Dissemination and 3D Visualisation of Hampshire Aerial data

Final Report – September 2015

Jason Sadler + Homme Zwaagstra, GeoData

Project Outline

Hampshire County Council released 100GB of high resolution aerial photography and height data covering the whole county, under the Open Government Licence (OGL) and in various raw formats. The full data set was available to copy for processing and local use, but there were no resources to create an 'official home', despite there being wide interest in the data for research, survey or leisure purposes, with many potential applications and considerable public interest within environmental and geospatial sectors.

GeoData has particular expertise and a track record in the storage, management and delivery of open geospatial data, a high profile and expanding field with relevance to RCUK and H2020 research infrastructure priorities (e.g. Research Infrastructure, Open Research Data).

This project led to a new research collaboration with the Council (via the Hampshire Hub) to develop and host (within the University's Open Source Geospatial Lab) an online portal for disseminating readily accessible versions of the data through open source software and open standards for geospatial data and metadata. We incorporated leading edge visualisation technologies, enabling users to explore high resolution 3D representations of Hampshire using any modern standards-compliant web browser.

The collaboration with Web and Internet Science researchers in ECS promoted sharing of expertise and research leadership in semantic web and linked/open data disciplines, with FSHS expertise in geospatial data management and visualisation skills. The project has helped foster these relationships, leading to several substantial grant applications, detailed below, and linkages with the Southampton Web Observatory project and portal.

The portal and visualisation technologies, developed within the project have also been directly applied and further developed within several other key research projects. As well as highlighting the impact of this work, additional developments funded by these other projects are currently being fed back into the Hampshire Portal, providing ongoing value for money and sustainability for the portal as well as a sense of momentum to the collaboration with the Council's Research and Innovation unit (Hampshire Hub).

Activities and Outcomes

Initial efforts involved building a research collaboration with Hampshire County Council's Hampshire Hub Partnership team to develop the aims of the project, define useful functionality and explore how the outputs may best be exploited mutually. This relationship developed very positively and has led to several substantial grant applications, building directly on the concepts developed within the project.

The majority of technical effort was spent developing tools to process the raw data and producing a series of working prototypes of the web application aimed at fulfilling the data dissemination and 3D visualisation requirements of the project. This has successfully been achieved and the main project outcome, the Hampshire Aerial portal, is available at http://hampshire.geodata.soton.ac.uk. Some illustrative screenshots are included in Appendix I.

The re-use of the technologies on other funded projects will allow us to continue to improve this in collaboration with the Council, and they will shortly be linking directly to the portal from their own Open Data catalogue, for example from commonly used Area profiles.

Meetings were also held at various points during the project with the ECS Co-I (Simperl) and with some of her co-researchers and PhD students, developing discussions of the Web Science and Linked Data elements

of the project, identifying ideas for relevant grant applications and to feed ideas into other ongoing research within ECS.

Technical Architecture

The underlying goal driving the development of the portal software has been to create a scalable and generic 'micro service architecture' and tool for the dissemination of raster datasets to many concurrent users. To be scalable it needs to be deployable on a range of hardware whilst not swamping the available resources, which is a danger when processing large datasets in parallel. A detailed description of the technical architecture is provided in Appendix II.

Expenditure

The project budget of £11,296 was fully utilised, with 7.5 days of Sadler and 45 Zwaagstra expended, and £115 used for travel & subsistence to meetings with Hampshire and WSI events in London. Many additional in-kind days have effectively been contributed to the project, where the technologies have been utilised and further developed on other projects, with improvements feeding back into the Hampshire Aerial portal.

Meetings and Collaboration with External Stakeholders

A major success of the project has been the development of a strong collaborative relationship with Hampshire Hub Partnership, the open data group within Hampshire County Council who commissioned and provided the raw aerial photography and height data, and who fully engaged with the project. Through a series of meetings with the Hampshire team, at the University and in Winchester, the development of the new portal was been steered and its impact disseminated via the Hampshire Hub website and blogs.

As well as the grant proposals already submitted as a consequence of the project, detailed below, we continue to collaborate with the Hampshire team to disseminate the project, develop further research proposals and seek joint funding. Project meetings with Hampshire identified potential key areas for development and example use cases outlined, in particular for Horizon 2020 calls.

The new architecture was identified early in the project as the potential basis for at least 3 other active and forthcoming University research projects (see the section below), with meetings and demonstrations being held with internal and external stakeholders from each.

Initial joint discussions have also been held with a local Remote Sensing and Space technologies SME who have carried out example UAV flights over the River Test flood area near Romsey, with the potential to use the Hampshire Aerial portal to deliver ultra-high resolution imagery and 3d visualisations for a limited area, to support applications in flood prevention, warning dissemination and social machine analytics.

Channel Coastal Observatory: We have delivered this national coastal data management programme for a number of years, bringing in over £400k research funding. A new major phase is due to be contracted later this year and it is crucial that we are able to deliver to the funders expectations. Key components of the new portal have been demonstrated to the clients (3d visualisation), with other elements highly relevant to improve the existing data management website (immersive interface, download queue architecture). [potential research income £250k]

WorldPop: We have been working closely with Andy Tatem within FSHS on the processing and delivery of normalised world population data, used by many high profile projects and with existing extensions for Maternal Health (Norad + Google funded). There is great potential and desire to expand the data visualisation and dissemination element of this project, with significant World Bank and Gates Proposals in preparation. The new portal architecture is ideally suited to deliver this aspect of the project. [estimated potential £50-100k or more]

Oman Social Indicators Database: A proposal has been submitted with Faculty colleagues by invitation to the The Research Council (TRC) of the Oman Sultanate for a high-impact research project developing a Social Science data observatory for Oman, used for disseminating a range of open data for evidence based policy delivery. The portal code will be ideal for delivering the geospatial components of this project. If awarded the project will require an intense period of IT development over a 12-18 month period; the relationship with TRC is likely to be long term and there would be many potential spin-offs to other projects, both within our own area of expertise and within the Faculty in terms of PhD students, papers and further research grants. [initially £311k]

OpenPop: An Open Gridded Population Dataset for England and Wales is being developed by GeoData as part of the Open Source Geospatial Laboratory, based on the Office for National Statistics (ONS) 2011 Census. OpenPopGrid uses dasymetric techniques to distribute the 2011 Census residential population onto a 10m grid using building footprints derived from the Ordnance Survey VectorMap District dataset and the ONS Postcode Directory. The resulting dataset aims to be the most accurate estimate of the 2011 residential population distribution attainable using data published under the Open Government Licence. Our plan is to make it available as open data under the OdBL using the Hampshire Aerial portal architecture and to use the project as an exemplar impact project and to seek RCUK funding for further development or similar dataset applications.

Impact Review

The main impact of the project has been the creation of a new open data portal allowing a wide range of users to access and view data which was previously effectively locked. Several key components of the 3D visualisation architecture (Cesium Terrain Builder and Tile Service) have been made available to the open source community via GitHub and there has been considerable interest and use in the tools from a range of users and groups internationally.

The project has led to a the development of a number of successful research collaborations with various groups, described above, and in particular the development of a strong partnership with Hampshire County Council for future joint research collaborations.

Dissemination and impact related activities are ongoing, including tweets, blog posts and case study contributions to the Hampshire Hub website and presentation at a regional meeting (April 2015) of Maptime, an international network of user groups promoting open source geospatial technologies. Hampshire Count Council are preparing press release and other marketing materials for the next software release, due shortly.

The contribution of the software to other research projects, outlined above, represents a major impact of the project in terms of income, impact and staff expenditure relating to those projects.

Specific Grant Applications

In addition to the contribution the WSI funding under this project has made to the projects described above, 2 Horizon 2020 bids (Call: H2020-INSO-2015-CNECT, Topic: INSO-1-2015) have been submitted with ECS and other European collaborators, one of which includes Hampshire Hub as a partner, the other as a named pilot stakeholder. These are described below, with summary information taken from the full proposal documents.

1. iMOVE - Intelligent mobility for city life events based on open governance

PI/Co-Is: Elena Simperl, Nigel Shadbolt (ECS), Jason Sadler, Oles Kit (GeoData)

External Collaborators: Universities, Public Authorities and Companies in Italy, Germany, Austria, Spain and UK, including Hampshire County Council (Hampshire Hub) as a named partner

Outline:

Governments keep and will continue in the near future experiencing financial pressures due to the crisis, on one side, and growing expectations and demands from citizens and businesses, on the other side. They recognize that the availability of open data and open services in an open government setting improve public service design and delivery as well as enhance transparency. Specifically, councils of the cities of Berlin, Trento, Salzburg, and of the Hampshire county promote the open government approach, possess some common open assets and face common challenges, such as improving public mobility services in the context of city life events. This affects: citizens, benefiting from public transport and accurate real-time data; governments at various tiers, involved in decisions about public transport, and infrastructure decisions; and businesses, benefitting from open access to reliable and accurate data. iMOVE delivers a validated open source platform for the cocreation of fit-for-purpose, mobile, public cross-country services in the context of city life events, based on the open government and open innovation approaches, supported by a sustainable service delivery business model. The platform is extensible, robust, easy to deploy and use. It features EU level data harvesting, cloud-based, entity-centric, privacy preserving, dynamic data integration and analytics functionalities. iMOVE is multi-disciplinary, by involving social sciences and ICT domains and by addressing cross-country/cross-cultural differences arising from Germany-, Italy-, Austria- and the UK-based real-world business cases.

Objectives:

1. To create a **positive case for trust in public administrations** (PAs) by delivering a number of services promoting the open government approach in the extended mobility scenario.

2. To promote the **open innovation** approach instantiated through quadruple helix stakeholder collaborative networks as an enabler for **co-creation, selection and delivery of public services**.

3. To integrate, tailor and enhance existing technologies for intelligent data integration in the context of open government in order to provide an extensible, multilingual, **cloud-based, entity-centric, privacy-preserving, open and re-usable platform** able to process **large volumes of data**, offering real-time data analytics functionalities that enable seamless data and service re-use, and thus, creation of new **pro-active and more personalized** (based on richer user profiles) **cross-country services** that should better suit user needs.

4. To deliver **fit-for-purpose public services for city life events** through adaptation and personalization of existing ICT solutions that take into account **user experience** and **social behavior** patterns:

   a. **Mobile applications** for Android and iOS that facilitate collaboration and participation of citizens and businesses in public administrations as well as increase their transparency.

   b. **Open services** (APIs with open specifications) of the platform enabling programmatic access to its functionalities in a modular way, thereby allowing development of new (public) services.
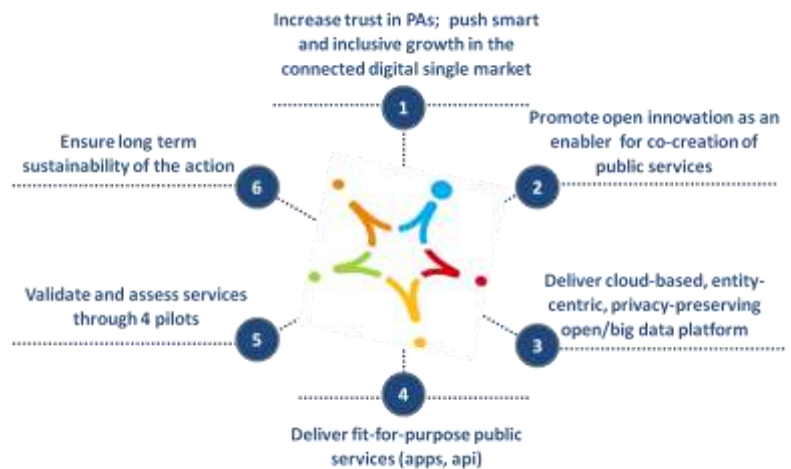
5. To validate the iMOVE platform through **4 pilots run in 4 countries for a duration of 12 months** each and to **assess effectiveness** and **socio-economic impact** of the action.

6. To ensure **long term sustainability** of the action, by establishing a self-sustainable non-profit foundation fostering exploitation of open data and iMOVE enabled services through an innovative **business model** based **public-private co-operation**.

<u>Budget</u>: Total € 2,926,062 including € 348,750 (Soton) and € 137,375 (HCC)

<u>Timescale</u>: 36 Months from early 2016 start date

<u>Impact</u>:

Impacts are identified within the proposal with respect to the work programme topic objectives:

- Stimulating the creation, delivery and use of new services on a variety of devices, utilizing new web technologies, coupled with open public data
- More personalized public services that better suit the needs of users
- Reducing the administrative burden of citizens and businesses
- Increased transparency and trust in public administrations

Hampshire Pilot exemplary impacts:

Hampshire expects to gain a better understanding of transportation challenges in the county, by considering specific development projects, and the importance of tourism. This specifically applies to a location originally identified as a possible future Ecotown the Whitehill/Bordon development, which has subsequently moved to become a Major Development Area for East Hampshire District Council. Collection of data around the new transport strategy that will form part of the major development will help the parties involved with delivering the masterplan for Whitehill to assess, monitor and possibly amend the strategy, based on the user experience in IMOVE.

Hampshire county council will exploit the results of the project to improve the public services it offers to its citizens as well as by promoting the iMOVE potential value and results, through the Hampshire Hub, to its partners and on organizations in logistics and transportation services industries. This will foster the awareness and the usage of iMOVE both in public and private sectors.

2. COMPRISE - Collective policy monitoring through personalized e-Participation services

<u>PI/Co-Is</u>: Elena Simperl, Yunjia Li (ECS), Jason Sadler, Oles Kit (GeoData)

<u>External Collaborators</u>: Universities, Public Authorities and Companies in Italy, The Netherlands, Slovenia, Portugal, Turkey, Poland and UK, including Hampshire County Council (Hampshire Hub) as a named stakeholder.

<u>Outline</u>:

Public policy issues are extremely complex due to the continuously growing amount of data and global aspects they should take into account. In this setting, enabling eParticipation of different stakeholders for monitoring the policy performance and suggesting policy directions is essential to increase transparency of policies and trust in government. Existing participation platform do not consider differences among participants concerning their prior knowledge on the political topic, cultural differences, attitudes and roles. They provide the same access to data and services to all of them and analyse outcomes in a homogeneous way.

The COMPRISE project delivers personalized services for eParticipation, aimed at collecting data for policy monitoring, suggestions and opinions on policy directions and for educating citizens on policy topics. These services will be customize on participant's profile and will make visible and understandable a vast amount of open data, opinions, and news from press along with their geographical and temporal location. Also, COMPRISE will create services for enabling e-participation through both web portals and mobile apps.

The impact of the innovation proposed in COMPRISE will be assessed on three pilots: one in the Emilia Romagna Region of Italy by extending an existing eParticipation tool called IoPartecipo+, one in Amsterdam where a large number of open energy data are available and the third in Hampshire.

The overall participation process, the awareness of participants, their ability to communicate their opinions, the influence they have on each other, the identification of opinion leaders as well as collective reputation will be analysed before and after the pilots.

COMPRISE is based on a solid consortium of Public bodies, SMEs and academic partners, leveraging the expertise acquired and tools developed in previous EU projects to come up with applicable and immediately exploitable results.

Objectives:

eParticipation tools enable bottom-up, evidence-based policy making……

eParticipation has been recently recognized as an essential component of the policy design and making process, a bottom-up activity which some1 are calling a step change in the way we think about politics and policy making.

… providing benefits to a number of stakeholders…

The benefits of eParticipation are clearly outlined in the European eaPartication Summary Report delivered in 2009, where advantages are identified by taking into account citizens, various stakeholders and policy makers.

…but deliver these services to participants with no attention to their profile.

Before collecting opinions and ideas from participants, eParticipation tools should enable access to policy related (open) data and services.

COMPRISE solves this issue by delivering a personalized eParticipation platform…

COMPRISE addresses both these aspects by providing personalized access to data and services based on the "user profile", taking into account previous knowledge, user interest, level of past participation, but also psychometric indicators. When delivering personalized services to a user, it is critical to get an accurate map of the user's "psyche". That is we need to know what kind of information the user values, pays attention to and wants to engage with. For this purpose, psychometric tools go well beyond demographic data.

…..for accessing open data and services ……

The platform conceived in COMPRISE integrates and manages large amount of open data related to the policy issue considers and provides personalized access to them, shaping visualization and communication means suitable for the user profile. Second COMPRISE integrates and extends a number of existing services (coming from successfully concluded EU projects) related to information extraction and decision support tools, but is open to the integration of additional services in an incremental way.

…through both mobile devices and web-portals.

COMPRISE conceives the access to open data and services using multi-modal devices: through mobile apps and through web-portals.

The assessment of COMPRISE is performed on three pilots….

…..also through a deep social analysis of user engagement, awareness and trust.

<u>Budget</u>: Total € 3,098,398 including € 350,042 (Soton)

<u>Timescale</u>: 36 Months from early 2016 start date

<u>Impact</u>:

As above, detailed Impacts are identified within the proposal with respect to the work programme topic objectives:

- Stimulating the creation, delivery and use of new services on a variety of devices, utilizing new web technologies, coupled with open public data
- More personalized public services that better suit the needs of users
- Reducing the administrative burden of citizens and businesses
- Increased transparency and trust in public administrations

With additional impacts set out against an additional objective:

- Policy Engagement and Policy communication enhancement

Appendix I – Application Screenshots

The development portal images below show the home page, download definition, area selection, 3D view and mobile device testing.
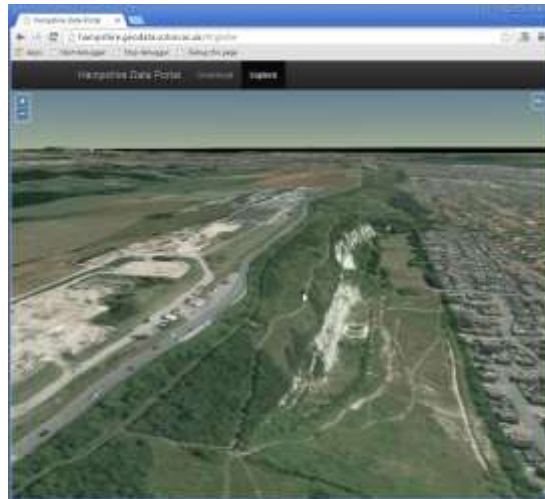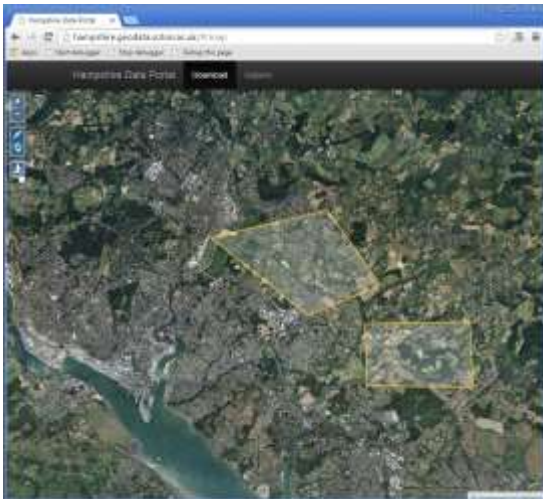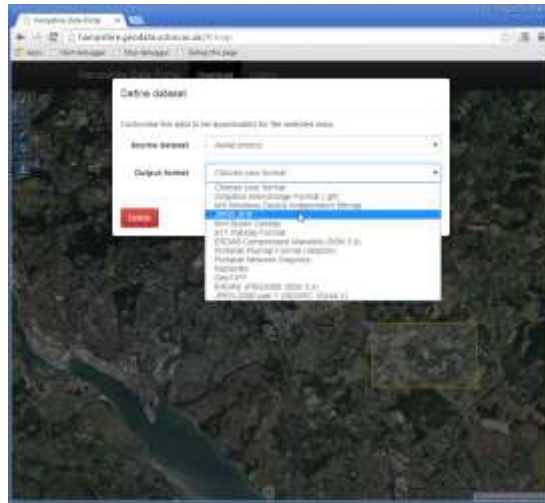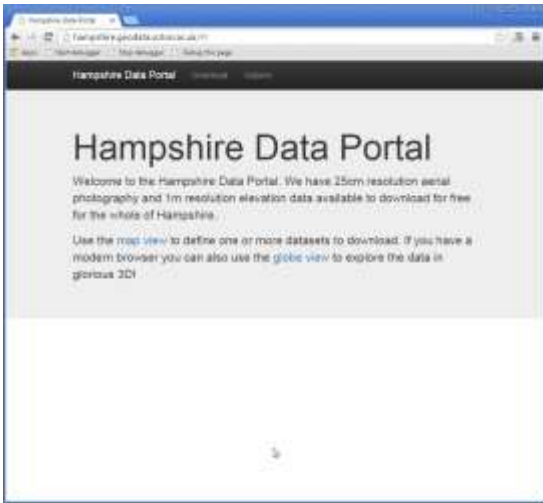
Appendix II – Technical Architecture

An overview of the new architecture chosen to deliver the Hampshire portal and 3D viewer is shown below in Figure 1.
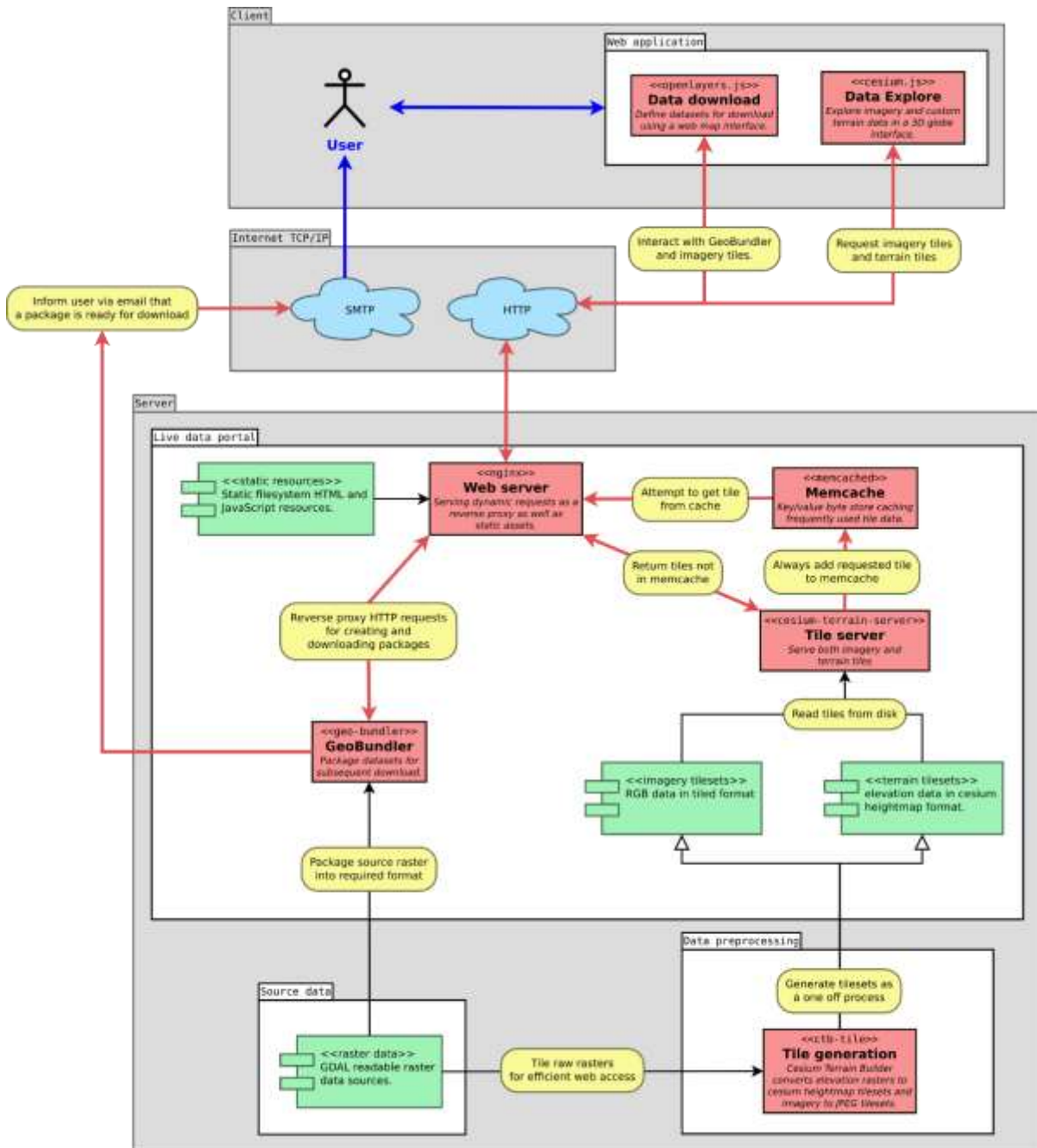

Figure 1. Overview of Software Architecture

The two key components of the application are the Data Download and the Data Explore. Both make heavy use of maps. Data Explore uses a 3d globe based on Cesium.js to interactively navigate through the data, overlaying aerial imagery on top of elevation data to create the 3d terrain model. The Cesium 3d globe architecture is based on WebGL browser standards. Data Download provides a 2d map interface to the data with controls allowing users to define datasets to download in terms of the source raster, the output format and the output extent. Each definition is termed a 'Data Window' and together the data windows form a data bundle, represented as a zip archive of the output rasters.

Both aspects of the application make heavy use of tiled resources, in terms of both imagery and terrain data. The terrain and imagery tiles are generated by in house developed software called Cesium Terrain Builder (released as open source under a previous project on GitHub) in a data pre-processing step. They are then served to the clients using a custom HTTP tile server. In order to optimise the browsing experience each request to the tile server is cached in memory using Memcached. The reverse proxy server is configured to check memcached for the resource before falling back to the tile server. In this way common tiles are always readily available in memory, minimising the need for disk I/O and considerably speeding up the browsing experience.

The client facing web server is also responsible for reverse proxying requests for generating the data bundles themselves. For scalability reasons, the application responsible for creating data bundles has been abstracted behind a distinct service called GeoBundler, providing a range of well-defined application interfaces behind a 'microservice architecture'. GeoBundler is built around a job queue paradigm in order to control resource consumption: only a limited number of jobs can be run concurrently and all others are queued waiting for previous jobs to complete. Users are notified when their bundle is ready to download via email.

Figure 2 illustrates the GeoBundler job queue and the interactions required to create data bundles.

All the programs displayed in Figure 2 are long running, persistent servers, with the exception of the bundler-admin program. Apart from the geo-bundler REST API which performs the initial bundle validation and submission, all of the servers listen to a job queue run under the beanstalkd job management software. The form a chain of processing steps with each step performing a discrete task: there is a strong analogy with UNIX pipes.

This approach allows jobs to be well managed and controlled, achieving the aim of scaling. The various steps are run as discrete services, which also provides a lot of flexibility in how they are deployed and communicate with the job queue and database.

In practice the software is bundled into Docker images and run as Linux containers on the deployment system. Deployment configuration is managed in a separate Git repository and uses the Ansible orchestration software to perform deployments. Again this makes for a flexible and yet robust and highly configurable deployment paradigm.

The web client is a Rich Internet Application (RIA) leveraging AJAX to a responsive, single page interface within the web browser. It is build around Twitter Bootstrap, a CSS and Javascript HTML toolkit that enables the construction of clean, user friendly interfaces that automatically adapt across various client devices, including mobile. This allows users to define and create bundles in one client (e.g. on a mobile device whilst travelling to work) and download them in another (e.g. on the office desktop computer).
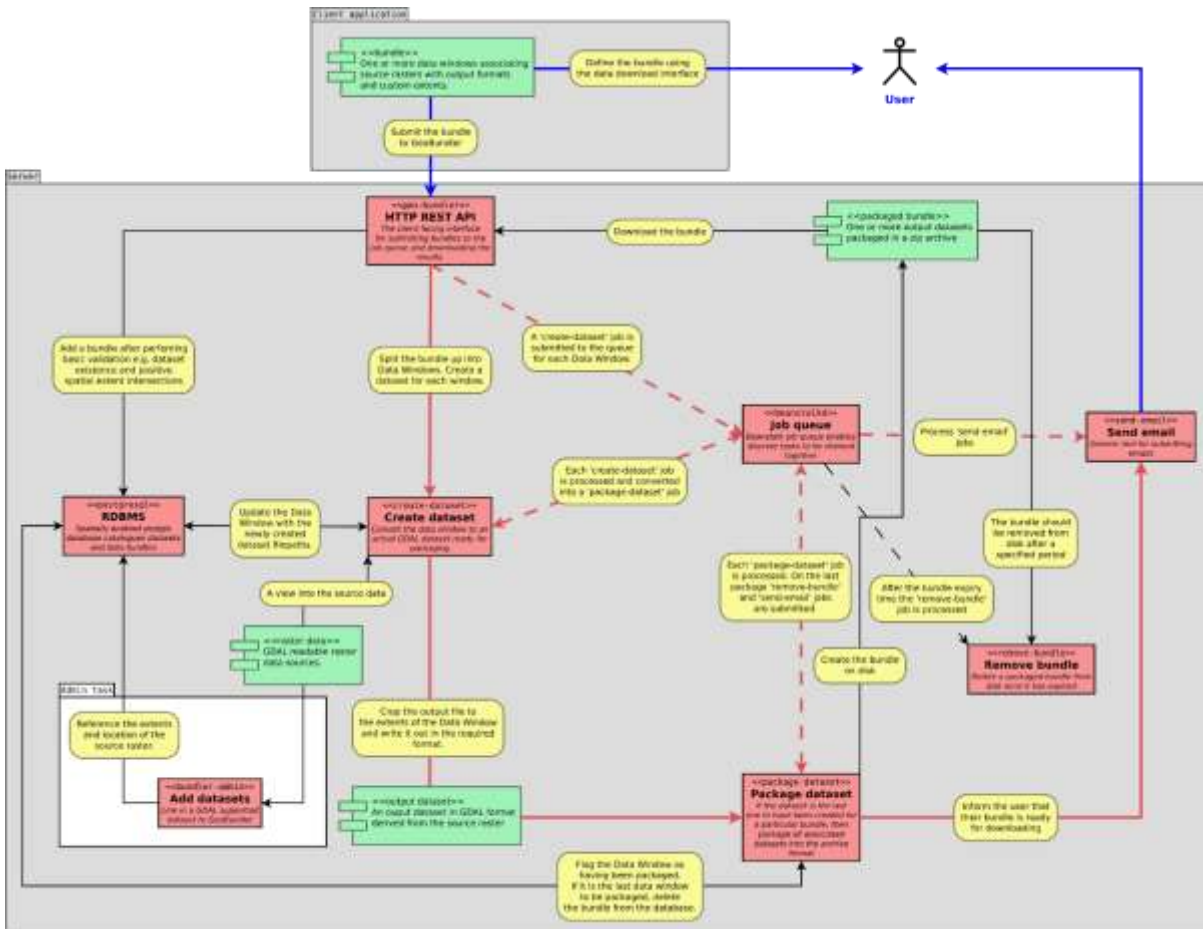
Figure 2: Overview of the GeoBundler Raster Packaging Software