

# How Historical Documents Become Discoveries at Ancestry

Mike Tolman // Product Manager – Content Systems // Ancestry

[mtolman@ancestry.com](mailto:mtolman@ancestry.com)

## Introduction: Discovering Your Past, Inspiring Your Future

At Ancestry, we add around 50 million images and records to our website every month. While we cannot disclose many of the specific techniques and strategies we use to acquire records for new regions and countries, we will talk about what we acquire and what we do with that content once we get it. At a very high level, these historical records are the fuel that feeds the fire of discovery at Ancestry and is the source for the rich experiences we provide to our subscribers. In this presentation, we will talk about our partnerships with archives and other resources throughout the world. We will also discuss how we ensure that the images and data are of the highest quality possible. Finally, we will highlight some revolutionary and innovative technologies that we utilize to improve original source content and to make that content more useful and impactful. All of that said, our entire mission is summed up in our commitment to empower journeys of personal discovery to enrich our user's lives.

## The Acquisition Process

By closely monitoring customer feedback and market trends, Ancestry is in constant contact with archive partners and data resources throughout the world to continue to enhance and enlarge our image and record collection databases. Once a record sharing agreement is reached, Ancestry will send document preservation specialists or technologists to our partners and we will engage in the work to digitize those records. Those records are then brought into our systems to be processed and enhanced to ensure an optimal viewing experience for our customers.

## Quality at Ancestry

Every month, over 20 million images are processed by Ancestry. Every single image is viewed and analyzed to ensure optimal quality. We even utilize specialized photographic and scanning equipment that give us best in class and cutting-edge capabilities that are unmatched in the industry. Once those images have been processed, they are sent along to our keying vendors around the world who create the records that make these images and data searchable and hintable. These records go through various stages of quality checks before they are published to our website. Even then, we allow our users to make suggestions and corrections to improve the quality of that data.

## Visual Technologies to Improve on the Original Source Content

Through the years, Ancestry has developed some innovative solutions to preserve historical documents and to even save the information that would have otherwise been lost. One example is how we regularly encounter damaged documents that are difficult to read because they have faded due to exposure to water, mold, fire, excessive light, or natural fading. We solved this using multispectral analysis to view parts of the document that are only visible in alternate spectrums. Other techniques we use are adaptive auto-contrast algorithms, image binarization, and adaptive enhancements.