University of Pittsburgh

SCHOOL OF Information Sciences

Preparing the workforce for digital curation: the iSchool perspective

Reflections & Challenges

IDCC Panel, February 2014

Professor Liz Lyon, School of Information Sciences, University of Pittsburgh

Flavours of data scientist

(Lyon 2012)

- ***data engineer -*** focus on software development, coding, programming, tools
- ***data analyst –*** focus on business/scientific analytics and statistics e.g. R, SAS, Excel to support researchers and modellers, business
- ***data librarian*** – focus on advocacy, research data management / informatics in a university / institute
- ***data steward –*** focus on long term digital preservation, repositories, archives, data centres
- ***data journalist –*** focus on telling stories and news

Likely degree backgrounds

- ***data engineer -*** computer science, engineering
- ***data analyst*** **–** mathematics, statistics, business studies
- ***data librarian*** – arts & humanities
- ***data steward*** **–** arts & humanities
- ***data journalist*** **–** journalism, media studies, communications studies

Flavours of disciplinary/ domain (X-) informatics

- Archival informatics
- Bioinformatics
  - Bioimage informatics
- Biodiversity informatics
- Business informatics
- Cheminformatics
- Community informatics
- Computational informatics
- Development informatics
- Disease informatics

- Ecoinformatics
- Education informatics
- Engineering Informatics, an interdi
- Environmental informatics
- Evolutionary informatics
- Forest informatics
- Geoinformatics, an interdisciplinar
- Health informatics
  - Consumer health informatics
  - Imaging informatics
  - Public health informatics
- Hydroinformatics
- Irrigation informatics
- Laboratory informatics
- Legal informatics

- Materials informatics
- Medical informatics
- Music informatics
- Neuroinformatics
- Pervasive Informatics
- Social informatics
- Technical informatics
- Translational research informatics

STEM: strongly represented….

“Very few librarians are likely to have specialist scientific or medical knowledge - if you train as a research scientist or a medic, you probably won’t become a librarian.”

RLUK/Mary Auckland: Reskilling for Research 2012

https://www.flickr.com/photos/23312112@N04/9108008669/i[illegible]hotolist-e[illegible]Fe-9dju9H-f3rUKg-h2kzk1-8[illegible]-ahsWNa-a8CXWA-6SEFzn-[illegible]DMRoo-4Zudts-TpNW5-4jh869-2MhSt[illegible]-8XQvXo-8s9Uup-6GB7QU-995KZ7-7uG7vL-[illegible]1-77mSoG-7LGBw4-at7uYC-ghoME2-jfbsjF-8rHmyB-khyact-7gWFgw-968cHa-i8gDyd-jrvv7v-hu8KBH-5X3pmT-8LBseT-dMXgXq-fmNe5N-dNuWKB-dNMux[illegible]brmocc-b9djCp-yKP5Q-dsQ5xw-9HtSRR-eRVdi8-5BncXP-apYD4x-6gQ4mJ-4N6WQv

Curation : domain disconnect ?

Science

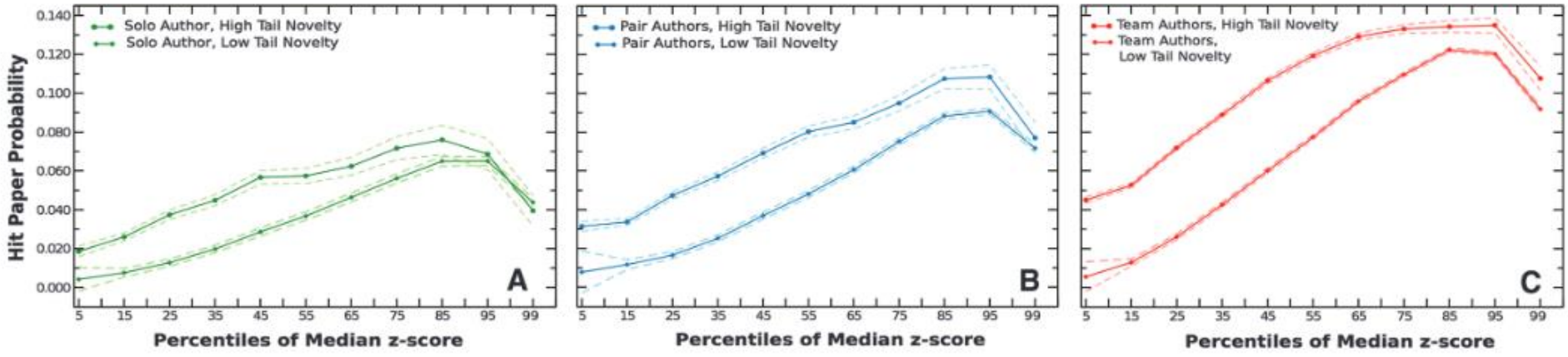
AAAS

Atypical Combinations and Scientific Impact

Brian Uzzi *et al.*

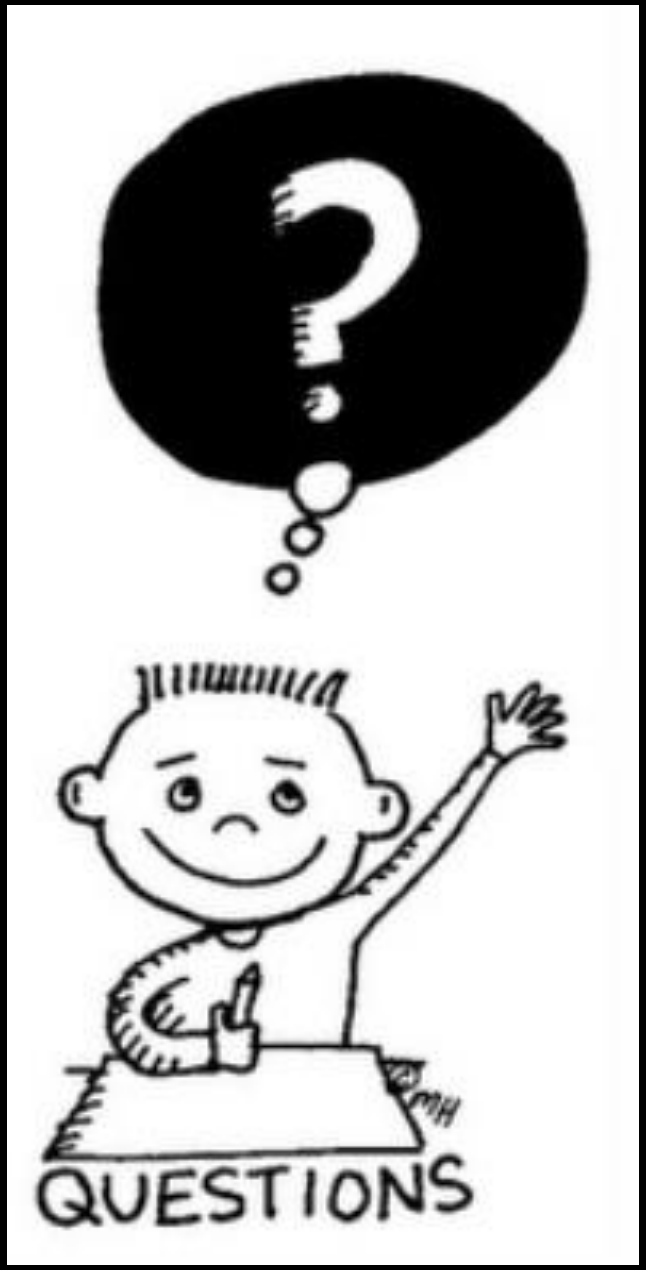
Science **342**, 468 (2013);

DOI: 10.1126/science.1240474



"…teams are 37.7% more likely than solo authors to insert novel combinations into familiar knowledge domains".

Teams: multidisciplinary?
+/- data curation? informatics?



“If science becomes more multi-disciplinary, should I aim towards multi-disciplinary skills or towards specialisation?”

Curation skills?

Workforce challenges in 2014

1. Who should we attract to the data curation workforce?

2. How best to address data curation and team science?

3. How to balance multi-disciplinary vs domain specialisation skills?