

Map mashups, Web 2.0 and the GIS revolution

Michael Batty^{a*}, Andrew Hudson-Smith^a, Richard Milton^a and Andrew Crooks^b

^aCentre for Advanced Spatial Analysis, University College London, London, UK; ^bCenter for Social Complexity, George Mason University, Fairfax, VA, USA

(Received 31 January 2010; final version received 8 February 2010)

Mashups, composed of mixing different types of software and data, first appeared in 2004 and ‘map mashups’ quickly became the most popular forms of this software blending. This heralded a new kind of geography called ‘Neogeography’ in which non-expert users were able to exploit the power of maps without requiring the expertise traditionally associated, in the geographic world, with cartography and geographic information science, and, in computer science, with data structures and graphics programming. First we suggest the need for a typology of map mashups while arguing that such a typology is premature. We then discuss the need for standards and formats, moving on to questions of security, privacy and confidentiality. We follow this by introducing the key issues of creating spatial data for mashups through crowd-sourcing. To ground this presentation in applications, we explore some classic exemplars from our own and related work with map mashups and portals such as MapTube (<http://www.maptube.org/>). We then point to extensions to other graphical media, to 3D, to virtual worlds and beyond. In conclusion, we speculate on what all this might mean for GIS software and geographic information science.

Keywords: map mashups; crowd-sourcing; Google maps; MapTube; Neogeography; map standards; 3D mashups

The context: Web 2.0 and beyond

Right from its inception in 1992, the great innovation of the web was to communicate and disseminate graphical information. Map applications appeared almost immediately, first as backcloths on which to display locational information. By the late 1990s, various products that enabled users to ‘find their way’ – gazetteers and atlases – had appeared such as MapQuest (2010) and in the early 2000s, value was being added to these interfaces as they began to be customised to provide new layers of spatial data that users could query. A good example in the United Kingdom is ‘UpMyStreet’ that was first created in 1998 and now contains a wealth of local information targeted around the search for property and local services (UpMyStreet 2010). In February 2005, however, the field was revolutionised with the introduction of Google Maps (2010) closely followed by its applications programming interface (API) in June of that year that let users embed their own varieties of Google Map within their own web pages. The many applications that have followed define the fast-moving field of ‘Map Mashups’ that will be described here.

The context for the rapid development of these new user-orientated technologies in general and map mashups in particular is the rise of Web 2.0 and the emergence of Neogeography which we need to define at the outset. We can do no better than quote the academic’s *bête noir* Wikipedia (2010) when it comes to defining Web 2.0. Then:

The term ‘Web 2.0’ (2004–present) is commonly associated with web applications that facilitate interactive information sharing, interoperability, user-centered design, and collaboration on the World Wide Web. Examples of Web 2.0 include web-based communities, hosted services, web applications, social-networking sites, video-sharing sites, wikis, blogs, mashups, and folksonomies. A Web 2.0 site allows its users to interact with other users or to change website content, in contrast to non-interactive websites where users are limited to the passive viewing of information that is provided to them.

The term is closely associated with Tim O’Reilly because of the O’Reilly Media Web 2.0 conference in 2004. Although the term suggests a new version of the World Wide Web, it does not refer to an update to any technical specifications, but rather to cumulative changes in the ways software developers and end-users use the Web. Whether Web 2.0 is qualitatively different from prior web technologies has been challenged by World Wide Web inventor Tim Berners-Lee, who called the term a ‘piece of jargon’ – precisely because he intended the Web to embody these values in the first place.

In terms of the development of the web, then it is clear that despite Berners-Lee’s assumption that the web should enable users to both extract and create (read and write) content, until 5 years or so ago, this was largely not possible for ‘ordinary users’. Thus the earlier era might be called Web 1.0 which begs the question that we implicitly allude to, as to whether or not a Web 3.0 is on the horizon. Nevertheless this definition from Wikipedia focuses directly on the

*Corresponding author. Email: m.batty@ucl.ac.uk

technologies we are summarising here which are those that enable non-expert users to create maps and manipulate map data in user-friendly ways. These applications are opening up these technologies to a much wider audience of potential users than anything available hitherto. In this sense then, Web 2.0 moves geographic representation, manipulation and analysis beyond conventional GIS systems, thus stretching the bounds of geographic information science.

In the context of geospatial data and mapping, the context has been further spun into user-orientated services through the rise of what Eisnor (2006), one of the founders of www.platial.com, called in 2006 'Neogeography'. She defined this as '. . . a diverse set of practices that operate outside, or alongside, or in a manner of, the practices of professional geographers'. This obviously addresses the fact that it is now possible for users other than professional geographers, geographic information scientists and cartographers to create their own map content, and this has the potential to broaden the domain of interest and applications quite radically. Rather than making claims about scientific standards, methodologies of Neogeography tend towards intuitive, expressive, personal, absurd, artistic or maybe just simply idiosyncratic applications of 'real' geographic techniques (Turner 2006). This is not to say that these practices are of no use to the cartographic/geographic sciences – indeed they clearly are as this article will emphasise – but they usually do not conform to the protocols of professional practice (Haklay *et al.* 2008).

Although we see this as a key to the renaissance of geographic information (Hudson-Smith and Crooks 2008), the term Neogeography is perhaps of its time, in a similar manner to that of 'Cyberspace' which is now rarely used. Its importance is the trend towards the immediacy of data use in map terms without worrying, or indeed caring too much about standards. Indeed, the concept is very much in the spirit of Web 2.0, which is more than a set of 'cool' new technologies and services, important though these are. It has, at its heart, a set of at least six powerful ideas that are changing the way people interact digitally. These have been crystallised by Anderson (2007) from an earlier statement of principles from O'Reilly (2005) and as these six ideas overlap considerably in terms of geospatial information, we will list them here:

- Individual production and user-generated content
- Harnessing the power of the crowd
- Data on an epic scale
- Architecture of participation
- Network effects and finally
- Openness,

all of which pervade discussion in the rest of this article. Indeed the overlap of geospatial computation with Web 2.0 is so noticeable that the term GeoWeb is more opportune and timely than Neogeography.

The third concept that lies at the heart of Web 2.0 is the mashup in general and the convergence of this idea on the 'map mashup' as its exemplar par excellence (Butler 2006). Originally the term was used to describe the mixing or blending together of musical tracks and is seen in its quintessential form in DJ DangerMouse's 'The Grey Album'. The term now refers to websites that weave data from different sources into new integrated user services as first noted by Hof (2005). In many ways, the GeoWeb and mashups are synonymous, with a style of programming that increasingly no longer requires raw datasets in, for example, .csv format which can be displayed online via a map with other data sources overlaid. This is the power of mashups and we would argue that it is the true definition of Neogeography in a world where little or no programming is required to visualise information spatially. Such a prospect is only just beginning to exist with the period from 2005 to 2009, a time of data mashups with the GeoWeb still firmly in the hands of those who were able to code, run their own servers and engage in all the arcana of professional computation. Although the prospect exists of truly interactive map creation for people who know nothing about the web and its data display, current creators of such mashups need know their way around XML, cartographic map projection and geographic information systems functionalities at a high level. Despite all the talk of Neogeography, it is only now in 2010 that various data sources are coming on-stream with tools to enable immediate and direct visualisation.

In the article that follows, we make the point time and again that the current state of map mashups is extremely primitive in terms of the potential for non-expert users to create their own map content and tailor their cartography to very specific applications. Although the rate of change in this field is dramatic, almost everyone who tries their hand at creating a map produces a new form of application that defies classification. The field is inchoate and shows little sign of convergence. We realise that in professional geographical circles, the crediting of Google in creating a new era in spatial information is controversial. Yet it does need to be stressed that the majority of the features developed so far in map mashups are simply tools for the display – visualisation – of basic or simply derived geographic information. They do not provide any of the complexity of spatial analysis *per se*, merely the visualisation of spatial data whose financing is under-pinned by income generation through advertising. This is symptomatic of Web 2.0 but there is rapid change in that users are beginning not only to create more sophisticated maps tagging the type of information contained therein but defining where such information is produced, who uses it and at what time it is created and applied. Such extensions are fast-becoming killer applications that promise to move this field, as Hudson-Smith (2008) suggests, well beyond the domain of visualisation.

In this article, we will first attempt to define the variety of map mashups, suggesting we need a typology while

arguing that such a typology is premature. We will then note standards and formats, moving on to questions of security, privacy and confidentiality. We follow this by introducing the key issues of the creation of spatial data through mashups by harnessing the power of the crowd: crowd-sourcing, which has been called by Goodchild (2007) volunteered geographic information, and then to ground this presentation in applications, we explore some classic exemplars from our own and related work with which we are most familiar. We then point to extensions to other graphical media, to 3D, to virtual worlds and beyond. Finally we speculate on what all this might mean for GIS and geographic information science.

A typology of map mashups

It is almost impossible to now count the number of map mashups that have been developed since Google released its API in June 2005 for it would appear that our ability to mix and match different data and software for different applications depends on a wide spectrum of programming skills that has no coherence with respect to how such mashups are produced. MacDonald (2008) suggests that in August 2008, there were 1740 spatial mashups (see <http://www.programmableweb.com/tag/mapping/>) and in February 2010, this site suggests that the number has risen to 2153. Although the sourcing of this information is not clear, it is all we have. What is however clear is that so far there have not been any attempts to classify these and thus we must be content with sketching a rudimentary typology based not on technical structures but more on patterns of usage and practice which invoke the six principles noted above.

This inability to classify relates specifically to the fact that most mashups do not follow the classic pattern of blending two or more sources of software. In fact, the majority blend software and data, possibly adding a little personalised code or scripting that makes the mashup distinct. In fact, this domain is quite well structured in macro terms. It consists of a few major players ranging from Google (Maps), Microsoft (Bing Maps) and Yahoo (Maps) to agencies such as the UK Ordnance Survey (OpenSpace) and freeware initiatives such as OpenStreetMap (OSM) (OSM 2010) who provide platforms, APIs which let users customise their own content without giving access to source code, content in the form of data, and often advice, all the way to individual map mashups that focus on applications such as crime mapping, property and so on which often use the products of the main players. Within this range, there are specific software products such as our own GMapCreator (which lets users produce tiles to be layered on Google Maps) and the portal that we use to embed it which is called MapTube (CASA 2010a). GeoCommons is another such portal which 'delivers analytics through maps' (see <http://www.geocommons.com/>) locking the user into a sequence that enables them to add a little more GIS functionality than

most. What is urgently required is some map of this terrain so that potential users can at least navigate the massive number of possibilities that are now on offer. We do not see this paper as comprising that map for the field, as it must develop a little more before the sort of structure that is clear to a multitude of possible users is clarified.

To summarise how we see the field now and before we note the details of some of the main developments, we will characterise the spectrum of types in the following fourfold categorisation:

- Basic portals that provide maps as a backcloth which users can customise either directly within the web page or using elements of the API software that such portals often provide. The obvious example is Google Maps with its various customised products such as MyMaps.
- Applications immediately built on the basic platform using the methods provided through the portal or platform. These represent by far the largest class of mashups.
- Secondary Software, often called middleware, which provides methods for tiling or layering the backcloths provided by basic portals, as for example with GMapCreator in MapTube for tiles or MapMaker which Google themselves provide for adding vector data to Google Maps. To an extent these software developments are being extended to provide GIS functionality although this is early days with respect to such analytics.
- Basic software, sufficiently different from the basic portals above, that lets users themselves create content, as for example in OSM which is based on crowd-sourcing.

All mashups can be seen as some combination of these four variants which are all tempered by the way they are used. We do not include in this paper the parallel set of web mapping services that are largely built around internet GIS (for these are now less fashionable anyway) or the various open source and/or free GIS products that are increasingly available online. Most of these latter products in fact require more professional GIS and programming skills than the map mashups that we focus upon here.

With respect to the basic mechanics that power the portals providing functionality for users to create their own maps, the APIs for 2D web-based maps that sit on HTML pages fit into two groups: lightweight Javascript-based APIs (such as Google Maps or OpenLayers) and those based around a more complex technology such as ActiveX, Silverlight, WPF or Flash (which are used in Microsoft's Bing Maps and Yahoo Maps). The 3D maps, which we will come to much later in this paper, are always built using a more complex technology due to the overhead of 3D rendering (as for example in Google Earth and the

Google Earth Flash Plugin, Microsoft's Bing Maps, Silverlight and WPF).

In order to display 2D map data on an HTML web page, a mapping API is used to run code (Javascript or Flash or ActiveX) on the page which gives the user the familiar ability to zoom or pan the map by clicking and dragging. This is increasingly being called a 'slippy map'. All these systems work by reducing the map to a set of tiled images. At the first zoom level, a single 256×256 pixel tile covers the whole world. At the next zoom level, there are four tiles, then 16, then 64 etc. (according to the structure of the relevant quadtree). With commercial maps from Google, Microsoft and Yahoo, the tiles are already rendered and stored at the portal as the user has no access to the raw data used to create them (e.g. the vector dataset representing the road network). With OSM, users have access to an open source of this data, so they can render their own tiles. This does reduce clutter (such as the number of colours) on a choropleth map showing important data, and it also enables the user to change the base map projection. At present, all the major 2D tiled map systems use the same map projection, namely 'Spherical Mercator' (EPSG:3785, 4326 and 900913), a UTM projection which assumes the world to be a perfect sphere rather than an ellipsoid. Although this works adequately for the most populated areas of the world, any data shown on the map above or below 85 degrees north or south is inaccurate. Some users, involved for example in weather forecasting or climate science, require data to be truly global where a polar stereographic projection would make more sense. For these applications, a custom tile renderer with a different projection could be constructed using the OSM data, or other open sources of world outline files. To date, little use has been made of custom tile renderers using a different projection, although this is being discussed for visualising environmental data.

Although the commercial maps have their own bespoke tile renderers, open sources of data require open source tile rendering software to function. A common combination is OSM data with the OpenLayers API for the map, using Mapnik or OSMarender to render the data for the map using a rules file, or style which defines how the OSM data are drawn. Prior to the introduction of Google Maps in February 2005, web-based maps of this type had utilised an Open Geospatial Consortium (OGC) standard called WMS for maps sent to the client as images, or WFS for maps where vector data was sent to the client. Neither of these solutions is scalable to large numbers of users due to a lack of ability to cache requests and this is implicit in the way the standards are written. Google uses a tiled map where the tiles are in fixed geographical locations for any request made by any user. This allows the caching of tiles both by the system for rendering and storing the tiles and by the web server and client browser. The OGC (2010) now has a standard called Web Map Tile Service (WMTS), which adopts the fixed location tiles implemented by

Google and adopted by all large scale web mapping systems.

It is also important to note that while OpenLayers and the Google Maps API are intrinsically similar, OpenLayers is just an open source Javascript library. The Google Maps API is also a library but is backed by other Google infrastructure. When Keyhole Markup Language (KML) vector data are overlaid on a Google Map, if these data come from another web site, then it will fall foul of a security restriction in web browsers called 'cross site scripting'. In essence this means that the browser will only show data that comes from the same site as the page that it is displaying. If the KML file is on another web site, then it cannot be displayed. Google allows this by using a KML proxy which is part of their 'free to use' web infrastructure. This does not exist in the case of the OpenLayers API as there is no such infrastructure. For our own MapTube site, we allow KML overlays on Google Maps using their proxy, whereas for the OpenLayers view, we have to use our own KML proxy developed by ourselves which runs on the MapTube web site as a web service. This allows the OpenLayers API to download any KML file on the internet by passing its address to this web service.

These technical details are important with respect to the nature of the mashup that ultimately emerges. They not only affect speed of access and size of tile that can be displayed, but also determine to an extent the presentation of what might be possible in any application. The rate of change has been so fast and so uncoordinated that the field has barely had time to pause, to take stock of what is now possible and what is the most efficient way of handling any application. In short, advising a user as to the best way forward is extremely difficult at present as there are few users and developers who have a clear view of the entire domain and certainly there is little advice as to best practice. We will return to this point throughout this paper.

Standards, formats, security, privacy and confidentiality

Standards and formats for software and data are in a state of flux in the geospatial domain, particularly with respect to map mashups as might be expected in an area that is dominated by non-expert users demanding better and easier functionality which the most expert of those users are able to provide as third parties. However, the OGC does adjudicate and recommend various standards which cover spatial data formats, protocols and structures for storing and accessing data, as well as various methods for querying, assembling and aggregating data. The basic *de facto* standard with respect to spatial data is probably the 'shapefile' which is a proprietary ESRI binary format for vector data, but it is well documented and supported by almost all GIS software. It handles large amounts of geographic data in any projection but it is not supported by OGC. They have however

recommended the Geography Markup Language (GML) and the KML as basic standards (OGC 2010).

The OGC's set of OpenGIS Standards lists 42 different standards which 164 companies that have signed up to and a number of new standards are currently being agreed. For example, CityGML is an encoding standard for the representation, storage and exchange of virtual 3D city and landscape models. It extends the GML3 schema to model 3D vector data along with other semantic data related to a city. It aims to provide a rich and extensible language for describing the features of a city. GML is an XML markup for geographic data-defining points, lines, polygons and coverages. Community specific application schemas are used to extend GML for use in a particular domain. The Grid Coverage Service (GCS) refers to data which is raster in nature rather than vector. Examples include satellite images, whether visible light or any other sensor, digital aerial photos, Light Detection And Ranging (LIDAR), elevation and terrain data. The Grid Coverage Service document defines standards for requesting, viewing and analysing raster data.

Simple Features for SQL refers to a definition for the storage and retrieval of geographic feature data in SQL databases. At present, the following spatial databases support this standard: SQLite, Microsoft SQL Server 2008, MySQL, PostGIS, Oracle Spatial, ESRI ArcSDE, Informix and IBM DB2. A Styled Layer Descriptor (SLD) is a document for user-defined symbolisation and colouring of geographic feature and coverage data. It provides users and software with the ability to control how geospatial data is visualised. Tile server software like Mapnik or GeoServer use an SLD document to define how the map tiles are drawn from the geographic feature data. In the case of thematic or choropleth maps, the SLD is extended by the Symbology Encoding specification to provide rendering of data that is not provided for in the base SLD specification. The Symbology Encoding document defines how feature and coverage data is portrayed visually on the map. This is an XML encoding using symbolisers and filters to define how attribute data is displayed.

The Web Coverage Service defines a standard interface for access to coverage data e.g. satellite images, aerial photos, digital elevation and terrain data, LIDAR or any other raster-based instrument, whereas the Web Feature Service defines a standard interface for access to vector data in the form of points, lines and polygons. Data falling within a bounding box can be queried and the raw vector data returned to the client. The Web Map Service is a simple HTTP interface for requesting maps using layers. The maps are drawn by the server and returned to the client as images (e.g. jpeg or png). The client specifies the bounding box of the map, together with the layers required and receives the map back as a single image, unlike the WMTS service which returns a set of images as tiles. Last but not least, the Web Map Tiling Service is a standard intended to improve performance and increase the scalability of web

map services through caching but WMTS is still at the candidate stage. It is modelled around the large scale tiled map systems as used by Google, Microsoft and Yahoo where requests are made for discrete map tiles which can be cached both in the server and the client browser.

The other difficult area when it comes to the web and geospatial data involves questions of privacy and confidentiality against the backcloth of security. There are many different ways of publishing maps on the web, each with their own security considerations. In the case of the raw data used by Microsoft or Google to create their maps, the data used to create the image tiles is simply not physically accessible. OSM data are free for non-commercial use under a Creative Commons licence which we note below, so it is protected in that way. It is possible, however, to write a program to request all the map tiles from the server, thus 'stealing' the information, but there is a limit on the number of tiles that you can request in a set period although this misuse of the system is explicitly covered in the terms of use that any party has to agree to when signing up to use the system. Similarly, there is the same restriction when pointing your own site at another's tile set but this also violates the conditions of use. MapTube uses this idea to point at third party tiles sets on the Internet (Internet distributed file system) but with their explicit permission.

Where vector data is passed to the client, this can only be protected by limiting the amount that can be accessed and forcing users to agree to their terms of use and copyright. This type of data includes all the navigation information e.g. 'How do I get from A to B?'. When the system shows a route, it is giving away part of its network data. KML files also have the same problem, which is why Ordnance Survey vector outlines can never be placed into a KML file and published on the web. Google's Mapplets, maps which are sometimes referred to as 'mashups from mashups' is an interesting example of a map system requiring a complex security infrastructure. Ordinary users are able to make mashups including Javascript code which can be published as a Mapplet for other people to download. This falls foul of the cross site scripting restrictions, so the code is served from Google's Mapplets engine, after being scanned for any potentially malicious code.

Most map websites do not contain much information that is useful to a potential hacker (e.g. no credit card details), but they do suffer from all the other problems associated with publicly accessible web sites. Some common forms of attack for websites generally include SQL Injection attacks, buffer overflows, or exploitation of web services with malformed requests. In the case of our MapTube website, all of these have been tried without success so far. Most web servers and other framework technologies provide a level of protection against this type of activity e.g. asp.net includes protection against SQL Injection and other malicious content entered into forms.

In terms of copyright, there is little map data available on the web or in more remote digital format that is not without some copyright restrictions that apply to who is able to use such data. Most of this copyright relates to the intellectual property rights that are ascribed to such data but this is closely bound up with cost. For example, various public agencies such as local authorities, universities and other government bodies in the United Kingdom have special licenses that are negotiated with vendors such as Ordnance Survey which make the use of the data 'free' for set purposes, once the basic licenses have been paid for. This often does not extend to web use but the picture is changing with initiatives such as the current UK government's policy of 'Making Public Data Public'.

All data shown on Google Earth or Google Maps is protected by US copyright laws. This includes any derivative products, but the license for Google Earth and Google Maps allows for non-commercial personal use e.g. websites and blogs. Bing Maps (formerly Microsoft Virtual Earth) and Yahoo Maps have similar copyright restrictions and non-commercial personal use exemptions. OSM is the exception to this, being covered by a Creative Commons Attribution-ShareAlike 2.0 license meaning that the source vector data used to make the maps is available for download, unlike the commercial data sources. This dataset has been built by processing various open sources of map data and collecting GPS track data during 'Mapping Parties' to build up road networks for world cities. Data is 'crowd-sourced'. Although this method has been used to build road networks and boundary files, there is currently no 'open' source of satellite imagery.

Creating spatial data from scratch: OSM, crowd-sourcing and geo-networking

Increasingly map mashups are being used to capture map data through various procedures involving the user community which are increasingly referred to as crowd-sourcing (Howe 2006, Shirky 2008). This form of user engagement and interaction exploits what Surowiecki (2004) calls *The Wisdom of Crowds* in his book of the same name. In fact, crowd-sourcing of map data is not the only way in which map data can be acquired but individuals have enough local knowledge to be able to sense this using various handheld devices which capture the local geometry of streets and plots as well as the wider landscapes in which they exist. We have already noted OSM (2010) as the classic example of crowd-sourced data with the simple yet far reaching aim of creating and providing free geographic data such as street maps to anyone who wants them. The project came about as the majority of maps currently used, including those online, have the sort of legal or technical restrictions on their use in creative and/or productive activities that we sketched in the two previous sections. Founded by Coast in August 2004, OSM now has over 200,000 members and is funded entirely

by donations. In essence, any map created by OSM is free to use for whatever purposes the user wishes. It can be downloaded, repackaged, used offline or customised to fit a brand or style, a feature that sets it apart from almost any other extensive geographic dataset.

The development of OSM is notably simple in its concept. Contributors take handheld GPS devices with them on journeys, or go out specifically to record GPS tracks. They record street names, village names and other features using notebooks, digital cameras and voice-recorders (see the OSM Wiki 2010). Once the event is complete, the GPS tracks are uploaded, detailing the tracks allowed to the user which are then added to the central database. Additions such as street names, type of path, links between roads etc are added via the use of notes taken on route.

This data is subsequently processed to produce detailed street-level maps, which can be published, freely printed and copied without restriction. As such OSM is the exemplar of a crowd-sourced community mapping project. Anyone is able to take part if they have a GPS unit and the desire to see their work as part of the map. However, since 2006, Yahoo have allowed OSM to use their aerial imagery to aid in the creation of maps, and to an extent, this has lessened the need for GPS traces but still requires the community effort of gathering street names and providing details of road types, road restrictions etc. An excellent recent example dates from 2007 when OSM began to use Yahoo Imagery to map the streets of Baghdad, Iraq, via remote sketching of the imagery combined with calls to participants in the vicinity to aid in refining the road layout information. Figure 1 details the layout which was completed by 5 May 2007 on all roads which are visible in the sourced imagery.



Figure 1. The OSM Crowd-sourced Map of Baghdad, May 2007 (from <http://wiki.openstreetmap.org/wiki/Baghdad>).

Haklay (2010a) of University College London (UCL) has carried out a comprehensive comparison between OSM and Ordnance Survey Meridian 2 data and has found that the OSM data is accurate to an average about 85% for street/road overlap although he has generated many more detailed estimates. He also shows that the coverage in the United Kingdom has risen from 27% in 2008 to 65% as of October 2009. His blog contains many relevant details and it is worth noting that such map mashups are being used in many emergency situations such as the recent earthquake in Haiti (Haklay 2010b).

As we also noted above, an alternative to OSM has been introduced by Google in its Map Maker service which began in June 2008. This in many ways is similar in nature to OSM as it is designed to crowd-source maps in countries where current mapping data is unavailable or sketchy. In contrast to OSM however, its licensing terms for all maps created using Google Map Maker (2010) are the intellectual property of Google. Users are able to trace features in a way similar to OSM's use of Yahoo Data. These are sketched directly onto imagery with the ability to add roads, railways, etc. through to building layouts and business locations. Data quality is maintained by ensuring contributions are moderated by more experienced users. OSM operates in broadly the same way with data quality maintained by user checking and validation. Both OSM and Google Map Maker have played a role in the mapping of Haiti in light of the January 2010 earthquake. Each system has varying levels of accuracy and as Haklay (2010c) notes, there seems to be a friction between Google Map Maker and OSM as to which organisation will ultimately prevail amongst governmental and NGO users.

Crowd-sourcing is not simply about adding map data. It is being used for adding more general socio-economic data at the individual level which has obvious spatial content that can be mapped. In this way data about current social issues can be collected directly as a means to supplement existing data sets or to create new ones. We will outline our own MapTube portal for the creation of map mashups in the next section but suffice it to say that we can use these resources to generate maps which produce near real time responses mapped spatially in the fashion of an online, near real-time geographic social survey tool.

The system has been developed through a number of custom-written examples for the BBC to collect responses to topical questions emerging from discussion on radio or TV through the provision of related web sites that let users make relevant responses. Radio 4, BBC South, BBC Look East and BBC North have all used the system to enable users to respond to specific survey questions where they are asked to specify a postcode of their location so geographic maps can be created. In the cases in question, geographic 'Mood Maps' have been created (Hudson-Smith *et al.* 2009a). The process was first used to create a mood map of the economic recession in the United Kingdom; working with BBC Radio

4 and BBC TV NewsNight, a survey was created where people were asked to choose from one of six options as to factors affecting them most during the current recession. No personal information was collected with respect to the 23,000 total responses making up the survey. We show the typical content in Figure 2 which is produced in real time on line through the MapTube portal (<http://www.maptube.org/>) constructed in map mashups using the Google Maps API.

The kind of crowd-sourcing as we have defined it in relation to creating maps or displaying maps is based on the purposive collaboration of the crowd and the map makers or map masher. In fact, with the emergence of many social networking sites such as Facebook, Flickr, Twitter and such like systems that enable millions of users to create their own data and to respond to other users of these systems through various ways of online communication, there is the prospect of tagging these profiles and responses with respect to location (Economist 2010). There are various mashups building up pictures of places from Flickr which are quite well developed and illustrated in map terms but these are from using a consistent series of tags such as geocodes which are often added after the pictures are produced or when the pictures are uploaded to the data set. A much more radical form of crowd-sourcing is to take the geo-locations from real time responses such as from the texts that define 'tweets' in Twitter if the user is willing to active the GPS sensing technology in their devices. We are experimenting at present with monitoring such data in different places developing a toolkit to replicate the ability to crowd-source data for any user in any geographic area world wide. Data can be pulled in directly from such social network sites for specific phases, locations or trends.

Classic exemplars: from GMapCreator to MapTube

Most map mashups are created by users blending a map data source which includes some geometry and their attributes such as, say, the boundaries of local authorities and some census data such as household types, with some map base that already exists in software that contains functions that let the user import the map data in question. Invariably providers such as Google, Microsoft and Yahoo make such functions available as, for example, in Google's MyMaps and most mashups are created in this way. Other users have more programming ability and are able to piece together maps and data using specific functions that they themselves develop, whereas the most organised of these systems involves developing specific functionality in relation to some map platform which many users can use, thus providing functions that the vendors of the map products in the first place do not provide. The development of these middleware functions is not usually accomplished in cooperation with the map providers but the map providers open their software through APIs which make the development of this middleware by third party providers possible.

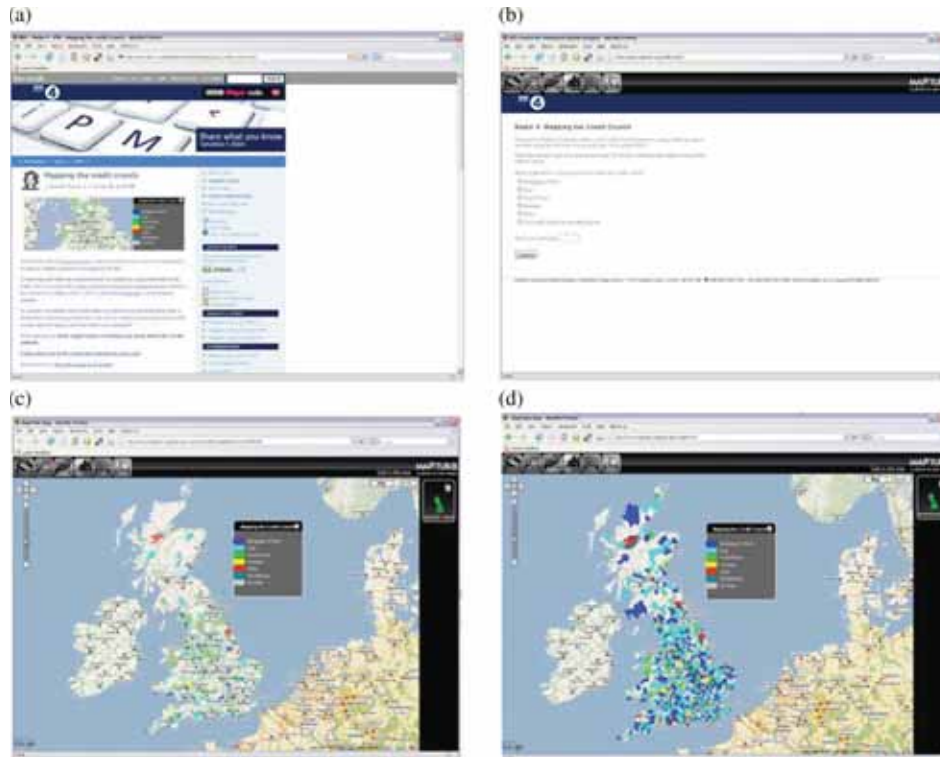


Figure 2. MapTube and Crowd-Sourcing: The Credit Crunch Mood Map. (a) Radio 4 iPM Web Page on the Mood Map for the Credit Crunch, (b) The User Web Questionnaire, (c) Early Response Distribution, (d) After 23000 User Responses.

This is the path that has resulted in the most specific applications which are tailored to particular disciplines, problems, applications and policies. The MapTube portal built by ourselves at the Centre for Advanced Spatial Analysis (CASA) in UCL essentially lets a user develop a map mashup which, if the user agrees, is indexed on the MapTube site. The mashup is made using a Java program called GMapCreator which essentially takes an ESRI shape file and converts this into a map composed of different tiles based on a user defined level of resolution. The tiles are in fact rasterised images of the original vector and attribute data which is contained in the shape file but there is now an option to import KML files into such maps. These tiles can then be overlaid onto a Google Maps base. The user in fact defines a range of scales which is then measured by the number of zooms that the user requires noting that Google Maps has 19 levels of zoom to play with (21 levels for some satellite coverage). Colours relating to the numerical attribute scales of the map are fixed by the user and GMapCreator then generates an overlay and displays this as a Google Map in a web page format. Many layers can be created in this way. The user is then asked if they would share their file with the MapTube site and if they agree, we usually define a pointer to the URL where their map they have just created is stored. If it is clear that no copyright has been infringed, then the user might store the file directly on the MapTube server. The system does not require any specific intervention on the part of the researchers who have

developed this and in this sense, it is an open resource based on freeware. The source code is not available, however, because of the experimental nature of the site and the limited ability of the developers to support users (Gibin *et al.* 2008b).

MapTube is part of the work undertaken by the Geographic Virtual Urban Environments (GeoVUE) team based at UCL's Centre for Advanced Spatial Analysis. GeoVUE was a research node of the National Centre for e-Social Science (NCeSS) funded by the Economic and Social Research Council (ESRC) to investigate how innovative and powerful computer-based infrastructure and tools developed over the past 5 years under the UK e-Science programme can benefit the social science research community. The focus in GeoVUE was on visualisation and the node has now merged with the MoSeS node at the University of Leeds to augment the visualisation capabilities based on maps with the development of spatial and geographical models that require such visualisation. Further development of the system has taken place under the 'National e-Infrastructure for Social Simulation' project (NeISS) which is funded by JISC as part of its Information Environment programme. GMapCreator has been used to develop a site called London Profiler (<http://www.londonprofiler.org/>), which takes map data for London and displays this as a series of overlays on Google Maps while also enabling the user to import other map layers into the scene (Gibin *et al.* 2008a). The maps in MapTube are in

principle available for anywhere and for anything but there is a London version which is similar to that contained in the London Profiler (<http://www.maptube.org/london/>).

We illustrate a series of London maps in MapTube in Figure 3. This option simply lets the user select all the maps for London that are in the system but of course Google Maps are world-wide and any area can be mapped. Because users of the site who put up or index the maps they create in MapTube do not in general collude or cooperate, the site presents a massive archive where a user can find other users who have created maps that might cover the same area as themselves but are unaware of. In this sense, maps are created without needing the knowledge or approval of anyone else.

As yet, there is little functionality in MapTube to engage in spatial analysis. What functionality exists is based on visual comparison and extraction of attribute data. A simple overlay facility is built in so that one can shade in and out of as many map layers as the user considers, need be active but in practice, the system works best for a couple or three layers. Apart from the crowd-sourcing that has been attached to the system for the BBC programmes referenced above, the system has not yet been blended with other real time software imports although there is plenty of scope to merge MapTube with other real time data sources. In fact this is one of our clear conclusions: potentially map mash-ups will always exist wherever two or more software and/or spatial data sources can be blended in different ways. In this sense, the user is always in control.

One last point is relevant to these kinds of middleware that support map mashups. They can be used not only for maps but also for any data that needs to be displayed in 2 dimensions and which requires the functionality that is offered by the basic map platform, of which pan and zoom

are the obvious features. Pictures and related artwork are the obvious source. A variant of GMapCreator called GMapImageCutter has been used for some very interesting displays and we note these in passing. It was used by scholars at Harvard to display ancient Greek manuscripts on the web where several levels of zoom were required and it was used by researchers in dentistry at the University of Helsinki as part of their web microscope project to illustrate the many details of a wisdom tooth extraction (Hudson-Smith *et al.* 2009b). It has been used by the Kramer collection in Cologne to show works of 'Old Masters' where zoom and pan are required. These are those we know of and there may be more but it illustrates the fact that these kinds of mashup are not restricted to maps but pertain to any spatial data that requires exploration and visualisation. As a spin-off from GMapCreator, we have developed an ImageCutter (CASA 2010b) and a PhotoOverlayCreator (CASA 2010c) for these purposes (for details see the software section at <http://www.casa.ucl.ac.uk/>).

From maps to related media: 3D, second life and beyond

The rise in computing power, specifically graphic card technology, collaborative techniques and changes in data licensing models is rapidly moving map data into the 3D environments. These were originally represented by Computer-Aided Design (CAD) but have now been extended into a range of multimedia, particularly virtual worlds and a wide range of gaming environments that are being opened up for the addition of external content. Free mapping sources based on 'virtual globes' from the biggest software vendors have also been extended to 3D with Google Earth and Bing Maps (formally Microsoft Virtual Earth) being the most widely known but with others



Figure 3. London Maps Imported into MapTube.

such as World Wind from agencies such as NASA providing ubiquitous ways to access geographic information utilising the third dimension on a global base (Craglia *et al.* 2008).

In fact the extension to 3D is a world unto itself and we cannot do justice to this here. All we will do is point the way for there is little doubt that the styles of map mashup that are now routine in the 2D mapping world are fast extending to 3D, particularly as these extensions are being pursued in parallel with links to geocoding of social networks and the delivery of all this content in whatever environments and devices are now available in which to use and view it. Of note is the fact that Google Earth has an iPhone app allowing 3D information to be viewed and overlaid with data while on the move and Bing Maps is being integrated with ESRI's flagship proprietary GIS – ArcGIS 9.3, thus allowing two and three dimensional data to be ported into a professional level geographic information system. In many ways, this is moving the entire concept of data mashups in a geocoded world forward towards the 'GeoCloud', whereby data is held and manipulated using Cloud-based services which are accessible regardless of location. Google Earth is typical of this trend where its data is now currently a mix between custom-created 3D cities via automated photogrammetry techniques and crowd-sourced models via its free SketchUp and Google Building Maker modelling applications. Google Building Maker in fact was introduced in late 2009 allowing the user to model directly on top of oblique aerial imagery using a range of simple shapes. The technique is reminiscent of the CANOMA software tool by Adobe, released in 1999, but now operating over the web using pre-defined imagery.

Google SketchUp was released in 2006 to complement the professional version which allows direct integration of GIS tools with a number of file input and export options. SketchUp is a wider ranging application than Google Building Maker and allows any type of 3D model to be created, as opposed to simply buildings, with the option of integration into Google Earth provided by the Google 3D Warehouse. Users are encouraged to model their local neighborhoods as part of a crowd-sourcing exercise to provide 3D content where automated processes are cost prohibited. The process is similar in many ways to Google Map Maker, which operates in similar terms and under similar conditions. Model submissions are reviewed internally by Google as and when the user selects the option in SketchUp that a model is 'Google Earth Ready'. The model is checked to determine if the building is 'real, current and correctly-located'. If the model passes the review process, it is added to the '3D Warehouse Layer' making it publicly viewable in Google Earth when the box in the sidebar that is labeled '3D Buildings' is checked. So far users have been encouraged to model sections of the earth via a series of 'model your town' competitions where Google exhorts the user to 'Show your civic pride (and maybe win a prize) by creating a 3D portrait

of your community and sharing it with the world. You have the power to get your town on the map – and there's no bigger map than Google Earth' (Sketchup 2010). Such an approach is typical of crowd-sourcing, although with more stringent terms and conditions than OSM but with a much more focussed and controlled aim in mind. Whether we can include these as map mashups takes us to the very edge of our interest here but at least all this is representative of new ways in which non-expert users can create their own geographical content for their own use.

In fact, any user can import map data into the 3D environment of Google Earth if they are able to represent their data as a KML file. There are now plenty of free plugins to do this and many professionally structured GIS systems are now able to import and export KML files. The Free Geography Tools web site contains a variety of such converters not only for Google Earth but also for OSM and other mapping systems (see <http://freegeographytools.com/>). We have produced a GEarthCreator which enables users to convert files into KML and display them in Google Earth and we show an example of this for world GDP in Figure 4a. But we can also use products such as Google Earth directly with respect to other software where we are able to take map data and directly convert this into KML form while activating Google Earth while the original software is still running, thus exploiting the power of the 3D software to augment other software which does not have such 3D capability. An example is shown in Figure 4b for one of our land use transportation models of Greater London in which 2D data is plotted continually as the users explore the model data, outputs and predictions but also wish to see the data in 3D. A link to Google Earth enables the user to add additional data available in Google Earth from third party suppliers and compare this with the data that is exported from the user's own analysis. In this sense, Google Earth also acts as an archive for data being created from other software.

In extending our abilities to mix, match and visualise data in 3D, it is worthwhile briefly noting multi-user environments such as Second Life which was launched in 2003 with little more than a few kilometres of simulated space. It now covers more than 750 kilometres and via its scripting language (LSL), it is possible to import geographic data from a variety of sources. These so-called Mirror Worlds and the emerging ParaVerses have the potential to move how we share, visualise and communicate geographic data to a new collaborative level. We expect such systems to remain niche in the short term, indeed it could be argued the hype surrounding Second Life and its use in academia has now passed and its initial promise has yet to be fulfilled. It is our view however that the display, manipulation and communication of data in a three dimensional collaborative space, including the current examples such as Second Life, are worth pursuing for academic usage. We show a demo of our porting map data in 2D and 3D into Second Life in Figure 5.

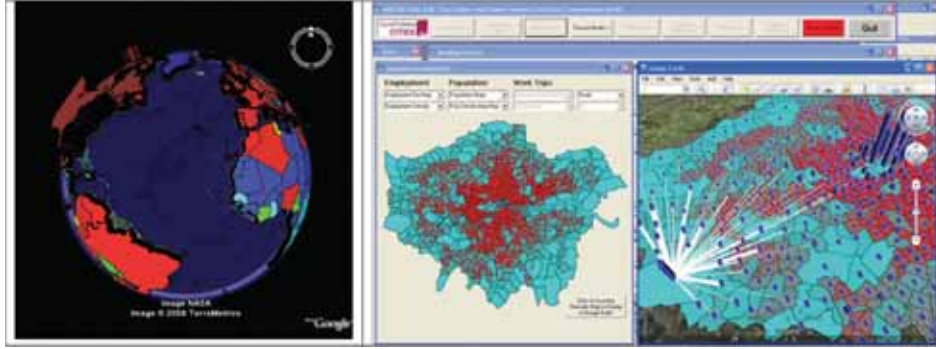


Figure 4. Three-dimensional Mashups Using Google Earth. (a) Conventional Import of a KML file of GDP (b) Exporting 2D Thematic Maps from a Land Use Transport Model into Google Earth.

The revolution for GIS and beyond

The potential audience for map mashups is extremely wide and diverse as is clearly apparent from this summary of the state of the art. So far, there are few services in place to advise potential users especially at the novice level as to how such mashups might be achieved and much depends on the ingenuity of the user in searching the web. Over the last 5 years, Google Trends (2010) indicates that the number of searches involving the word 'map' worldwide has risen by 200%, with domination by the English speaking world and Europe quite clear. In terms of the word 'mashups', the rise in the use of this term is about 40% per year but this looks as though it is now rising at a decreasing rate. This is likely to change as maps and mashups themselves become more internationally based and new data sources for other parts of the world emerge. We do not have good data on the absolute number of any of these types of search with respect to maps but the JISC Geospatial Working Group state in their vision statement that 'In 2004, map-based searching was the top activity online in the United States' (JISC GWG 2010).

What we have not discussed here is the proliferation of open source GIS software or freeware GIS. Many such systems now exist and there are many tools such as GeoTools for example (see <http://www.geotools.org/>) that allow professional users and programmers to embed functionality in their own programs or devise strings of modules that can be used to enable more focussed tasks than those that are accomplished in more generic systems. Moreover, increasingly tools are available that bypass traditional GIS software. For example, the mashup that we illustrated above in Figure 4b which is based on a visually driven land use transportation model for Greater London contains rudimentary mapping capabilities designed specifically to quickly and continuously illustrate spatial data being manipulated and produced throughout the modelling process. None of this visualisation has any of the basic features such as pan and zoom that are standard in GIS or computer cartography. But by linking this directly to Google Earth through the import and export of KML files, then all of these capabilities and more can be easily grafted onto the system.

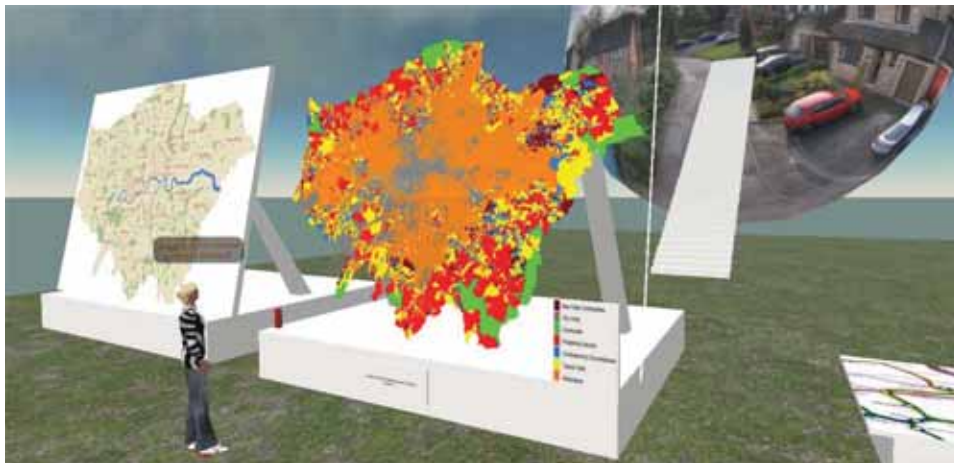


Figure 5. Importing Mashups Made Using GMapCreator into Second Life.

Although these more technical developments in which users can now develop their own systems from many basic modules which are effectively free and in the public domain will change GIS and its science, it is the more basic access to mapping capabilities that we have in mind here. In fact, we do not believe that any of these developments will undermine the continued development of professional GIS. In fact it is more likely to present new technical and scientific challenges to the community thus adding to the panoply of possibilities for visualisation of 2D and 3D maps across many different domains, fields and applications. Although the field is changing and without doubt the structure of the GIS industry will change with this, the tide is still rising and the field is still expanding out. Map mashups which have formed the subject of this article are likely to merge into the general background of Web 2.0 technologies and beyond. We have not speculated directly on the emergence of Web 3.0 but if Web 1.0 was about extracting information and Web 2.0 was about creating it as well, Web 3.0 will be about adding intelligence to this process. By this, it will be ever easier to create a map and the stages that users now have to go through will become almost entirely automated but at the same time will be fit for purpose. In a sense, this is no more or less than the semantic web that Berners-Lee always had in mind (see reference to Wikipedia 2010).

A related focus for Web 3.0 is ‘the location-aware and moment-relevant Internet’ which will have profound implications for GIS in that this focus will force the field to grapple with questions of time, which in terms of the science are still largely absent. In short, it is likely that the developments reported here are likely to move the scientific field of GIS towards dealing with space-time data in ways that are quite different from the simple temporal layer approach that still dominates methods of dealing with time in terms of maps. The development of locationally aware devices and the sampling of temporal data in real time, as for example in the Twitter feeds that are now being captured, is likely to produce new ways of visualising space and time. It is entirely likely that we will see many more animated sequences of spatially varying activities in and outside of real time and in the near future temporal map mashups will become the focus, with dramatic implications for GIS.

References

- Anderson, P., 2007. *What is Web 2.0? ideas, technologies and implications for education*. Horizon Scanning Report, JISC Technology and Standards Watch. Available from: <http://www.jisc.ac.uk/whatwedo/services/techwatch/reports/horizonscanning/hs0701.aspx> [Accessed 31 January 2010].
- Butler, D., 2006. Mashups mix data into global service. *Nature*, 439, 6–7.
- CASA, 2010a. Available from: <http://www.casa.ucl.ac.uk/software/googlemapcreator.asp> [Accessed 31 January 2010].
- CASA, 2010b. Available from: <http://www.casa.ucl.ac.uk/software/googlemapimagecutter.asp> [Accessed 31 January 2010].
- CASA, 2010c. Available from: <http://www.casa.ucl.ac.uk/software/photooverlaycreator.asp> [Accessed 31 January 2010].
- Craglia, M., et al., 2008. Next-generation digital earth. *International Journal of Spatial Data Infrastructures Research*, 3, 146–167.
- Economist*, 2010. A world of connections, a special report on social networking, 30 January 2010, 1–20.
- Eisnor, D., 2006. What is neogeography anyway? Available from: http://platial.typepad.com/news/2006/05/what_is_neogeog.html [Accessed 31 January 2010].
- Gibin, M., et al., 2008a. Exploratory cartographic visualisation of London using the Google maps API. *Applied Spatial Analysis and Policy*, 1, 85–97.
- Gibin, M., et al., 2008a. *Collaborative mapping of London using Google maps: the London profiler*. Working paper 132, Centre for Advanced Spatial Analysis, University College London, London, UK. Available from: http://www.casa.ucl.ac.uk/working_papers/paper132.pdf [Accessed 31 January 2010].
- Goodchild, M.F., 2007. Citizens as sensors: the world of volunteered geography. *GeoJournal*, 69, 211–221.
- Google Map Maker, 2010. Available from: http://www.google.com/mapmaker/mapfiles/s/terms_mapmaker.html [Accessed 31 January 2010].
- Google Maps, 2010. Available from: <http://maps.google.com/> [Accessed 31 January 2010].
- Google Trends, 2010. Available from: <http://www.google.com/insights/search/#q=google+maps&cmpt=q> [Accessed 31 January 2010].
- Haklay, M., 2010a. Available from: <http://www.slideshare.net/mukih/beyond-good-enough-spatial-data-quality-and-openstreetmap-data> [Accessed 31 January 2010].
- Haklay, M., 2010b. Available from: <http://povesham.wordpress.com/> [Accessed 31 January 2010].
- Haklay, M., 2010c. Available from: <http://povesham.wordpress.com/2010/01/18/haiti-how-can-vgi-help-comparison-of-openstreetmap-and-google-map-maker/> [Accessed 31 January 2010].
- Haklay, M., Singleton, A.D., and Parker, C., 2008. Web Mapping 2.0: the neogeography of the geospatial internet. *Geography Compass*, 2, 2011–2039.
- Hof, R., 2005. Mix, match, and mutate: ‘Mash-ups’ – homespun combinations of mainstream services – are altering the Net, *Business Week*, 25th July 2005 Available from: http://www.businessweek.com/@@aMrzTYQQQucvyQMA/magazine/content/05_30/b3944108_mz063.htm [Accessed 31 January 2010].
- Howe, J., 2006. The rise of crowdsourcing. *Wired Magazine*, 14 (6), 161–165.
- Hudson-Smith, A., 2008. *Digital geography: geographic visualisation for urban environments*. London, UK: Centre for Advanced Spatial Analysis, University College London.
- Hudson-Smith, A. and Crooks, A., 2008. *The renaissance of geographic information: neogeography, gaming and second life*. CASA working paper 142, University College London. Available from: http://www.casa.ucl.ac.uk/working_papers/paper142.pdf [Accessed 31 January 2010].
- Hudson-Smith, A., et al., 2009a. Mapping for the masses: accessing Web 2.0 through crowdsourcing. *Social Science Computer Review*, Doi: 10.1177/0894439309332299.
- Hudson-Smith, A., et al., 2009b. NeoGeography and Web 2.0: concepts, tools and applications. *Journal of Location Based Services*, 3, 118–145.
- JISC GWG, 2010. Available from: http://www.jisc.ac.uk/media/documents/jisc_collections/annex_d_gwg_vision.pdf [Accessed 31 January 2010].

- Macdonald, S., 2008. *Data visualisation tools: Part 2 – spatial data in a Web 2.0 environment and beyond*. DISC-UK data share project. Available from: <http://www.disc-uk.org/publications.html> [Accessed 31 January 2010].
- MapQuest, 2010. Available from: <http://www.mapquest.com/> [Accessed 31 January 2010].
- O'Reilly, T., 2005. *What is Web 2.0? Design patterns and business models for the next generation of software*. Available from: <http://oreilly.com/web2/archive/what-is-web-20.html> [Accessed 31 January 2010].
- OGC, 2010. Available from: <http://www.opengeospatial.org/standards/> [Accessed 31 January 2010].
- OSM, 2010. OpenStreetMap. Available from: <http://www.openstreetmap.org/> [Accessed 31 January 2010].
- OSM Wiki, 2010. Available from: <http://wiki.openstreetmap.org/wiki/> [Accessed 31 January 2010].
- Shirky, C., 2008. *Here comes everybody: the power of organizing without organizations*. London: Allen Lane at the Penguin Press.
- Sketchup, 2010. Available from: <http://sketchup.google.com/competitions/modelyourtown/index.html> [Accessed 31 January 2010].
- Surowiecki, J., 2004. *The wisdom of crowds: why the many are smarter than the few and how collective wisdom shapes business, economies, societies and nations*. New York: Little, Brown and Company
- Turner, A., 2006. *Introduction to neogeography*. O'Reilly, PDF Publication. Available from: <http://www.oreilly.com/catalog/neogeography/> [Accessed 31 January 2010].
- UpMyStreet, 2010. Available from: <http://www.upmystreet.com/> [Accessed 31 January 2010].
- Wikipedia, 2010. Available from: http://en.wikipedia.org/wiki/Web_2.0 [Accessed 31 January 2010].