1891 England & Wales Census

Ancestry.com brings every-name index
with linked images to the Internet

Greg Parkinson - Architect
Laryn Brown - Product Manager

Problems to Solve

- Scanning and image prep
- Compression and storage
- Mask creation, tiling, and thumbnail generation
- Browse table creation, jurisdiction normalization
- Indexing 30 million names
- Quality assurance and correction feedback

Scanning & Image Prep

- Film acquisition and copyrights
- Scanner speed and resolution determination
- File type? TIFF, JPEG 2000, GIF, PDF, SID

Compression & Storage

- Storage costs are dropping, but still can be considerable
- High bandwidth becoming more available, but not widespread
- Future uses of the image may need higher resolution

Mask, Tiling, Thumbnail

- Mask is the ultra low res view prior to tiling
- Tiling allows users to see a portion of the image without the whole image loading
- Thumbnail view is helpful for navigation orientation
- Watermarks
- Programatically create all these

Browse Tables, Jurisdictions

- Browse tables allow users alternate access to images
- Browseable images can precede keyed indexes
- The crux of browse tables is a understanding of jurisdictions and their relationships.

30 Million Name Index

- File prep and transfer to keyers
- Keying every name and nearly every field on the census return (1.5 billion characters)
- Field authorities (Name, place, relationship)
- Linking images and indexes
- Audit

Quality Assurance

- Images - missing, bad scans, wrong links
- Index - testing automation
- Post scan audits, post keying audits, in-house audits, functional testing, testing on live
- Customer feedback mechanism
- Rescans

What We Learned

- Take great care with the scanning. Bad quality or the wrong decision on file size-compression can be costly.
- Leverage your work (1930, Newspapers)
- Choose your viewer technology carefully
- Allow customers to correct your mistakes
- Create tools for production and testing as part of the project
- Hire great engineers

Live Demo

- Browsing
- Searching the index
- Html based viewer
- Ancestry Advanced Viewer