Using Cloud Technologies to Complement Environmental Information Systems

Thorsten Schlachter¹, Clemens Düpmeier¹, Rainer Weidemann¹, Renate Ebel², Wolfgang Schillinger²

Abstract
Cloud services can help to close the gap between available and published data by providing infrastructure, storage, services, or even whole applications. Within this paper we present some fundamental ideas on the use of cloud services for the construction of powerful services in order to toughen up environmental information systems for the needs of state of the art web, portal, and mobile technologies. We include uses cases for the provision of environmental information as well as for the collection of user generated data.

1. Introduction
Governmental authorities in Europe are committed to actively provide environmental information for the general public (EU 2003). For that reason a lot of environmental information systems (EIS) exist which mostly provide their environmental information for web browsers at desktop computers. Only some EIS additionally offer specialized web services or even implement service oriented architectures which can be used by mobile applications. But the provision of standardized interfaces and services has become an increasingly important issue since most environmental portals and mobile apps depend on data queries and services.

From the perspective of environmental authorities and operators of EIS adding cloud technologies to their own web infrastructure can address issues like high availability, performance and scalability of services. The possibilities to re-use or combine data (e.g. data mashups) from different cloud services can add value to existing applications or even can create new ones. In order to respond to current events (e.g. disasters), generic lightweight cloud services can enable the quick implementation of required information services.

For a user of an environmental portal or a mobile application, it would be desirable to have access to the entire range of environmental information collected by authorities. Unfortunately, for different reasons not all available data are being published on the web or as services yet, and therefore cannot be displayed in portals and mobile apps.

Cloud services can help to close the gap between available and published data by providing an infrastructure for storage, services, or even whole applications. Within this paper we present some ideas on how cloud services could be used for the construction of powerful environmental services that add state of the art web, portal, and mobile technologies to existing EIS.

¹ Karlsruhe Institute of Technology (KIT), Institute for Applied Computer Science, Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany, email: {thorsten.schlachter | clemens.dupmeier | rainer.weidemann}@kit.edu
² Baden-Wuerttemberg State Institute for Environment, Measurements and Nature (LUBW), Griesbachstraße 1, 76185 Karlsruhe, Germany, email: {renate.ebel | wolfgang.schillinger}@lubw.bwl.de
As a first step common use cases will be discussed in the next chapter. Then possible implementations will be discussed by evaluating some (commercial) cloud services, in particular for the storage and provision of spatial data and maps. Finally some general conclusions will be drawn from our experiences.

2. Some Typical Use Cases of Cloud Services for EIS

Prior to the implementation of environmental laws and directives environmental information often was collected in professional systems that were not available to the public. To fulfill the obligation of an active distribution of environmental information many of these systems have been enhanced to support querying and display of data on the Web. Unfortunately, in many cases this has led to isolated solutions that don’t work together. They lack interfaces or services for the access from external systems, and if they did, most of them would not provide the performance needed for many parallel requests. Some (hopefully few) professional systems even do not have been opened to the public at all.

For all these types of EIS cloud services could be used to decouple the internal usage of data from their public provision by hosting copies of publishable data in the cloud, clearly separated from more detailed or maybe sensitive and non-publishable internal representations that should only be used internally by the authority’s professionals. For this use case casual one way synchronization of data is needed. A major advantage of this procedure is that existing professional systems do not need to be changed. They just have to offer data export functionality in order to provide the cloud service with some data excerpt. In many cases infrequent synchronization of data may be sufficient, e.g. if data change seldom, like for watercourses, nature protection areas, or floodplains.

For other types of data, e.g. current observation values such as flood levels or air quality data, frequent and prompt updates of data are of vital importance. For this use case an export (and import) of whole databases are cumbersome and impractical. Cloud services often provide an API and some mechanism (i.e. Fusion Tables in the case of Google) for the on-going update of single data sets or data fields and the underlying EIS has then to implement an interface to trigger these updates. For this use case static metadata and frequently changing measurement data can be separated within the cloud representation, so that only frequently changing measurement data attributes have to be actualized.

“Better” EIS already implement standard interfaces and services, but looking at the capabilities of their underlying IT infrastructure and the potential usage of their services by other portals and/or mobile applications at greater scale the question of their scalability and performance arises. Also for this use case, duplication of data and services in the cloud may be a good solution for such scalability and performance issues.

A last use case we like to discuss here is the (temporary) storage of data collected by mobile applications through crowd sourcing. In many cases collected data is supposed to be added to existing professional systems at a later point in time, but user collected raw data should be first kept separate until they are reviewed or quality assurance (QA) procedures are executed on them. Again, cloud services can be applied. Cloud APIs and interfaces can empower mobile apps to write their data to the cloud, e.g. structured data or media files. Respective backend APIs can be used to implement QA and data interfaces, so only little effort has to be put into the enhancement of existing professional systems. Also authorisation and account management for the crowd sourcing applications can be kept separate from normal authorisation systems used internally. Existing cloud and social network accounts can be used for identifying and engaging users.
3. Going to the Cloud

In order to evaluate the potential of cloud services for the above use cases, the Ministry of the Environment, Climate Protection and the Energy Sector and the State Institute for Environment, Measurements and Nature Conservation of Baden-Wuerttemberg have selected and licensed some commercial Google products including Google Maps Engine (GME) (Google 2013c), Google Maps and parts of Google Apps for Business (Google 2013b).

Both, GME and Google Apps for Business are cloud services running on Google’s own cloud infrastructure. It was a conscious decision to use a computing infrastructure of an external provider. In principle cloud services can also be provided by an own cloud infrastructure (private cloud) which has certain drawbacks, mainly additional efforts in the fields of operation and purchase of suitable hardware (thinking of sufficient scalability).

But using public cloud services raises questions, too. In particular, many legal and privacy issues are not yet entirely clear. For example, according to the judgment of the European Court of Justice only limited data may be transferred to the U.S., where, however, the infrastructures of the major cloud computing providers are located (Babcock 2010). Some cloud service providers address this problem by offering special license agreements assuring exclusive hosting in European data centres. But safety in terms of data availability and security in regard to unauthorized data manipulation is still very important.

4. Storing Spatial Data using Google Maps Engine

Google Maps Engine (GME) is a tool for the storage of spatial data and the display of this data as maps. Using GME it is easy to address our use cases in respect to spatial data:

- Duplication of public data from EIS without own service interfaces to a Cloud service providing good service APIs.
- Duplication of existing, non-scalable services to provide better availability and scalability.
- Easy provision of data that have not been publicly available before (e.g. due to system limitations).

For this purpose the GME allows authorized users the author to define data sources. A data source normally is generated by the upload of data. This data can combine both spatial information (geometries) and corresponding according factual data. There are a number of supported import formats, e.g. ESRI shape files, KML, MapInfo, or CSV. As another type of data source bitmaps (e.g. orthophotos) may be uploaded, too.

Uploading information is one way to get data into GME. Since early summer 2013 Google provides an own API for the manipulation of GME data (Google 2013d). Now data (e.g. assets, layers, maps) can be queried, inserted, updated, and deleted from external sources such as professional systems or sync tools.

Once data has been uploaded to a data source in GME, the author may define how this data are to be retrieved and displayed. For example, he can create filters and define styling information depending on the data and/or a zoom level. Result of this step is a well-defined map layer. Multiple map layers may then be combined into an integrated map view.

Layers and maps can then be used in different ways, e.g. as map widgets (HTML), as WMS service, as data objects for the Google Maps API and they are well integrated with other products such as Google Maps and Google Earth.

GME is a powerful tool which meets most needs of the above use cases where no frequent updates of data are needed. As a main disadvantage the strong Google-centering of the offered map interfaces (few open standards like WMS) should be mentioned.
5. **Storing Structured Data using Google Fusion Tables**

As a second cloud service we evaluated Google Fusion Tables (Google 2013a). A Fusion Table can be considered as a combination of a database table and a spreadsheet. Fusion Tables further support the definition of columns with spatial information, e.g. addresses, latitude and longitude, or KML. Fusion Tables can be joined like relational database tables.

To create a Fusion Table data can be uploaded in the style of GME, several data formats are supported, e.g. spreadsheet formats like MS Excel, and CSV. Additionally Fusion Tables can be edited manually in a spreadsheet application using a web browser (Figure 2). Furthermore they provide a data manipulation API with SQL-style commands that may insert, update and delete data.

Fusion Tables may be queried and used in a variety of representations, e.g. raw data (JSON), map layers, heat maps, KML, diagrams/charts, etc.
With this set of features Google Fusion Tables can be used to implement another set of our use cases:

- Ad-hoc storing and editing of (geo referenced) data, e.g. a (short) list or map of (located) items such as energy agencies.
- Storing and display of measurement data including frequent updates of single data sets or data fields.
- Dynamically combining (join) master data and actual measurement values.
- Storage and display of simulation data, e.g. spread of particles in a water body (Figure 3).

Again Fusion Tables appear to be a powerful tool for the handling of tabular and spatial data. But some drawbacks have to be mentioned: Fusion Tables are in beta state, there is neither a guarantee for stability of the platform nor for the existence or development of the whole product in the future. The free version of Fusion Tables has several limits concerning the number of objects (rows) and the overall amount of data they can store. We also faced limits in the presentation of Fusion Table layers in Google Maps, where simulation data was not displayed completely if too many objects were to be displayed on a single tile.
Figure 3
Simulation of the diffusion of water contamination using Google Fusion Tables and Google Maps API. (Courtesy of U. Lang, Ingenieurgesellschaft Prof. Kobus und Partner GmbH)

6. Storing User Provided Data using Google Fusion Tables

Our last use case, the caching of raw crowd sourcing data, has also been implemented using Google Fusion Tables. Since Fusion Table cells cannot store binary data (BLOBs), this type of data is uploaded to Google Drive, another part of Google Apps for Business (also available as a free of charge version). Again Google Drive provides an API for the uploading of files. Files can be referenced by an URI which then can be used to combine structured data in a Fusion Table with binary data stored in the Drive service, e.g. collected data from the observation of a certain species with the respective photos or videos.

7. Displaying Cloud Data

Once data have been made available as cloud services, they can be used in manifold ways. Using Google’s cloud services may imply the use of Google Maps as a map frontend, but data may be displayed by other frontends, too. For example, the open source tool OpenLayers (OpenLayers 2013) is able to display GME-Layers (using the WMS representation) as well as Fusion Tables (using the OSM layer class with an appropriate projection).
Of course, Google Maps already provides out of the box support for most of Google’s own cloud services. So it is really easy to add a GME or a Fusion Table layer to a map just by using the corresponding layer type of the Google Maps API and its additional libraries. Specialized layer classes like `google.maps.visualization.MapsEngineLayer` or `google.maps.FusionTablesLayer` provide easy to use integration for the respective formats.

There are different ways of including maps into own applications, varying in the degree of coupling. Automatically generated web widgets can be integrated to a web page loosely and without programming by means of an `<iframe>` tag. Using the API customized map applications may be assembled using various data sources and providing special functionality to the user (Figure 4).

Figure 4
Combination of different layer types (GME, Fusion Table, Google Weather) for the display of rivers, flood reports and weather forecasts using Google Maps API v3.

Irrespective of the used product, JavaScript-based map frontends can be used for both web applications and mobile apps. Especially when mobile apps are based on web technology (web apps or hybrid apps) they can share code with (existing or new) web applications, and thus save expenses for development.
It has to be mentioned: Even if the Google Maps API (as well as lots of other services) to a certain degree are free of charge, exceeding these limits their use is subject to a paid license. Often, however, the free versions may be sufficient to perform serious testing.

8. Conclusion and Outlook

In summary, the evaluated cloud services met the vast requirements of our different use cases. Out of the box cloud services can help to reduce load for existing EIS and professional systems, especially when additional requests from portals or mobile applications cannot be handled by these systems themselves. The evaluated services are able to handle existing data in many standard formats and grant access to cloud-stored data in other formats.

Especially to storage of data which is accessed very frequently and/or by an ever increasing number of users, cloud services offer an extraordinary availability and scalability. However, the simplicity of use is limited to their standard scope. Specific requirements for an application have to be implemented additionally, maybe with considerable expenses. In this respect, there still is need for specialized EIS and for professional applications with specific functionality for professionals and experienced users. These can be supplemented by certain cloud services, especially to satisfy the information needs of average users.

The presented cloud services are and will be used for the implementation of real life applications like environmental portals (Düpmeier 2011) and mobile applications (Schlachter 2012) providing the interested and responsible citizen with the appropriate environmental information as well as the ability to report his own collected data to the authorities.

Bibliography


