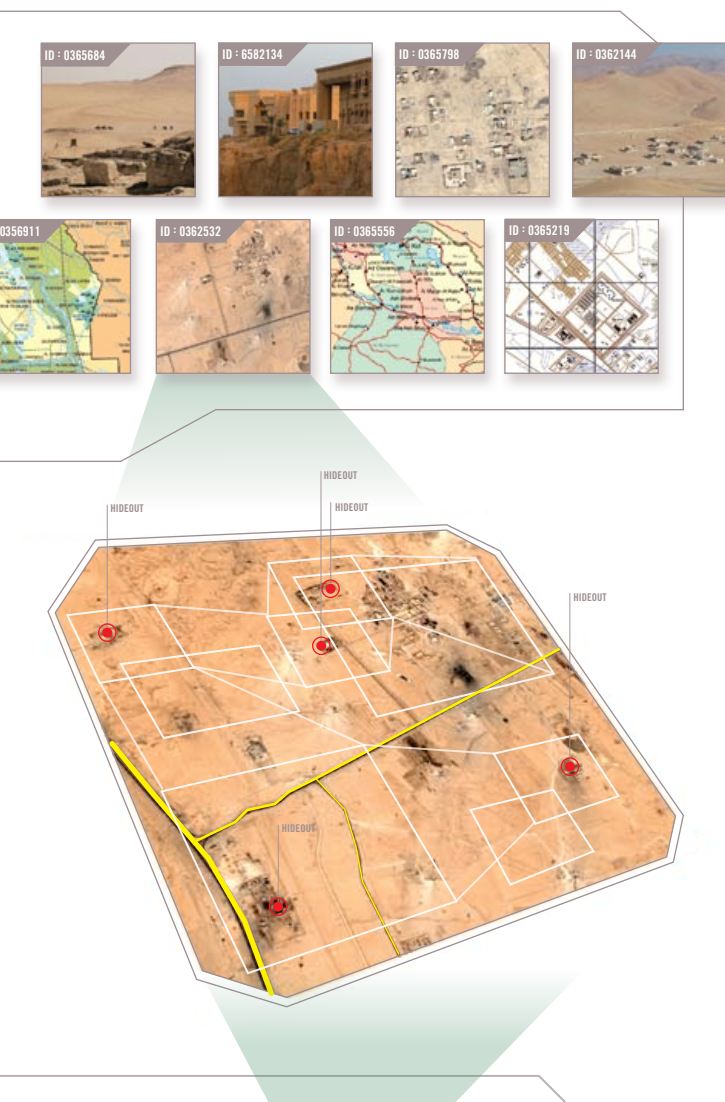


In the know

Documenting and re-creating top-level decision making processes is the focus of government historians.

Story / Barbara Jewett
Illustrations / Blake Harvey



Who knew what—and when—is on everyone's mind when it comes to high-level government decisions that affect hundreds of thousands of people. A team from NCSA is working with the National Archives and Records Administration (NARA) to understand preservation and reconstruction requirements for computer-assisted decision processes, and to prototype the necessary tools so future historians will be able to answer that question.

Peter Bajcsy and his colleagues Sang-Chul Lee and Rob Kooper developed a simulation framework to gather information about decision-making processes based on examining images. Frederico Fonseca of The Pennsylvania State University used a summer fellowship at NCSA in 2006 to work with Bajcsy's team creating terminology consolidators, called ontologies, which allowed further applications of the developed software. The focus of the team's work is not only on preserving gathered information and reconstructing the process that led to the decision that was made, but also on understanding preservation and reconstruction requirements.

Analyzing the information

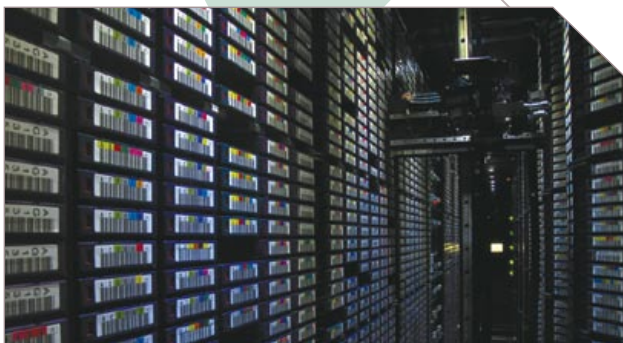
Let's say that high-level government decision makers are inspecting images, trying to identify terrorist hideouts in the course of deciding where to deploy troops. They process maps and photographs, geographical boundaries, and tabular statistics.

Recording the decision process

While any government decision process is complicated on its own, tracking the analysis of geospatial electronic records supporting government decisions adds another layer of complexity. Using NCSA's Image Provenance to Learn (IP2Learn), these critical decisions could be recorded at multiple levels of detail, preserved, and recreated later. IP2Learn tracks all operations of a user during the image inspection and reconstructs the image-based decision process. The software records and reproduces information about how a digital image is manipulated (for example, select, crop, magnify, or adjust the color) at varying levels of detail, from a textual summary of raw or interpreted image manipulation events to a full video replay of the actions taken in inspecting an individual image. This can be crucial in reviewing image analysis involving situations of national security; audits can answer questions like "Was every sub-area of an input image inspected?" or "Was the critical image sub-area inspected at the highest resolution?" or "Was it possible to recognize the object of interest from all viewed image sub-areas?"

Preserving the data

IP2Learn allows archivists to examine empirically the tradeoffs related to questions such as what should be preserved and how the data should be gathered, stored, and retrieved. They can also consider how the decision-making processes should be reproduced, what questions researchers will be able to answer using the reproduced information and, most importantly, the historical value of a recreated decision process. IP2Learn captures, summarizes and displays information about the costs (storage space and computing cycles) associated with each approach to capturing, preserving, and providing access to these records, facilitating the evaluation of cost-versus-value tradeoffs.





Balancing the differences

Fonseca came to NCSA because of his interest in expanding the IP2Learn framework with the concept of ontologies. Ontologies are databases that contain words and the mappings (terms, attributes, and relationships) between words as well as strict rules about how to specify terms and relationships. The mappings can be quite complex.

Besides inconsistencies in terms used, for example one person using “hill” and another “rise” when both mean an upward slope in the topography, the decision maker may use a different meaning than the analyst. Consider what comes to your mind when you hear the word “plane.” You might think of a jet, while another person assumes a single engine propeller plane.

Ontologies are applied at the moment when multiple people annotate the same image. The ontology-based “translations” integrate the meanings of the annotated images for full understanding of image content. Executing ontologies provides a comprehensive description of an image area, allowing a decision to be made quickly and with confidence.

The future

NCSA continues to gather data about the use of computers in decision making where high confidence in the information the decision is based on is imperative. Thus, IP2Learn and other technologies developed at NCSA, like CyberIntegrator (http://isda.ncsa.uiuc.edu/ecid/ECID_cyberintegrator.htm), could be used for gathering software execution information and expanded to other applications. For instance, the IP2Learn software could be used in medicine, in law enforcement, for emergency planning and responses to epidemics, for auditing and quality control in a variety of government and industry situations, as well as for education and training purposes.

The IP2Learn research is funded by the National Archives and Records Administration. The work with Frederico Fonseca was funded by the NCSA summer fellow program.

Team Members

Peter Bajcsy
 Frederico Fonseca
 Rob Kooper
 Sang-Chul Lee

