distance of roads from Sydney from 1818. It played a central role in the subsequent surveying, mapping and planned expansion of the Colony from the early 1800s beyond the current extent of New South Wales. Surveyors measured and laid out the line of many roads. Distances in the County of Cumberland were measured from this Obelisk. Those distances were also recorded at the side of the road on milestones or other distance markers. Road plans prepared by surveyors show these distances as well. Public works officers and workers were responsible for forming and making the roads, but it was the surveyors who laid them out, thus providing a direct link to the Macquarie Place Obelisk.

Joanna Sassoon:

Its an important point Chris. Here is the zero marker in Perth. It has moved from the side of the building to the pavement and undergone several design changes. So, perhaps it is no as longer accurate as when it was on the building rather to its side, but it's interesting.
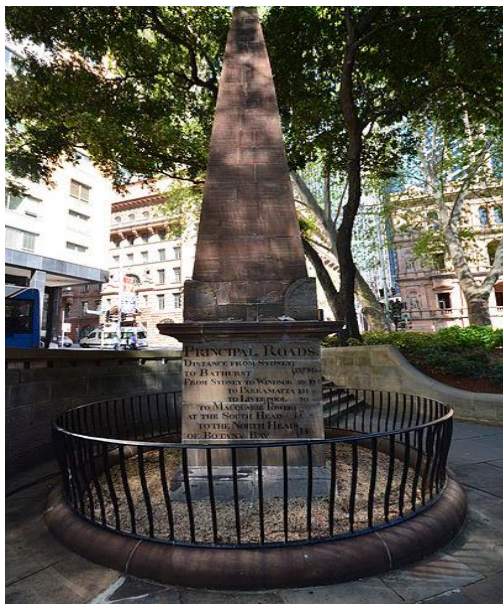
Point Zero, Sydney

Point Zero, Perth

Chris Hurley:

And just to confuse things, here is an extract from a report for the City of Sydney by Casey & Lowe (Archaeology & Heritage Consultants):

> **3.25 The Obelisk as Marker**
> An issue for the project was to determine if the obelisk is still used as the zero point for the measurement of roads in New South Wales. There is information to suggest there was a change of measuring point in the nineteenth century from the obelisk to the GPO.
> The 'Addendum' by B T Dowd puts together information about this change, explaining why the location of some of the milestones was further along the Parramatta Road than it should have been, if they had been measured from the obelisk.
> On 11 September 1846, the Deputy Surveyor-General S. A. Perry wrote to the Colonial Secretary suggesting that the milestones on the Liverpool and Parramatta Roads needed replacing because they were so mutilated and asked from what point the measurement should commence - from Macquarie Place, the Post Office or the boundary stone of the city? The reply from the Colonial Secretary's office dated 23 September 1846 was that the Governor approved of the expense of new milestones 'the distance being measured from the General Post Office in George Street which appears to His Excellency to be a central and therefore

eligible point of commencement'. Tenders were called for the new milestones in December 1846.

Dowd also makes reference to a letter dated 11 November 1839 from the Postmaster-General James Raymond to the Surveyor-General suggesting that the GPO be used as commencing point for the measurement of the Penrith and Windsor Roads. In December 1839 the Surveyor-General 'issued instructions to one of the assistant surveyors to mark out the places for the milestones, with the zero point at the centre of the Post Office building in George Street'. Dowd says that he has not been able to trace any record of the completion of this proposal. He suggests that because it was decided in 1846 to use the GPO, the earlier scheme probably did not eventuate.

Consultation with the RTA on this issue to understand current practice in relation to the obelisk has revealed that it is a designated survey mark and is still used as the zero point marker for major main roads as New South Wales including the Great Western, Princes and Pacific Highways. Minor roads are measured from the GPO. This marker has been an important point for the measurement for roads from 1818 to 2003 after 185 years.

When I was in Melbourne, I asked about the Victorian datum but no one seemed to know. The suggestion was made that it was the spire on the clock tower of the Melbourne GPO. As Andrew suggests, this stuff isn't straight forward and, for those of us with that kind of mind, it is endlessly fascinating.

I love a story about the building of the Opera House. When the site was opened, a stone was laid where the Premier & Utzon cut the ribbon and this stone, it was announced, was the datum for the building. After years of delay and cost over-runs, Utzon was sacked and new architects and builders took over the half completed structure. The first thing they asked to see was the foundation stone. They were told it had been getting in the way and that it had been moved into a nearby storage shed. They were horrified until they were told that the builders had blasted a rivet into the concrete apron and had been using that.

2024, May 7:

In case you thought all this was just nerdy, musty old stuff (rather than the endlessly fascinating thing that it is) how's this from Land Vic :

Australia is in the process of modernising the Australian Geospatial Reference System (AGRS) to support advancements in precise positioning and enhance alignment with underlying spatial information. This national program is being delivered through the Intergovernmental Committee on Surveying and Mapping (ICSM) and involves updating fundamental components of the national coordinate reference system:

Victorian Railways has its own approach:

These days track distances are measure from one point, just south of Spencer Street. Prior to conversion to metric measurements in the 1970s (Weekly Notice 21/73 lists that "All traffic operations will be metric" from the 1st July 1973), the was three track datum points.
Datum A - Just South of Spencer Street, this is also the datum point for all kilometre distance measurements. This datum applied to all lines "Western", ie heading North away from Spencer Street.
Datum B - This point was opposite the end of Elizabeth Street at Flinders Street Station. It was 59 chains and 52 links from Datum Point A. This datum applied to the St.Kilda and Port Melbourne Lines.
Datum C - This point was between Elizabeth and Degraves Streets at Flinders Street Station. It was 61 chains and 23 links from Datum Point A. This datum applied to all lines "Eastern" ie heading East away from Flinders Street/Princess Bridge.

But they appear to be a bit grammatically challenged.

And in three dimensions already

On 5 May 1971, Geoscience Australia, on behalf of the National Mapping Council of Australia, carried out a simultaneous adjustment of 97 230 kilometres of two-way levelling. Mean sea

level for 1966-1968 was assigned a value of 0.000m on the Australian Height Datum (AHD) at 30 tide gauges around the coast of the Australian continent. The resulting datum surface, with minor modifications in two metropolitan areas, has been termed the Australian Height Datum and was adopted by the National Mapping Council at its twenty-ninth meeting in May 1971 as the datum to which all vertical control for mapping is to be referred. The datum surface is that which passes through mean sea level at the 30 tide gauges and through points at zero AHD height vertically below the other basic junction points.

A useful essay on the AHD can be found in *Wikipedia*

> The **Australian Height Datum** was introduced in 1971 as the official vertical datum for Australia, and thereby serves as the benchmark to which all height measurements are referred. The Australian Height Datum is an amalgamation of decades of spirit levelling work conducted by numerous state and territory authorities across the country, and was corrected to align with the mean sea level observations of thirty tide gauges positioned around the entire coastline. While it remains the published vertical datum for all surveying and engineering operations performed throughout Australia, newer technologies have uncovered numerous deficiencies, offsets and distortions within the Australian Height Datum, leading to discussions about defining a new Australian vertical datum.

Now, what price the DAD (Descriptive Archival Datum)? The daddy of all descriptions. Something I once termed THE BIG ONE (viz. everything) – a special instance of the URO (Universal Recordkeeping Super-Type) . What a concept. Not going to happen, is it.

2024 May 9:

> **<<… what price the DAD (Descriptive Archival Datum)? The daddy of all descriptions. Something I once termed THE BIG ONE (viz. everything) – a special instance of the URO (Universal Recordkeeping Super-Type) . What a concept. Not going to happen, is it.>>**

They tell me the reason it's not going to happen is that it's too theoretical to be of any practical use.

Pish!

It's essentially the Home Page of the *Modest Proposal*. I could write it myself in five minutes – probably have done somewhere in my writings. The problem does not lie in the drafting of it. The question is whether it ever could or would be used.

Every description has **a** DAD, but it's not something that most descriptive archivists acknowledge or are even aware of. It's the orientation of the archivist's mind, the point of view, the slant, the bias, call it what you will. The archivist is like Eratosthenes "simply picking somewhere convenient to him and decreeing that the prime meridian", even if he isn't aware that he's doing it (or she). But we don't have **the** DAD because there's been no archival equivalent of Chester Arthur.

PS. Less Books: I am trying to get rid of some of mine. I have been trying to unload my duplicates. I have hundreds of them. If you see a title that's interesting, and you buy a lot of books, it's likely you buy the same one twice if you forget you've already got it (a thing that happens more often as you get older - and you have more books). The moment you walk out of the store, they're worth next to nothing. Actually, that's exactly what most second-hand dealers will offer - nothing. Anyone else in this situation may be interested in the Book Grocer's Sell Your Books page. They don't offer much but I got $600 for some of mine recently and that's better than chucking them out. Meanwhile, is anyone interested in a free copy of *Archives: Recordkeeping in Society* in Japanese? Not even the Book Grocer would give me anything for that, I imagine. They kindly sent me a copy because my chapter is

included in the translation. It's probably impolite to give it away "but it's an unjust world, and virtue is triumphant only in theatrical performances."

Alan Ventress:

I went through the same process when I retired from paid work in 2012. I had collected an enormous library of Australian history books which I kept at work. There was no room to store them at home so I decided to give them away. Quite a few were taken by historians, for example, a complete set of the Australian Dictionary of Biography and the 1958 Alec Chisholm edition of the Australian Encyclopedia. The remainder I gave to Peter Tinslay at Antique Books and Curios, Crows Nest and to Dalwood Homes for their regular fetes. I still have a large collection of books in my office/library but will leave those for my children to throw away upon my demise. (photo attached) I stopped buying books about 10 years ago.

<<the 1958 Alec Chisholm edition of the Australian Encyclopedia>>

This is an excellent reference work, far exceeding in scholarship the other editions. A few decades ago, they were throwing them away. I have two sets, one of which was given to me by Alan Ives (a bibliographical bower bird of epic stature) and the other I picked up myself for a few dollars. They are now selling online for $hundreds on eBay, e.g.

- Second edition $499 AU
- 1963 printing $140 AU
- 1st edition (1858 sic) $520 AU

Some of these asking prices seem a little high. So far as I can tell, it is not digitised and available online. So, you have to possess one or go to a library. But many libraries will have thrown them out. I believe that some "historical" encyclopaedias can be very valuable (intellectually, not monetarily) long after that part of the content that dates quickly has been overtaken. I also have a copy of EB 1911, generally acknowledged (along with EB 1929) to be a monument to scholarship (albeit outdated) - everything explained that is explainable.

> scholars and lovers of good prose still laud the 1911 edition, which has articles by famous scholars and writers Edmund Gosse, J. B. Bury, Algernon Charles Swinburne, John Muir, and Bertrand Russell … And it has a certain snob appeal. I LOVE the prose of the early 20th century. The essay on Virgil is more vivid than clumsily-written articles about "the ploughshare in the *Georgics"* and other abstruse subjects published in classical journals.

J B Bury is one of my personal favourites and I have several of his books. My executors will have to dump everything. Isn't that a sad thought?

🏠 MUCH ADO ABOUT DESCRIPTION

What are researchers to do who want to read these things in future if they aren't digitised and online? My only comfort is that I won't be around to find out and, if the *Wikipedia* author (below) is correct, it won't much matter anyway.

Wikipedia says that EB 1911

> is readily available on the Internet. Its use in modern scholarship and as a reliable source has been deemed problematic due to the outdated nature of some of its content. Modern scholars have deemed some articles as cultural artifacts of the 19th and early 20th centuries. Additionally, the 11th edition has retained considerable value as a time capsule of scientific and historical information, as well as scholarly attitudes of the era immediately preceding World War I.
>
> Sir Kenneth Clark, in *Another Part of the Wood* (1974), wrote of the eleventh edition, "One leaps from one subject to another, fascinated as much by the play of mind and the idiosyncrasies of their authors as by the facts and dates. It must be the last encyclopaedia in the tradition of Diderot which assumes that information can be made memorable only when it is slightly coloured by prejudice. When T. S. Eliot wrote 'Soul curled up on the window seat reading the *Encyclopædia Britannica*,' he was certainly thinking of the eleventh edition." (Clark refers to Eliot's 1929 poem "Animula".) It was one of Jorge Luis Borges's favourite works, and was a source of information and enjoyment for his entire working life.

After 1929, continuous revision made commentary on "editions" difficult. One of the banes of serving as reference librarian late at night at NLA in the early 70s, was fielding queries from harassed parents being hounded at home by encyclopaedia salesmen with limited time offers asking if you recommended that they buy it. We were forbidden from recommending anything but I always told them that they should consider it very carefully and they usually took this to mean they shouldn't.

2024, June 19:

Meanwhile, Paris Point Zero has apparently become inaccessible because of the Notre Dame fire. If I were still travelling internationally and I found myself in Paris, I would console myself with a visit to the Shakespeare and Company Bookshop (nearby) and to Musee Cluny to sit once again (and find a little peace) before the Lady and the Unicorn.

And stay at The Esmeralda as I once did when young. Some things, I'm glad to say, seem not to have changed. Migrant arrivals, on the other hand, have surged, they tell me –

- Overseas migration 2022-23 – net annual gain of 518,000 people
- Migrant arrivals increased 73% to 737,000 from 427,000 arrivals a year ago
- Largest group of migrant arrivals was temporary visa holders with 554,000 people
- Migrant departures decreased 2% to 219,000 from 223,000 departures a year ago.

And I think most of them have settled on the Central Coast. It's getting crowded up here. The two 5-minute peak hours we used to have are now submerged in heavy traffic most of the day, the kind of thing I came here to get away from.

2025, January 23:

<< 1 June, 2024: … Descriptively, topography is an important dimension … We speak of European and Indigenous. In view of our increasingly diverse ethnic makeup, it might be more accurate to refer to New-Comers and Old-Timers … I imagine many of our archives nowadays incorporate Old-Timers' place names … alongside those of the New-Comers … If ever there was an argument for a DAD to sort out at least this part of the landscape, surely this is it. PS One problem with historical place names is whether or not to treat them as timeless or time-bound … >>

<<6 May, 2024: … in October 1884 … forty-one delegates from twenty-six countries … picked a line that ran through the capital city of just one of their number … the international scientific community had already agreed that the new meridian should 'pass through an astronomical observatory of the first order'. That effectively limited the options to … Paris, Berlin, Washington and Greenwich. The dingy London suburb won … [But] The prime meridian isn't there at all … To find the *actual* prime meridian, you need to move slightly to the east [where] you'll find a bin, suggesting itself as a repository for any dog faeces you happen to have about your person … >>

Fast forward to 2025 and it seems that international standardisation of place names, in the news now because Trump II has renamed the Gulf of Mexico, is just a matter for co-operation and goodwill:

> The Gulf of Mexico doesn't belong to just one country — but the majority of it is divided up between Mexico and the US … There are a bunch of treaties about the division of the Gulf of Mexico dating back decades, which the US and Mexico have ratified … The executive order only changes how the US refers to the basin — but that doesn't mean the rest of the world has to follow suit … The International Hydrographic Organization (IHO) is a global body that works to ensure all the world's oceans are surveyed and charted uniformly, trying to make sure all nautical charts are the same : "*It is recommended that where two or more countries share a given geographical feature (such as, for example, a bay, strait, channel or archipelago) under a different name form, they should endeavour to reach agreement on fixing a single name for the feature concerned*" … The IHO also says it should cooperate with the United Nations Conferences on Geographical Names with respect to the standardising of names.

Even within Australia, there is no over-riding jurisdiction over how places are named - just "co-ordination" and "co-operation" between responsible authorities. Happily, in our remote corner of the world, this is likely to be done amicably and without too much danger of interference from foreign bumkins. I would note, however, that a New Zealander standing on the west coast will look out over the Tasman Sea while an Australian looking back at him from our east coast is likely to say he is gazing out over the Pacific Ocean.

<<Andrew Waugh: LOL. The quoted portion of the executive order is very careful to only rename the portion of the Gulf of Mexico within US territory, despite any impression that you may have gained from the announcement or the media.>>

I wonder if Trump himself read the fine print before signing.

2025, January 31:

<< The quoted portion of the executive order is very careful to only rename the portion of the Gulf of Mexico within US territory …>>

Up to the limit

> Mexican President Claudia Sheinbaum has written a letter to Google asking the firm to reconsider its decision to rename the Gulf of Mexico … it will only appear on Google Maps with the new name for people based in the US - elsewhere in the world it will retain its current name, which has been used for hundreds of years. There is no international organisation responsible for the naming of bodies of water. But ***Mexico argues the U.S. cannot***

legally change the Gulf's name because the United Nations Convention on the Law of the Sea dictates that an individual country's sovereign territory only extends up to 12 nautical miles out from the coastline …

Google has not yet responded to the BBC's request for comment. But in a statement on social media on Monday it said: "We have a longstanding practice of applying name changes when they have been updated in official government sources." It will also rename Mount Denali as Mount McKinley in the US, following another order from Trump. "When official names vary between countries, Maps users see their official local name," it said …

There has long been fuss over place names that are considered unusual (e.g. Fugging) and, more specifically, ones that are considered offensive (e.g. Black Gin Creek). And then there's ones that are disputed (e.g. Taiwan/China; Israel/Palestine). I guess the Gulf of Mexico is on the cusp of becoming one of those.

The name of a Perth suburb (Innaloo) has been the butt of jokes and unsuccessful attempts to change it.

PS. Trump's decision re Mt McKinley is actually just one more round in a long-running dispute.

<p align="center">ॐॐॐॐॐॐॐ</p>

2020, December 11: What, if anything, is a handbag?

Oddspot: Included in a "handbag exhibition" at V&A is Winston Churchill's despatch box. Rucksacks, too, apparently qualify. But, on closer examination, the exhibition's title (Bags: Inside Out) reveals a wider scope and a curatorial subtlety that the *Guardian's* sub-editors found to be resistible. Perhaps because some things like handbags, cod-pieces, nappies, etc., correctness notwithstanding, just seem to be intrinsically funny (like Boris Johnson).

2022, May 27: New to My Website

These are some odds-n-sods I have discovered while cleaning out my hard drive. Can't be sure they haven't been incorporated somewhere in longer essays already uploaded. Duplication can't do any harm I suppose.

Electronic Series (2002)

Why I wrote this or for whom it was intended escapes me now but it doesn't appear anywhere else on this website so far as I can tell (although I fear it may be buried somewhere in one of the longer essays). Its matter is along the same lines (or, at least, in the same pumpkin patch) as The Hunting of the Snark: Searching for Digital Series (2011) so, although this piece was written several years earlier, I have attached it as an addendum to the later work. Thematically, it is also congruent with the multi-part Relationships in Records (2001–2004), alas unfinished, and my essay on the Series in the Encyclopedia of Archival Science (2015). Taken together, they open a door to a theme I always hoped to explore further and it is one of my major regrets that I never have.

The argument concerns the ethos of the record and, by extension, that which differentiates an archive from a collection or from toilet paper for that matter. An instant Document arises from a connection with an event or circumstance of which it is evidence – that is its purpose. It is capable of being put to other uses (as an historical artefact, for example, or a promotional gimmick) but that is not what a record is for. Putting it to another use may obscure the purpose for which it was created but it cannot change it. That shared purpose is what binds instant records together as a sequence or series – even if they are not kept together physically (cf. dockets and computer records) or even purposefully and even if as estrays the chain is broken. That shared purpose establishes a relationship between them and an intellectual structure that is a defining characteristic of the record.

But that structure (between instances) cannot be found solely in the relationships (based on a shared common purpose) that subsist between instant records. What also binds them together is the relationship that these instant records have with the accumulation (series/fonds) to which they belong – even in the curious case of Robinson Crusoe's Diary where the instance and the accumulation are one. It is a relationship that can be shared with no other artefact outside of the accumulation even if the content matter is the same (or even identical as in the case of a replica). A land title is a singular proof of ownership and the collective (register of titles) shares the same purpose consolidated by the authority and assigned responsibility of the "collector".

What makes a Series, then, is not the accumulation (collection) of like instances but the shared purpose subsisting between the instances and the accumulation. This is what differentiates an archive from a collection, the ethos or purpose of which is to bring together instances indiscriminately or on the basis of whim as to the common purpose shared between the instant records and the Series to which they belong, a recordkeeping purpose that is in operation before any process of collection begins. (Note: This is a distinction between the essence of the Archive and the Collection, not between the roles of Collectors and Archivists).

This contextuality (the purpose shared between the instant record and the collectivity) is further enhanced by relationships with Agents (Doers) and the Activity (Deed) that is the embodiment of that purpose. But the term "purpose" is misleading. Paradoxically, these seemingly contrived recordkeeping alignments are what the old books meant by "naturalness" which need not result from a purposeful intent at all on the part of the Doer – hence what I have called the accidental record. Note: The distinction between the "natural" archive and an "artificial" collection has been explored by Geoffrey Yeo in "The Conceptual Fonds and the Physical Collection" *Archivaria* 73 (2012) 43-80.

Scaleability (2002)

How could you give me life, and take from me all the inappreciable things that raise it from the state of conscious death?

Nearly two hundred years ago, Charles Dickens (one of my favourite authors) wrote a novel called Hard Times (1854). In it he devastatingly satirised regard for data without understanding:

"Now, what I want is Facts. Teach these boys and girls nothing but Facts. Facts alone are wanted in life. Plant nothing else, and root out everything else. You can only form the minds of reasoning animals upon Facts; nothing else will ever be of any service to them.

"Respect for Facts ("mere facts") would appear to be a hallmark of archival thinking but it is understanding we really need. Dickens "did not decry the wholesale usage of statistics … [but] how this information can be subjected to perversion and abuse, for purposes of subjugation and creating statistics that are class-biased".

> *[Bounderby] now stepped forth. A mighty man at cutting and drying, he was … in his way (and in most other people's too) a professed pugilist … He was certain to knock the wind out of common sense, and render that unlucky adversary deaf to the call of time. And he had it in charge from high authority to bring about the great public-office Millennium, when Commissioners should reign upon earth … "You are to be in all things regulated and governed … by fact. We hope to have, before long, a board of fact, composed of commissioners of fact, who will force the people to be a people of fact …"*

Dickens was <u>not against factual knowledge</u> but rather:

> *against statistics as a form of social knowledge, a way of knowing which necessarily constitutes the object of its knowledge - in this case the working class and their conditions of life - in particular ways and which thereby dictates particular approaches to it. It is statistics as what Michel Foucault would call a disciplinary technology of knowledge, as a mechanism for moral and political surveillance and restraint.*

I do not think it fanciful to liken this to our corrective idea that evidence and interpretation cannot easily be separated and that objective Truth can be abused. Indeed, that idea has become common place. In none of his books does Dickens look to Dogma, Institutions, or Systems for an answer. Not only are factories and trade unions portrayed as instruments of oppression but also <u>workhouses,</u> <u>schools,</u> <u>the law,</u> <u>bureaucracy,</u> <u>mobs</u>, the <u>weight of evidence,</u> <u>creeds,</u> <u>theories,</u> "<u>smelly little orthodoxies</u>", <u>families</u> even. The corrective lies not in an orthodoxy of our own but in diversity rather than homogenisation and that is where we too must find it, even if the orthodox hate us for it.

In simple terms, that means our view of the case must self-consciously encompass the multiplicity of change and perspective. The whole purpose of archival description is to make the facts being observed submit to an understanding of their meaning – a true understanding but one which may involve dissonance or contestation (the "<u>Grey Zone</u>"). The Grey Zone is <u>not a comfortable place</u> – disinformation and deception abound alongside the dialectic. Dogmatists want us to take sides:

> *… Dickens conveys that organized labour was so much self-deceiving agitation, which in passing squashed the rights of individuals … He knew that it was not so, for the above eye-witness account was his own, from his article "On Strike" … The more we find out what actually happened at that time, the more we realize that militancy was a lifeline – a well-spring of hope, a channel for popular energies, as well as an indispensable lever … if one tried to imagine the great industrial novel that never did get written, one might suggest that the masters cried out to be satirized, the mass of the people to be presented with clear-eyed realism.* (David Craig, *Introduction to the Penguin edition*).
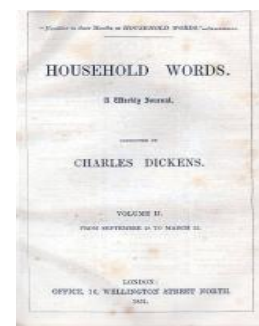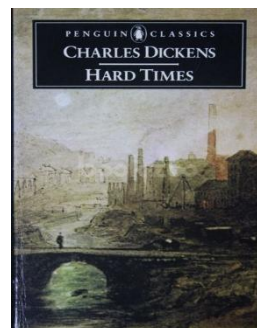
But that is not at all what Dickens was about. In the <u>article referred to</u> (relating to the Preston Strike of 1853/1854) he declined to choose: "*Masters right, or men right; masters wrong, or men wrong; both right, or both wrong; there is certain ruin to both in the continuance or frequent revival of this breach*" and he predictably concluded:

… into the relations between employers and the employed, as into all the relations in this life, there must enter … something of mutual explanation, forbearance and consideration … otherwise those relations are wrong and rotten to the core and will never bear sound fruit …

For all the ferocity of his life-long attacks on Dogmatism in all its forms and on the suffering it begets, Dickens longed for "an era of its being quite settled that the national dustmen have only to do with one another, and owe no duty to an Abstraction called a People …" (Hard Times, Book III, Ch.9).

The phenomena we describe (the entities) are not self-explanatory and how we portray and juxtapose them either illuminates or obscures their meaning (sometimes both) – never more so than when we show them standing in relationships with each other. This memorandum was prepared, at their request, for the ASA Descriptive Standards Committee twenty years ago. It was an early warning of the folly of building relationship data into the attributes assigned to descriptive entities.

> *I wouldn't say a word that could be reckoned as injurious,*
> *But to find a mother younger than her son is very curious,*
> *And that's the kind of mother that is usually spurious,*
> *Tara-diddle, tara-diddle, tol-lol-lay.*

My spelling of "scaleable" has been objected to. It is an allowable variant and the criticism (you will not be surprised to learn) has made me stubborn.

The Canonisation of Peter Scott (2019)

In preparation for a recent I-CHORA Conference in Melbourne, I was involved in a discussion amongst a group of Australian and New Zealand archivists (some young, some venerable) dedicated to articulating and handing on our shared understanding of what I continue to call (notwithstanding a certain amount of carping) the Australian ("Series") System. We felt a need to explain it better and to provide a springboard for further development by a new generation of archival thinkers. It was thought this could begin with conference papers and grow into a book, but no book has appeared so far as I know. My view was (and is) that we first need to establish a Canon rather like the 4th century Christians did when forming the New Testament – approving some things and discarding others. But what to approve, and what to discard, and who is to do it?

2022, June 1:

Re Scaleability

I am told that Dickens' approach to social problems is not "practical".

Let my old friend George Orwell respond to this:

... in the ordinarily accepted sense of the word, Dickens is not a 'revolutionary' writer. But his position here needs some defining. Whatever else Dickens may have been, he was not a hole-and-corner soul-saver, the kind of well-meaning idiot who thinks that the world will be perfect if you amend a few bylaws and abolish a few anomalies ... The truth is that Dickens's criticism of society is almost exclusively moral. Hence the utter lack of any constructive suggestion anywhere in his work ... There is no clear sign that he wants the existing order to be overthrown, or that he believes it would make very much difference if it were overthrown. For in reality his target is not so much society as 'human nature' ... His whole 'message' is one that at first glance looks like an enormous platitude: If men would behave decently the world would be decent ... It seems that in every attack Dickens makes upon society he is always pointing to a change of spirit rather than a change of structure. It is hopeless to try and pin him down to any definite remedy, still more to any political doctrine. His approach is always along the moral plane ... the strongest single impression one carries away from his books is that of a hatred of tyranny ... it is not at all certain that a merely moral criticism of society may not be just as 'revolutionary' — and revolution, after all, means turning things upside down as the politico-economic criticism which is fashionable at this moment. Blake was not a politician, but there is more understanding of the nature of capitalist society in a poem like '[I wander through each charted street]' than in three-quarters of Socialist literature. Progress is not an illusion, it happens, but it is slow and invariably disappointing ...

Consequently two viewpoints are always tenable. The one, how can you improve human nature until you have changed the system? The other, what is the use of changing the system before you have improved human nature? ... The central problem — how to prevent power from being abused — remains unsolved. Dickens, who had not the vision to see that private property is an obstructive nuisance, had the vision to see that. 'If men would behave decently the world would be decent' is not such a platitude as it sounds. ... A good-tempered antinomianism rather of Dickens's type is one of the marks of Western popular culture ... The ordinary people in the Western countries have never entered, mentally, into the world of 'realism' and power-politics. They may do so before long, in which case Dickens will be as out of date as the cab-horse. But in his own age and ours he has been popular chiefly because he was able to express in a comic, simplified and therefore memorable form the native decency of the common man ... Dickens voiced a code which was and on the whole still is believed in, even by people who violate it. It is difficult otherwise to explain why he could be both read by working people (a thing that has happened to no other novelist of his stature) and buried in Westminster Abbey ... in the case of Dickens I see a face that is not quite the face of Dickens's photographs, though it resembles it ... He is laughing, with a touch of anger in his laughter, but no triumph, no malignity. It is the face of a man who is always fighting against something, but who fights in the open and is not frightened, the face of a man who is generously angry — in other words, of a nineteenth-century liberal, a free intelligence, a type hated with equal hatred by all the smelly little orthodoxies which are now contending for our souls.

Grasp that and you have the reading. I do not recommend Dickens as a guide for dealing with social issues (though you could do a lot worse) but as an example of how to approach an understanding of them. A moral sense is indispensable to the observation and description of facts but meaning remains elusive. Multiplicity, of which scaleability is merely one aspect, is how we archivists are able to perceive, if we choose, two or more "tenable viewpoints" – which is just another way of reaching out for objectivity. Which is the more "realistic" basis for action? Conclusively adjusting a society corrupted by human frailty, employing hard facts combined with a belief in perfectibility to do so, or provisionally treating society's problems and well-meaning efforts to overcome them as both being equally corrupted by flaws in human nature?

"When life itself seems lunatic, who knows where madness lies? Perhaps to be too practical is madness. To surrender dreams — this may be madness. Too much sanity may be madness — and maddest of all: to see life as it is, and not as it should be!" [Miguel de Cervantes Saavedra]

2024, April 13: The Drop-Down Box

A coronial Inquiry has criticised Liverpool Hospital in relation to the death of a new-born child, inter alia, for recording incorrect information on the mother's condition:

> After an ultrasound scan in the lead-up to the birth, a doctor at Liverpool Hospital recorded the incorrect location of Carleton's placenta by accidentally using the wrong drop-down box to generate the report, Forbes found. However, even if the position was correctly recorded, it would not have changed the hospital's decision to admit the expecting mother into the operating theatre for the C-section, the coroner said. During the operation, which started almost an hour after an emergency caesarean was called for, the placenta was either unintentionally cut or it separated from the wall in a process called an abruption. This caused excessive bleeding.

Is this way of gathering data prone to error and carelessness? The drop-box is essentially form-filling. Both streamline a process. But they arguably lead to carelessness when the agent lets *the form do the thinking*. That can be a good thing when the intellect of the form-filler is in doubt but not so handy when that is also true of the form-user.
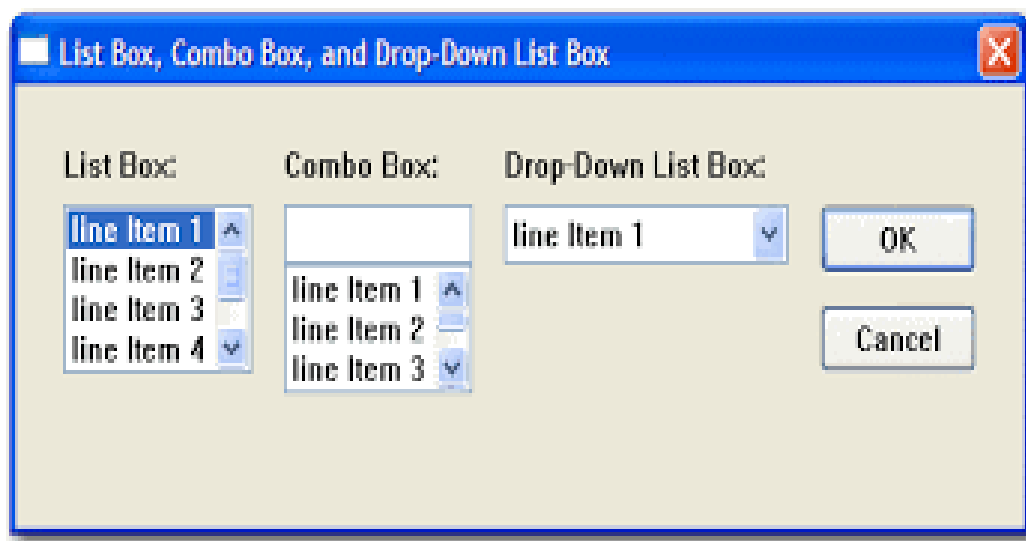
They are also easier to tamper with (cf. *The Verdict: Costello testifies she wrote the patient ate a full meal one hour before being admitted, contradicting the patient record, which states a nine-hour interval. On cross-examination, an incredulous Concannon asks how she can prove this. Costello reveals that her superiors threatened her with termination unless she changed the original record from "1" to "9,")*

Though this does not seem to have been an issue here, all forms suffer from the danger of *binary-blindness*: the delusion that complex situations can be easily reduced to simple yes/no answers (and yes, I do know I've just used a split infinitive and no, I don't care). The drop-down is still binary even if more than two choices are given and they take the form of words or phrases. Good form design may need to provide (where appropriate) for an "other" box that allows for a response in prose – with the attendant danger that you replace doubts about the intelligence of the form-filler with confidence in their judgement and concision (yes, concision is a word).

Of course, the binary (not just on forms) is beloved of lawyers and bureaucrats (and customs inspectors) because it binds you in advance to an answer you can't later wriggle out of.
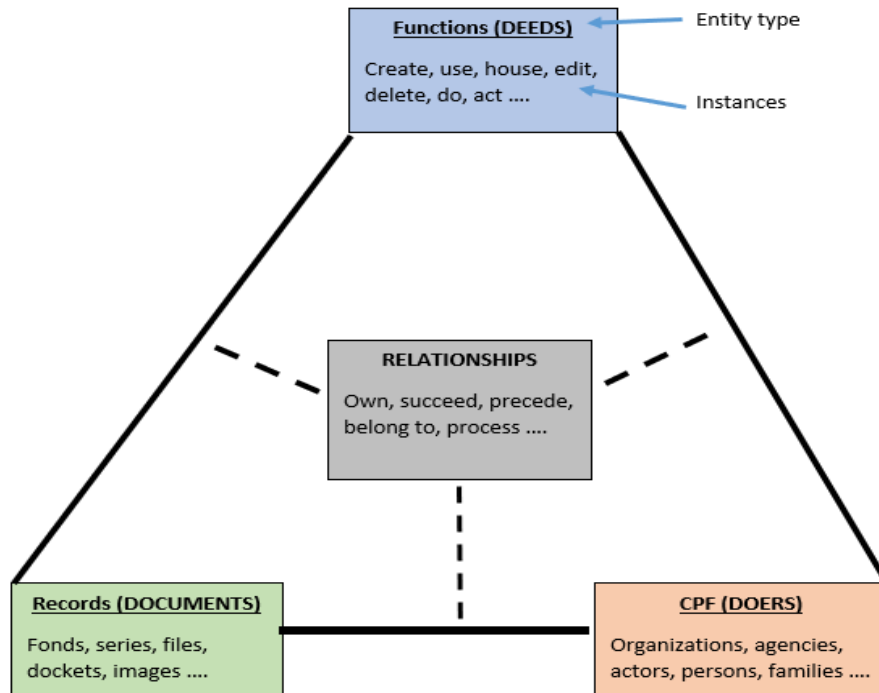
<<it binds you in advance to an answer you can't later wriggle out of.>>

Come to think of it, it might be a good idea to force our politicians to use forms when making their reckless and misleading promises.

APPENDIX

Records in Context (RiC) 1.0 – Comments on First Draft

| Relationships in RiC (1) | posted *Sat Sep 17 05:32:37 EDT 2016* |

Someone has certainly been busy - 792 relationships and still counting. Phew! I read somewhere that a diligent German historian was only able to find 210 reasons for the fall of the Roman Empire. We certainly got that beat. This is a list of implementation options rather than a conceptual model – some of the logical possibilities when designing and implementing an application. To explore the full range of possibilities, two things are needed :

  1. the underlying relationship-types must be identified;

  2. the terms must be defined (cf. p.39) so that we all interpret the words in the same way. Then we can pay more attention to refining or expanding those concepts that are currently being most contested (e.g. "create") and to discovering additional instances (e.g. "received by" under Transmission, "involved party" under Formation, "adopted (by)" under Existential Features, etc.). But it is more important to conceptualise than to itemise, therefore (by way of example):

> One could begin with a thesis (inviting the antithesis) that provenance is to be found in Relationship-Type : Formation (see below). This could be tested by examining whether the 63 instances listed so far are, in fact, acceptable statements of provenance and whether any other ideas about provenance, of the kind that have been put forward lately in the literature, can fit within the instances listed or require additional instances to accommodate them. Is provenance only to be found within Formation? Are there formative relationships that are not allowable statements of provenance? Can provenance be found in other Relationship-Types? Does a formative relationship between Agents ("establish", for example) confer ambient provenance vicariously on a document-type? If so, how would that differ from "uses [agent-delegate]" which I have nominated as Existential? Alternatively, should ambience and provenance be kept conceptually separate? Does the Relationship-Type framework assist or hinder in (re)defining or (re)imagining our core concepts such as provenance.

🏠 MUCH ADO ABOUT DESCRIPTION

I have trouble with two of the proposed entity-types (viz. Date and Place) of which more anon, so I can't yet come to terms with those proposed relationships involving one or other or both of those (204 out of the total). Interestingly, I singled these two out as problems long before I reached p.91 where Date and Place are also nominated as "properties" of relationships so maybe I'm not alone in needing to think some more about them. And I don't think it's worth dwelling long over the relationship-type "associated with" (292 out of the total). We've used that for years as a cop out for making links where we are too lazy or too uncertain to be specific. Anything can be associated with anything and, once you've said that, there's not much more to say and little benefit from saying it 292 times. Of the remainder, here is my first attempt at a categorisation into relationship-types (without the benefit of certainty as to what any of the terms mean):

- **Relationship-Type : Formation** (63 instances) viz. "create/created by"; "authored"; "collect(ed); "wrote/written"; results from/in"; "accumulate"; "assemble"; "arrange"; "establish".
- **Relationship-Type : Governance** (42 instances) viz. "owns/owned by"; "rights held"; "controls"; "directs"; "manages"; "superior/subordinate".
- **Relationship-Type : Succession** (22 instances) viz. "successor/predecessor"; "parent/child".
- **Relationship-Type : Belonging** (30 instances) viz. "part/part of"; "member of"; "is/has example".
- **Relationship-Type : Possession** (12 instances) viz. "held/holder".
- **Relationship-Type : Transmission** (4 instances) viz. "sent by".
- **Relationship-Type : Documentary Features** (73 instances) viz. "copy of"; "draft/original of"; "subject of"; "addressee"; "documentary form"; "evidence of".
- **Relationship-Type : Existential Features** (57 instances) viz. "has/had functional relation"; "assumed identity"; "sibling/spouse"; "uses [agent-delegate]"; "pursues/occupies [position or occupation]"; "fulfils [function]"; "performs [activity]: "authorize(d)"; "required competency"; "defined/revised [by mandate]".

There is, of course, much room for debate (e.g. is "authorize" an instance of the Governance or Existential type?). Nevertheless, I would find discussion at that level more rewarding than simply multiplying instances before something like that has been done.

Relationships in RiC (2)	posted *Sat Sep 24 02:53:40 EDT 2016*

[Daniel Pitti commented: … I think you are correct in saying that the current list of relations under each of the high-level entities is not a conceptual model. I would say that working our way to the relations being properly conceptually modeled is underway but by no means complete… The need to classify and conceptually organized the types is very much on our agenda. I think your first pass very good, but others have been suggested…]

I have no problem with a long list that illustrates a concept. The RiC 1.0 list of relationships could easily stretch from 792 instances to 7,920 and beyond. Thinking up new instances could become a parlour game for archivists. My interest is in what principle(s) the instances illustrate. My suggested categorisation was derived from what is there in RiC 1.0 and is not what I would have come up with if I'd started with a blank page, so "something to live with" would indeed be most welcome. What I mean by implementation is that, w/o further explanation, one has to infer what the terms mean and how they might be used. Taking "creates", for example, and ignoring for the moment its diverse and often contested meanings (simply taking it as an unproblematic idea) it can be applied as a relationship thus:

[ACTOR A]<creates>[RECORD X]

and this seems to be the how RiC 1.0 means it to be understood.

But all recordkeeping is based on describing action and circumstance and "creates" is an action which can, therefore, be rendered as a FUNCTION rather than a relationship (as well as, not instead of). The descriptive statement "A creates X" can then be rendered differently within the RiC 1.0 framework, where FUNCTION M = creates, as:

[ACTOR A]<performs>[FUNCTION M]<to produce>[RECORD X].

It may be that somewhere in the list of possible relations in RiC 1.0 the option of using this second formulation is already provided for but, if so, only the diligent will find it and, absent more explanation, some of them may not understand that these are two allowable ways of achieving the same result. I agree, therefore, with those who have argued that it is important to draw out statements about how relationships are formed from the list of enumerated possibilities.

In the first formulation, according to RiC 1.0, Date & Place could be formulated as properties of a relation and also as instances of Entity-Types in their own right (instead of rather than as well as in any particular instance, I suppose). In the second formulation, it would be easy to link an instance of a Date-Entity and an instance of a Place-Entity to an instance of Function M. For those working a formed archive, the second formulation may seem unnecessarily complex but those involved in active record-making may encounter '000s of create transactions every day and a developer might find it a more effective way of reaching the same outcome (viz. a statement to the effect that "A creates X"). Developers are clever people and could, no doubt, come up with lots more ways of achieving the same outcome for every rule, taking account of the differing needs of their client populations, so long as we provide them with a robust conceptual framework.

Entities in RiC   posted *Tue Oct 11 13:46:05 EDT 2016*

Confusion between Recordkeeping Entities and Authority Records began with ISAAR. This seems an apposite moment to correct the misunderstanding. Four of the 14 proposed Entities (Documentary Form, Date, Place,
Concept/Thing) could be represented as properties of the ten remaining. There is no harm in having those four as entities if that is useful (though the utility eludes me) and many more besides. In some metadata schemas,
Relationships are nominated as entities, for example. But, if you're going to name four, you should make it clear that many other kinds of entity are possible and, if you're going to name those four, you should make it clear that they can (optionally) be treated as properties.

Alternatively, true Authority Records, like EAC-CPF and SNAC, could be built for Documentary Form, Date, Place, Concept/Thing, etc., etc. to control data content of the properties of Recordkeeping Entities. This leads on to the question whether we need to stipulate the properties of Authority Records used in recordkeeping. The other ten Entities proposed in RiC 1.0 are true Recordkeeping Entities whose properties can be controlled by Authority Records of one kind or another (or not, as the user decides). These ten entities can be conceptualised as instances (not the only possible ones) of three basic Entity-Types that are particular to recordkeeping :

- **DEEDS**: events or circumstances that give rise to recordkeeping – e.g. functions, functions (abstract), activities, mandates, processes, responsibilities, products, etc.;
- **DOERS:** actors who undertake the Deeds - e.g. agents, occupations, positions, corporations, agencies, processes, persons, families, etc.;
- **DOCUMENTS**: memories of Deeds undertaken - e.g. records, record components, record sets, series*, fonds*, documentary objects, processes, artefacts, legends, myths, etc.

I deliberately include "process" under all three types to illustrate the point that the same thing can be described in more than one way, using different Entity-Types as appropriate. I

have already suggested the use of Relationship Types and I think using Entity Types is a better way also.

Four properties are common to all Recordkeeping Entity-Types in RiC 1.0 (Global Persistent Id, Local Id, Name, and General Note) and to those I would wish to add Date (either as a relationship or a property). Within the framework of an entity-relationship model, that would satisfy what I see as the mandatory requirements for all Recordkeeping Entities - viz. that they possess :

- **IDENTITY**: because every record is unique;
- **DATES**: because every record is time-bound;
- **RELATIONSHIPS:** because no record stands alone.

Other common properties, such as name, are useful but not essential in recordkeeping. If I were modelling RiC, I would represent the common properties as belonging to a Super-Type of the kind I have sometimes called the URO (Universal Recordkeeping Object), and more facetiously the HERO (Hurley's Enduring Recordkeeping Object). I think a good many more properties (e.g. Description) could be remodelled as common to all Recordkeeping Entities and brought into the URO either because they are already common to all Recordkeeping Entities in RiC or should be.

Other properties might be better handled in other ways, at least as alternative options. Some of these are trifling but "Accruals" (P24 & P25) should be given further thought. Accruals are part of a Process (viz. accessioning) and some people might want to document accessions as Record Sets (or Sub-Sets for incorporation into Sets). I would. That suggests that an option needs to be provided allowing accruals to be treated as Record Sub-Sets with relationships to Record Sets as part of the history of the formation of the Set and not merely as a property forecasting future possibilities. In the physical world, it was sometimes necessary to manage Transfers or Deposits as entities (Record Sub-Sets) separately from Accessions because they comprised one of more Accessions, formed before, during, or after relocation, and I imagine that similar entities might be useful during data migrations.

RiC: Quo Vadis?	posted *Mon Jan 30 00:42:38 EST 2017*

Just before the deadline for comment closes on RiC, here are some clumsy existential questions. They're not just questions for RiC, of course, or what EGAD's next steps should be - but they may apply to the future direction of description overall (RiC or no RiC) - at least I hope it may be so.

Query 1 (Structure):

Can we define an Entity/Relationship type as one containing instances that all operate according to the same recordkeeping requirements (allowing for extensions by sub-types that augment but do not conflict with the common requirements)? Can they be managed, in other words, using identical rules or practices (with extensions) that are set at the level of each type rather than each instance? I once theorised that an ownership relationship is a succession relationship in disguise - easily demonstrated (see below), but not so easily proven. Can we separate conceptualisation and implementation so that a proliferation of instances within each type would not matter. You could have 8, or 800, or 8000 instances of any type; the same standardised practice would govern all. Implementers could then select those instances that are useful to them, ignore the rest, and then apply the rules (or not) as appropriate. Could that approach be taken within an infrastructure (policies, procedures, roles, etc.) that is not particular to any one descriptive programme, jurisdiction, or prejudice?

Query 2 (Identity):

How should we think of the nexus between the description of an entity/relationship and the entity/relationship itself? How does an instance-in-action being described differ from the description of it? Is description simply a parallel universe, laying down a descriptive world

alongside an actual world? Does a recordkeeping (descriptive) system operate in a descriptive universe or in an actual universe or does it straddle the two? Where does our understanding of a corporation, for example, "exist" - in the actual world or within a descriptive (registration) system, or both? Can a description of an instance-in-action in the actual world (physical or virtual) be turned into a kind of avatar so that it can operate in a recordkeeping system as if it were the thing itself, not just a description of it? What is the difference (if any) between action in the virtual world and action in the physical world? How can two different descriptions of the same instance (in the descriptive world) be reconciled? Is there ever a case of a graphical representation for which no extant personality or actuality exists?

Query 3 (Validation):

How can authenticity be conferred on descriptions that operate outside of the source or native system? Could they be trust-worthily registered or validated using PKI and/or blockchain? What kind of recordkeeping system would be needed to validate them (viz. descriptions of description) and could that system be a source for persistent identification? To what extent would that require re-contextualisation? I once asked my friend Terry Cook when he was in full flight about top-down appraisal: How do you know when you're at the top? Reminds me of a great story I once heard about Hilary Jenkinson when he was interviewing a nervous young Oxbridge graduate for a job. Asked what had been his special field of study, the youngster replied, "The end of the 17th century, sir". Jenkinson growled, "Which end?" An archivist's question.

PS. Demonstration of a succession relationship disguised as an ownership relationship:

Consider a simple succession relationship:

AGENCY B------<succeeds>-----AGENCY A

Now, consider two ownership relationships:

FUNCTION G	FUNCTION G
<exercised/owned by>	<exercised/owned by>
(from 1901-1925)	(from 1925-1980)
AGENCY A	AGENCY B

The ownership relationships can be described in a table:

FUNCTION G ....

Dates	Exercised by….
1901-1925	Agency A
1925-1980	Agency B
1980-1995	Agency C
1995-date	Agency D

The ownership data captured in the Table is sufficient, without any need for further data input or description in the form of a succession statement, to generate a succession relationship:

AGENCY B--------<1925: succeeds in exercise of FUNCTION G>-----AGENCY A

Not only has an ownership relationship metamorphosed into a succession relationship, there is added value from depicting how and when the succession arises. The data table can, in fact, generate the following descriptive statements:

- AGENCY B--<succeeds>--AGENCY A in 1925 in exercise of FUNCTION G
- AGENCY A--<succeeded by>--AGENCY B in 1925 in exercise of FUNCTION G
- AGENCY A--<exercised/owns>--FUNCTION G from 1901 to 1925
- AGENCY B--<exercised/owns>--FUNCTION G from 1925 to 1980
- FUNCTION G--<was exercised/owned by>--AGENCY A from 1901 to 1925
- FUNCTION G--<was exercised/owned by>--AGENCY B from 1925 to 1980

PPS. The query "Does a recordkeeping (descriptive) system operate in a descriptive universe or in an actual universe or does it straddle the two? " was posed way back in the SPIRT Project <http://www.infotech.monash.edu.au/research/groups/rcrg/projects/spirt/deliverables/austrkms.html> (Business Recordkeeping entity class posited as a sub-set of the Business entity class). I never thought the answer was entirely satisfactory, but it was the right question to ask.