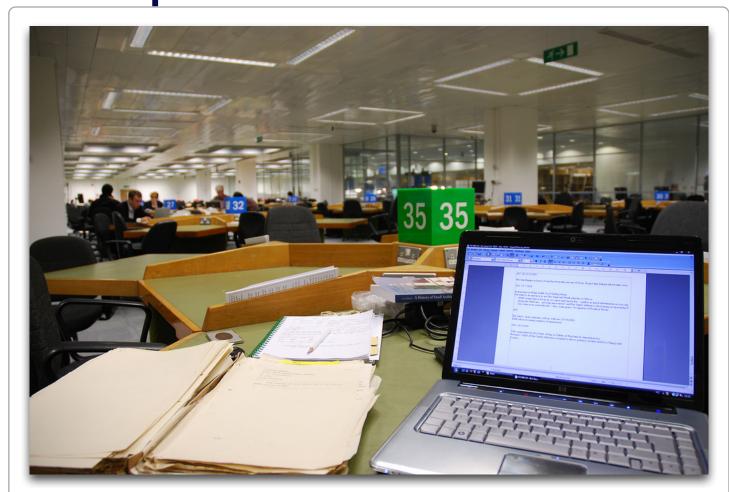
Creating a Digital History Commons through crowdsourcing and participant digitisation

Mia Ridge, Open University, UK

The premise...



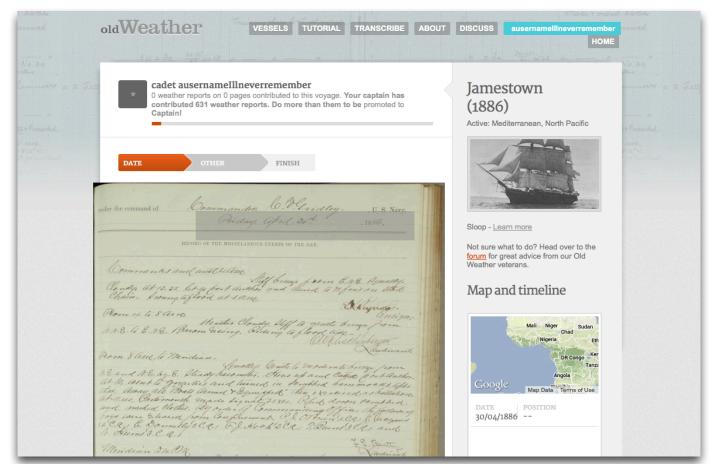
Simone travels to archives in her semester breaks, taking photos of most documents 'in case they're useful later', transcribing key text from others. Andre is at the next table. The documents he collected for his last research project would be useful for Simone's current book but he has no way of sharing them.



Martha and Bob joined their local history society when they retired. They're helping find out what happened to children from the village school's class of 1898 during World War I. They are using census returns and other online documents to add records to a database the society's secretary set up in Excel.



Daniel has his grandmother Sarah's letters and diary, describing her travels and everyday life from the turn of the century. He's transcribing them and wants to put them online to share with his extended family. One day he wants to make a map for his kids that shows all the places their great-grandmother visited.



Seema has two young kids and works for a local authority. She enjoys playing games on her mobile, and after the kids have gone to bed she transcribes ship logs on the Old Weather website while watching TV. Sites like Old Weather use 'microtasks' and crowdsourcing to digitise large amounts of text.

... a Commons could help connect everyone's efforts for the greater good

None of our friends above know it, but they're all looking at material from roughly the same time and place. Andre and Simone could help each other by sharing the documents they've collected over the years. Sarah's diaries include the names of many children from her village that would help Martha and Bob's project, and Seema could help everyone by transcribing sections of Sarah's diary. These materials could be aggregated in a Commons accessible to all. By enhancing the records with the techniques and productivity of crowdsourcing in combination with the contextualisation and research tasks historians already do, primary sources from anywhere could be made available in digital repositories.

Building a Participatory Commons

Official collections held by archives, libraries, museums

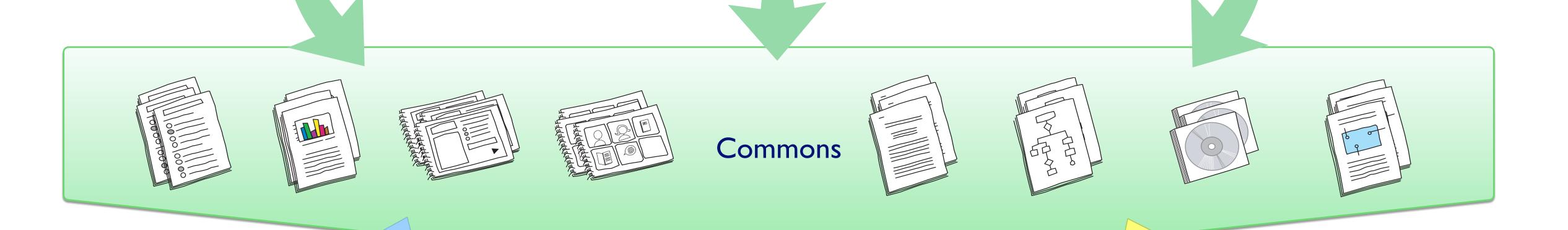
- Includes letters, diaries, personal photographs usually from prominent people or organisations; property transactions; births, deaths and marriages; wills; employment records; local newspapers
- Organised by source and purpose

Historians' personal research collections

- Includes texts, tabular data, images, artefacts from a variety of sources
- Enhanced with annotations, references, variable documentation at item level
- Organised by research project, collected over time for different purposes, uncertain copyright status

'Shoe box' archives

- Collections held by ordinary people inherited from family etc
- Sparse documentation e.g. personal recollection, album labels
- Includes letters, diaries, personal photographs usually from ordinary people; scrapbooks; artefacts and ephemera



Typical tasks historians perform

Take digital photographs

Organise by research question

Compare to other sources

Record document context

Link to people

Link to places

Structured text transcription

Descriptive tagging

Typical crowdsourcing microtasks

Full text transcription

Research Questions

- Is a Commons the best model for collaborative digital history resources?
- What is needed to ensure the reliability of resources from 'shoe box' archives?
- Which lessons in microtask design and task scaffolding from citizen science/crowdsourcing are relevant?
- How and why might historians resist using or contributing to collaboratively created digital history resources?
- How might historians be encouraged to share their personal research collections alongside traditional publications?
- How can the expertise of archivists be shared beyond physical visits to archives?