

Position Paper

Digital Gazetteer Research and Practice workshop

November 17, 2006

A Case-Study-In-Progress: How a Media Organization Tackles the Georeferencing Challenge/Opportunity

Founded in 1888 to increase and diffuse geographic knowledge, the National Geographic Society has grown into a multi-faceted organization producing editorial material in multiple media and across numerous business units. The Society, like other media companies, is working hard to make its digital assets more versatile and accessible amidst a rapidly evolving, increasingly fragmented media marketplace. The growing popularity of Web-based mapping platforms and consumer-oriented GPS applications have spotlighted a particular opportunity, namely organizing NGS content for presentation in the context of location. The Society's venerable brand, its renowned cartography, much of its editorial content, and the core of its mission are all about geography. For all these reasons, the notion of georeferencing Society content has been an easy sell.

The process of *implementing* georeferencing, however, has proven to be anything but simple. Challenges include the vast scope and variety of content at the Society, its archiving in scattered locations across multiple business units, its access via a variety of databases with little or no common standards or nomenclature, and widely varying rights restrictions. Despite these challenges, progress is being made in the creation of an enterprise-wide infrastructure, and in pursuing project-based initiatives fueled by specific business opportunities, most of which complement and strengthen the enterprise effort.

Overall goals of the enterprise-wide effort are to create an infrastructure for organizing and accessing content by geography, including archival content and new content as it is created in the field or archived at the office. The infrastructure must allow for a variety of media formats, several database and metadata systems, and a range of potential applications including print, Internet, video, film, and mobile/GPS. An additional goal is to integrate georeferenced content with National Geographic's cartography, potentially providing a unique competitive advantage.

The key to georeferencing legacy content, of course, is the placename or names associated with that content. National Geographic Maps' placenames database, which is essentially the same as the index to the Eighth Edition Atlas of the World, was identified as the basis for a "master gazetteer" against which other databases would be "harmonized" or cross-walked. Disadvantages of the atlas index are the small number of placenames (some 140,000) compared with other gazetteers, and the fact that the features those placenames are associated with are cartographic (adjusted for visual display) rather than truly spatial. Advantages include the high editorial quality of the placenames and their association with a

GIS and its attendant schema (hierarchies, feature types, etc.).

Off-the-shelf software will be utilized to cross-walk or conflate NGMaps' placenames with at least three other NGS databases, each with more than 20,000 entries: the Society's publications index, the Film Library database (video and film), and the Image Collection database (photography). The Getty gazetteer will be tapped as a reference tool to identify placement within the partitive hierarchy for those placenames not in the current Maps database. Meanwhile, the hierarchical structure of the Maps placenames index (continent-country-province, etc.) will be enhanced using standards and practices of existing online gazetteer services, but fine-tuned to suit the special requirements of NGS. Placenames will be matched with lat-long coordinates (points, lines, bounding boxes, and/or polygons), at which point it will become a bona fide gazetteer. Finally, the same placenames will be linked to a Web-enabled version of the new seamless, multi-scale GIS cartographic databases currently under development in NGMaps, making it a *GIS gazetteer* as well. The master gazetteer will be accessible within NGS via a Web interface, enabling archivists to use it as a reference tool for georeferencing new and future content. Thus over time the need for cross-walking will decline.

Inevitably, the Society will embrace the use of GPS-enabled cameras and video equipment. Meanwhile, NGMaps is using software tools to incorporate spatial information into image headers and metadata, tapping placenames, addresses, and descriptive information. National Geographic Maps, the Society's Digital Media group, and the Library and Information Services division are in discussions with MetaCarta about using their powerful text parsing tools to facilitate the georeferencing of text-based content and media metadata.

NGMaps and Digital Media have been collaborating with ESRI to incorporate georeferenced content into the MapMachine, National Geographic's mapping platform. The MapMachine will soon be able to display icons that will provide access to hundreds of photographs, articles, video clips, and sound files. This will provide a new means of access to the vast assets of the National Geographic website. It will also serve as a proof of concept and, we think, a beta version of an exciting new storytelling tool. For 91 years, cartographers at National Geographic have used maps to tell stories about the world. Linking dynamic maps and multimedia content through georeferencing will enable us to create authoritative, curated resources, including narratives, guided tours, news features, and curriculum materials, bringing our spatial storytelling skills to bear in new media environments for the benefit of a global audience.

Allen Carroll
Chief Cartographer
National Geographic Society