rethinking digital research

kaitlin thaney
@kaythanean ; @digitalsci
wolfram data summit
discovery
management
the “social”
discovery & delivery
changes at the workbench
annotation & curation
social & administrative
significant
(and somewhat silly)
gaps still exist
what we know, is research.
research is changing
(as is the research workflow)
discovery is still sub-optimal.
still the starting point
specific DBs no better
(in this case, likely worse)
name disambiguation

103-90-2
254-465-1
4-(Acetylamino)phenol
4-13-00-01091 (Beilstein Handbook Reference)
4-Acetaminophenol
4'-Hydroxyacetanilide
4-Hydroxyanilid kyseliny octove
Acetamide, N-(4-hydroxyphenyl)-
Acetamide, N-(p-hydroxyphenyl)-
Acetanilide, 4'-hydroxy-
APAP
N-(4-Hydroxyphenyl)acetamid
N-(4-Hydroxyphenyl)acetamide
N-(4-Hydroxy-phenyl)-acetamide
N-(4-Hydroxyphényl)acétamide
N-(4-Hydroxyphenyl)acetonilide
N-(p-hydroxyphenyl)acetanilide
N-Acetyl-4-aminophenol
N-Acetyl-p-aminophenol ...

and the list goes on ...
streamline search
our definition of “data” is changing.  
(or, at least, is long overdue to.)
mix of digital and analog mediums
the non-digital
the non-digital
Pork Spread Rillettes

The second characteristic imparts a

most particularly different punch. Accord

ing dishes to different salads of fish, salami, rillettes, etc. The following

cheese, black pudding, chitterlings, veal, veal

preparations available from characterise

Rillettes consist of seasoned pork

tiny pieces, bound together by fat.

They are eaten cold with Dijon

course or as a snack. The rill

duck, but most of the time pork will be ad

and make it the right consistency. The ril

crots and served as such with bread,

stock, thyme, bay leaf, etc., but the

water retains the true taste of the

watered tightly with plastic wrap and

longer the crocks have to be sealed with melted fat. When these

changes and the taste tends to grow rancid.

2½ pounds fresh pork from the chuck, butt or

neck (the meat should have about one-third

lean meat to two-thirds fat)

salt

coarsely ground black pepper
wasted ...
$$$

$$$ 
time
resource
our practices are limiting us.

(and in some cases, harming.)
3 Gb/year
9 Gb/PhD
3 Gb/year
30,000,000 Gb/all*

9 Gb/PhD

3 Gb/year

* roughly
how much made available?

30,000,000 Gb/all*

9 Gb/PhD

3 Gb/year

* roughly
make stupidly simple

... and part of the process.
“Non-reproducible single occurrences are of no significance to science.”

- Karl Popper
“That’s Dr. Arnold Moore. He’s conducting an experiment to test the theory that most great scientific discoveries were hit on by accident.”
We are not in a position to supply data for a particular country not covered by the example agreements referred to earlier, as we have never had sufficient resources to keep track of the exact source of each individual monthly value. Since the 1980s, we have merged the data we have received into existing series or begun new ones, so it is impossible to say if all stations within a particular country or if all of an individual record should be freely available. Data storage availability in the 1980s meant that we were not able to keep the multiple sources for some sites, only the station series after adjustment for homogeneity issues. We, therefore, do not hold the original raw data but only the value-added (i.e. quality controlled and homogenized) data. The priorities we use when merging data from the same station from different sources are discussed in some of the literature cited below. Parts of series may have come from restricted sources, whilst the rest came from other sources. Furthermore, as stated in http://www.cru.uea.ac.uk/cru/data/landstations/ we have never kept track of changes to country names, as it is only the location and the station's information and coordinates we are interested in.
rethink the way we rank researchers.

(the big, hairy “social” issue.)
What Des-Cartes did was a good step. You have added much several ways, & especially in taking ye colours of thin plates into philosophical consideration.

If I have seen further it is by standing on ye shoulders of Giants.

- Isaac Newton, 1676
locked in old mechanisms.

WONKA'S GOLDEN TICKET

GREETINGS TO YOU, THE LUCKY FINDER OF THIS GOLDEN TICKET FROM MR. WILLY WONKA.

PRESENT THIS TICKET AT THE FACTORY GATES AT TEN O'CLOCK IN THE MORNING OF THE FIRST DAY OF OCTOBER AND DO NOT BE LATE. YOU MAY BRING WITH YOU ONE MEMBER OF YOUR OWN FAMILY... AND ONLY ONE... BUT NO ONE ELSE......

In your wildest dreams you could not imagine the marvellous surprises that await YOU.
“Right now we're going through a Cambrian explosion of metrics.”

- Johan Bollen
there’s been a drastic spike in terms of sheer **volume** and **type**
citation / impact factor

\( h \) - index

weighted citations (eigenfactor, sjr)

“betweenness centrality”

alt-metrics, etc.
difficult to ... harmonise track / maintain / map understand (even still measure)
administrators / funders as *influencers* *drivers* of behavioral change
we’re battling tradition, not just technology.