

MapCruncher: Bringing Together the World’s Geographic Information

Jeremy Elson, Jon Howell, John Douceur, Danyel Fisher

Microsoft Research

Abstract

Current large-scale interactive web mapping services such as Virtual Earth and Google Maps use large distributed systems for delivering data. However, creation and editorial control of their content is still largely centralized. The Composable Virtual Earth project’s goal is to allow seamless interoperability of geographic data from arbitrary, distributed sources.

MapCruncher is a first step in this direction. It lets users easily create new interactive map data that can be layered on top of existing imagery such as road maps and aerial photography. MapCruncher geographically registers and reprojects the user’s map into a standard coordinate system. It then emits metadata that makes it easy for anyone on the Internet to find the published map data and import it. Interactive maps them become distributed, seamlessly composable building blocks – similar to images in the early days of the Web.

The Composable Virtual Earth Project

In the relatively short time since the introduction of Mashup APIs for sites like Google Maps and Virtual Earth, hundreds of user-created mashups have appeared. These mashups cover a wide diversity of subjects. For example, BusMonster plots public transportation routes; ChicagoCrime highlights dangerous areas of the city; RunwayFinder summarizes weather and airspace surrounding general aviation airports; HousingMaps.com shows real estate prices. These specialized sites each display the data from their particular application domain on top of maps and aerial imagery supplied by Google or Microsoft.

While useful individually, mashups can be far more useful if *integrated with each other*. Currently, this is effectively impossible; each Mashup is independent. Users who, for example, want to find cheap real estate in a low-crime neighborhood, or find the public transportation near a general-aviation airport, have little choice but to visit each Mashup individually and manually integrate the results. The goal of Microsoft Research’s Composable Virtual Earth project is to find new ways of constructing geographic Web mashups so that they can seamlessly interoperate.

Existing mashups are implemented largely in imperative code – JavaScript that runs in the client. This design gives site designers enormous flexibility, which led to the explosion of mashups that were creative and innovative. Well-known standards for describing geographically-tagged points, lines, and raster graphics had already existed for many years (e.g., GML, GeoRSS). The sudden appearance of the mashups suggests many applications are not well-served by those standards. The combination of HTML and JavaScript allowed developers to go beyond creating layers, to create *applications* – that is, to use

imperative code to customize exactly how their application operates, rather than simply creating layers declaratively whose user interactivity typically is limited to “on” and “off.”

A key design goal of CVE is to offer a Mashup framework that is sufficiently structured to enable composition, yet sufficiently flexible to admit innovation. This interoperability balancing act is common in distributed systems design, from domain-specific frameworks such as the Flux OSKit, xKernel, and stackable filesystems, to application-agnostic schemes such as Placeless active properties. We plan to exploit the geographical domain constraints to best achieve this balance.

MapCruncher

As a first step, we created MapCruncher, a tool that allows users to add custom raster overlays onto the existing road and aerial imagery provided by Virtual Earth or Google Maps. Overlays are typically detailed maps – such as a bicycle route map, building floor plan, or campus map. The resulting web site is an interactive web map that features both the user’s maps and the standard imagery. Like the underlying maps, user maps are pre-rendered into small image tiles at a variety of zoom levels, allowing the client to efficiently request the portions of a large virtual image that are needed for display.

MapCruncher first assists in registering the foreign map into the same (Mercator) coordinate system used by existing online map sites. Users select correspondence points between their own maps and existing maps – road intersections or other recognizable landmarks. Once enough points are selected, MapCruncher estimates the transformation from the original map’s coordinate system into Mercator by finding the best fit coefficients of a second-degree polynomial; while inexact, the error is typically small enough not to affect the results. MapCruncher then reprojects the original map and renders correctly registered and zoomed image tiles that can be seamlessly integrated with existing imagery.

Mashups created with MapCruncher do not restrict the developer’s freedom to write arbitrary JavaScript that customizes the experience of their end-users, satisfying one of our design constraints for CVE. However, MapCruncher also emits metadata *about* the Mashup, such as its geographic bounds, the file naming scheme for the tiles, and a brief description of the data as entered by the user. Because this data is semantically meaningful, it vastly simplifies later discovery and integration of the imagery into other applications. In addition, much of this data is encoded as specially constructed strings, allowing us to exploit normal search engines to find Mashups on the web that meet certain criteria. This combination of compositability and discovery takes us a step closer to our goal of a system that is capable of more seamless integration of geographic data on the web.