Towards an Interoperable Web Generalisation Services Framework
– Current Work in Progress
Theodor Foerster¹, Dirk Burghardt², Moritz Neun³, Nicolas Regnauld³, Jerry Swan⁴ &
Robert Weibel²
¹foerster@itc.nl
International Institute for Geoinformation Science and Earth Observation (ITC), the
Netherlands
²{burg,neun,robert.weibel}@geo.uzh.ch
Department of Geography, University Zurich, Switzerland
³nicolas.regnauld@ordnancesurvey.co.uk
Ordnance Survey Research Labs, United Kingdom
⁴jerry.swan@nottingham.ac.uk
Centre for Geospatial Intelligence, University of Nottingham, United Kingdom

Abstract
Web Generalisation Services are now receiving considerable attention from the
research community. In this context, interoperability has been identified as a crucial
beneficiary of the Web Service concept. This has inspired the formation of a working
group to establish consensus for Web Generalisation Services and to specify further
technical requirements. The intent of this working group is to enable a higher degree
of interoperability and thereby increase the possibilities for exchanging and sharing
generalisation functionality over the Web. OGC’s WPS interface specification has
been taken as a basis, since it provides a standardized means within the geo-spatial
domain to establish processes such as generalisation. This paper reports the current
status of the working group and the forthcoming work items.

1 Introduction
In recent years, different research groups investigated the possibilities of Web
Services in the context of automated generalisation processing (Edwardes et al. 2003,
Sarjakoski et al. 2005). The intent was to provide a common research platform for
generalisation and to enable automated and meaningful generalisation processing
based on Web Service technology. A first implementation of such a platform has been
published by (Burghardt et al. 2005; Neun 2007). Subsequently, Foerster and Stoter
(2006) proposed an implementation based on the Web Processing Service (WPS⁵)
interface specification of the OGC. Based on the experience with those frameworks, a
lack of interoperability has been identified as the limiting factor for automated
generalisation processing (Foerster et al. 2007; Regnauld 2007; Neun et al. 2008). The
workshop of the ICA Commission on Generalisation and Multiple Representation,
held in August 2007 in Moscow, was taken as a starting point to initiate a working
group addressing this issue.

The working group is formed by individuals from different organizations such as
research institutes, national mapping agencies, and software vendors. At the kick-off
meeting held in November 2007 at the Ordnance Survey (United Kingdom) different
goals were identified. The first goal was to establish interoperability through Web
Services in generalisation, the second goal was to increase the acceptance of Web
Services for automated generalisation outside the working group. To achieve the latter

⁵ OGC WPS website: http://www.opengeospatial.org/standards/wps
goal, a *Memorandum of Understanding* is underway, which is intended to be posted to the ICA. A special Task Force was formed, tasked with drafting and finally implementing the technical needs to achieve interoperability.

This paper reports on the current progress of the working group and introduces its main achievements to improve automated and meaningful generalisation processing in the context of Web Services.

The next section introduces the related work on Web Services within the research community about generalisation. It also elaborates on the advantages of interoperability in the context of web-based generalisation processing. Section 3 gives an overview about the activities of the working group and describes the purpose and the status of the Memorandum of Understanding. The technical issues, which have been addressed in the designated Task Force, are explained in Section 4. The last Section discusses the open issues and gives some guidelines for the future plans of the working group.

## 2 Purpose of the Project

A large number of generalisation algorithms, auxiliary data structures, cartographic constraints and measures has been developed and published within the generalisation domain in the past decades. Unfortunately these results are often only available either as isolated prototypes developed with different programming languages or they are integrated within monolithic closed systems. Thus the re-use of these generalisation tools and functionality is quite limited and often requires re-implementation. Therefore the idea of utilising an open generalisation platform has been discussed at various meetings of the ICA Commission on Generalisation and Multiple Representation (for an overview see Edwardes et al. 2007).

A first prototype of an open research platform called OXYGENE has been presented by Badard and Braun (2003). The platform was created to support the development of interoperable geographic applications based on ISO/OGC specifications. Following research concentrated on the usage of web services to solve interoperability issues in the generalisation domain. Harrie and Johansson (2003), Sester et al. (2004) and Sarjakoski et al. (2005) proposed a service-based generalisation framework within the GiMoDiG project for the harmonization of data types and the delivery of real-time data for mobile applications.

A service based interactive generalisation toolbox called WebGen suitable to use as open generalisation platform was presented by Neun et al. (2005; 2007) and Burghardt et al. (2005). Also, a classification for support, operator and processing services was proposed for the different types of generalisation services. Regnauld (2006; 2007) carried out first experiments for providing services in the WebGen framework from the Radius Clarity generalisation software (1Spatial 2007). Thus, from a technical perspective the utilisation of web service techniques for offering an open generalisation platform was shown to be feasible.

On the application side Harrower and Bloch (2006) implemented a web service prototype called MapShaper for browser based line simplification. Swan et al. (2007) developed a successful application of web services in generating schematic maps based on simulated annealing.

Foerster and Stoter (2006) presented a first generalisation service implementation based on OGC Web Processing standards. By supporting this standard, the services
offered by an open generalisation platform will be accessible also outside the
generalisation community and thereby inherit a higher degree of interoperability.
Additionally, Foerster et al. (2007) raised the issue of interoperability in the context of
WPS and Web Generalisation Services. Lemmens et al. (2006) investigated the
chaining of geographic services and recommended the usage of profiles to describe
syntax and the semantics of the service. Furthermore, Regnauld (2007) proposed the
usage of so-called geo-ontologies to formalise input data, user requirements,
cartographic knowledge, and service descriptions to build a generic system capable of
deriving generalised maps on demand.

The advantages of using a web-service architecture for an open generalisation
platform are manifold. From a research point of view re-implementations will be
avoided and generalisation functionality can be offered independently of
programming language or operating system and thus easily reused. Thus, more
complex research questions on the generalisation process can be investigated,
requiring a comprehensive set of generalisation algorithms, data structures,
constraints and measures. From an NMA perspective the availability of algorithms
published as web services would allow testing on their own data. This enables
benchmark testing and thereby comparison different generalisation algorithms which
are meant to solve similar generalisation tasks. Advantages for the software provider
include the possibility of offering generalisation independently of the particular map
production system used. New developments could be made available quickly and also
for testing purposes.

Interoperability is thus an indispensable prerequisite for a broadly applicable and
acceptable platform. Combined with an increased attention/acceptance of service
based architectures within the generalisation research community, this was the main
reason for founding this working group.

The platform is intended to provide generalisation functionality for a variety of
applications, different complexity and different kind of users. For instance, the
platform could provide expert users a means to evaluate and apply different
generalisation workflows to their data, but it could also be used to produce maps for
non-expert users. Such non-experts can access the platform indirectly through user-
friendly interfaces, such as demonstrated by a developed architecture for delivering
customized base maps (Foerster et al. 2008). Therefore, on reaching its final stage of
development, the platform should be publically accessible through the web. The
different web services will be provided at the initial stage by the different participants,
but the list of available services and functionality should later be open to anyone
interested in extending it.

3 Road Map for Constructing an Open Platform for Generalisation

At the workshop of the ICA Commission on Generalisation and Multiple
Representation (held in August 2007 in Moscow) a short session was devoted to
discussing once more the idea of developing an open platform for generalisation. It
was decided that a working group would be constituted for driving this forward. A
dedicated workshop was therefore organised at Ordnance Survey (Southampton, UK)
on the 19th and 20th of November 2007. It was attended by a mixed audience (National
mapping Agencies, Universities and GIS software companies), and considered how
such a platform could be successfully built.
After analysing the requirements of the different types of potential users of the platform, the participants of the kick-off workshop decided that the solution developed should be as generic (unrestrictive) as possible, and should be based on an international standard, preferably part of OGC’s suite of standards. A generic solution is required so that it can be used for a wide range of purposes. As the primary aim is to support research, the workshop participants wanted to impose as few restrictions as possible. As for the choice of the standard body, the GIS companies present expressed a clear preference for OGC standards, as they have a policy to follow its standards as much as possible. Therefore, it was decided to try to re-implement the WebGen prototype (Burghardt et al. 2005, Neun 2007) as an extension of the OGC standard WPS. A drafting team was formed and tasked to work on the specifications of the new WPS extension for generalisation (see Section 4).

Alongside the technical aspect, the participants also considered the problem of ensuring that the standard is adopted - the service-based open platform can only be successful if it is actually used in practice. The aim is therefore to gather some support commitments from a range of organisations, at least at the early stage. This early stage is the critical one, as it requires some investment (training plus the development of interfaces) to be able to use the platform. Once the volume of active users is large enough, the platform is expected to bring them enough benefits to justify its use. At the workshop in Southampton it was decided to write a Memorandum of Understanding (MOU) for supporting the platform. The MOU will stress the need to make the initial adoption investment, so that the platform is enriched and starts delivering benefits. The aim of this MOU will be to ensure the commitment of key organisations and to inform the broader GI-community about the benefits of using this platform. The commitment of these organisations can take several forms, such as promoting the use of the platform (for organisations like the ICA Commission on Generalisation and Multiple Representation); developing clients and/or servers for proprietary GIS (GIS software companies); enriching the platform by publishing new services (research institutes, national mapping agencies) or hosting a server.

Current activities therefore focus on developing an extension of WPS for generalisation processes (Section 4), as well as drafting the MOU. The future roadmap consists of a number of activities:

- Testing and demonstrating the use of the standard.
- Proposing classifications (operations, types of geographic features) for describing the services, which will be extensible.
- Building online facilities to support the community of users (access to the classifications, description of the services, feedback, Q&A, documentation, etc.).
- Organising tutorials to show potential users how to get started.

4 Work in Progress – Results from Zurich Meeting and Implementation Actions

In January 2008 the first meeting of the technical Task Force took place. Participants were Dirk Burghardt (University of Zurich), Theodor Foerster (ITC, the Netherlands), Moritz Neun (University of Zurich), Nicolas Regnauld (Ordnance Survey, UK), Jerry Swan (University of Nottingham) and Moritz Wittensöldner (University of Zurich). The goal of this meeting was (as stated above) to investigate the use of OGC WPS (OGC 2007) as a standard for the research platform and to draft an extension of WPS regarding the special requirements of generalisation services based on the experiences
with WebGen (Burghardt et al. 2005; Neun 2007). This draft should provide a
guideline for future client and server applications to perform meaningful
generalisation processing based on the WPS interface.

4.1 Introduction to WPS

Web service-based processing is a key requirement to provide web-based geo
information within a Spatial Data Infrastructure (SDI). After establishing
specifications for data services such as Web Coverage Service and Web Feature
Service as well as a mechanism to encode the data such as GML, standardized
processing became a demanding issue within the OGC community. This demand has
been addressed within the OGC Web Processing Service interoperability experiment
first in 2003, and led to the OGC implementation specification of WPS version 1.0.0
in 2007 (OGC 2007).

The goal of the WPS interface specification is to provide a generic means to
provide web service-based geoprocessing. It allows thereby processing any type of
process on any type of data. The communication with WPS is based on HTTP and
encoded by XML. Figure 1 depicts the three WPS-operations GetCapabilities,
DescribeProcess and Execute. The following subsection will introduce the operations
in more detail.

4.1.1 WPS Operations

The client-service communication involves three mandatory operations (Figure 1).

GetCapabilities

This operation returns the service metadata, such as name of the service, description
of the technical endpoint URLs and description of the service provider. Additionally it
returns a brief description of the available processes.

DescribeProcess

The briefly described processes in the GetCapabilities document returned by a WPS
instance can be queried in more detail by the DescribeProcess operation. It returns a
detailed description of the WPS including the different parameters and their data
types, which are required by the process in order to perform accordingly. The data
types are separated into Literal data types, which are simple ones such as Integer,
String or Double and Complex data, which can incorporate any data format.

Execute

This operation allows triggering the specific process according to the described
parameters of the DescribeProcess document and returns the result of the operation. In
this context, different modes of communication can be performed (Section 4.1.2).
4.1.2 WPS Features

Besides the above operations, it is possible to drive the mode of process execution by passing designated parameters. Those different modes allow the process to execute asynchronously. This means that the client can start a process on a WPS instance by the Execute operation and retrieve the process result by passing a process ID, which has been sent to the client upon the initial request. The process ID can also be used by the client to check the status of a started process. It is additionally possible for the client to pass the data by reference and thereby avoid physically sending the data. This improves efficiency of communication and enables service-side caching of data, by comparing the references for previously received data and using them instead.

4.2 From WebGen to WPS – Draft of WPS Extension for Generalisation Services

The WebGen platform was designed as a research platform for generalisation implementing the Web Service paradigm. The evolving WPS interface specification provides great freedom in terms of data models, processes and applications. While comparing the capabilities of WPS and WebGen as a research platform three main drawbacks of WPS were identified:

- Missing support of a service registry for WPS
- Missing standardized data models (for generalisation)
- Insufficient process descriptions

In the following subsections those three aspects will be addressed by presenting preliminary solutions.

4.2.1 A Registry for WPS

A research platform should foresee the presence of many heterogeneous and changing servers hosting different generalisation services. By taking care of such changing environments, the reliability and acceptance of such a research platform will increase significantly. As the GetCapabilities operation of WPS shows only the services of one server, a central Registry (Burghardt et al. 2005, Neun and Burghardt 2005) is needed.
at which all available services might be registered. This Registry lists all currently available services on the different servers maintained by the participating research groups and NMAs. Users of the research platform can access the registry from within their desktop application in order to select and execute one of the available services. The Registry can be queried via name of service, service authors, and service metadata.

Available services must have metadata attached, describing the service capabilities in a generic machine-readable (and comprehensible) way. Understanding the specific metadata by a computer is particularly important to establish meaningful and automated generalisation processes over the web. Therefore an operator classification as proposed by Foerster et al. (2007a) will be needed. This (semantic) metadata can be used by humans as well as computers to search for services doing a certain task. Finally it will help to build a fully automated and web service-based generalisation environment as proposed by Regnauld (2007).

The following tasks have been added to the WPS extension draft:

- A new Registry will be designed implementing a GetAllCapabilities method. This Registry is running on a central server for the research community. It collects the GetCapabilities of all participating servers. Whenever a new server is installed it must be registered at the Registry in order to be accessible through the research platform by the other users. This Registry cannot be accessed by standard WPS clients but is also not a prerequisite for interacting with the generalisation functionality hosted on distributed Web Generalisation Services.
- Each service listed by GetCapabilities and GetAllCapabilities has a metadata element containing keywords of an ontology describing the operator classification. This classification should define a hierarchical namespace concept like ica.genops.modelgen.Collapse or ica.genops.cartogen.Elimination. Further work in the research community will be needed in order to consolidate this classification.

4.2.2 A Standardised Data Model for Generalisation Services

The original WPS interface specification provides great freedom in terms of parameters and data. The specification imposes no restriction regarding the kind of format to use, neither for the input nor the output parameters. There might exist WPS instances working with vector data encoded as Shapefile or encoded as GML data or even services working with raster data (encoded for instance in HDF or GeoTiff format).

However, for use as a research platform a certain set of data formats and data models should be defined in order to increase interoperability and ensure that different services and clients are compatible. Furthermore, the future addition of more advanced and complex data structures such as graphs (Neun et al. 2008) or cartographic constraints (Neun 2007) should be foreseen with placeholders. Therefore the Task Force specified a model classification as a starting point, which will be added to the WPS extension. The following list gives an overview of the data types that should be supported within the research platform. It is important to note that the two main categories literal data type and complex data type are taken from the WPS
interface specification which distinguishes between simple (literal) and complex parameters.

- Literal data types (as parameters and attributes) such as String, Integer, Double, Boolean, Date/Time.
- Complex data types:
  - Geometry (encoded as GML2)
    - Point, LineString, Polygon
    - MultiPoint, MultiLineString, MultiPolygon, GeometryCollection (mix of geometries)
  - Feature (encoded as GML2)
    - collection of attributes (literal, complex e.g. Geometry)
    - FeatureSchema describing the attributes
    - optional metadata (e.g. semantic keywords from local or global classification of geographic features)
  - FeatureCollection (encoded as GML2)
    - Collection of features having the same schema
    - optional collection attributes (literal or complex, e.g. constraint or symbolisation)
    - optional collection metadata (semantic keywords from local or global classification of geographic features)
  - List
    - universal list structure
    - can contain any literal or complex type (e.g. Features)
  - Map
    - universal map structure with key value pairs
    - can contain any literal or complex type (e.g. Features) as value
  - Placeholders for more advanced data types:
    - Constraints (Neun 2007)
    - Trees & Graphs (Neun et al. 2008)
    - Symbolisation (e.g. Styled Layer Descriptors)
    - MesoObjects (Feature values and references)

4.2.3 Extension of the WPS DescribeProcess

The majority of the generalisation service instances process single features or collections of features as input and output data. In the DescribeProcess operation of a WPS the input and output parameters are specified in order to be read by the calling client. The WPS extension defines the application of GML2 as a default encoding for single features and collections of features. Such features will then be passed as complex data as shown by Foerster and Schaeffer (2007).

In WPS it is quite difficult to describe a specific schema and therefore only generic schemas have been established for WPS. This approach is not sufficient to describe the required feature (with its attribute) for a specific generalisation process.

In WebGen (Burghardt et al. 2005; Neun and Burghardt 2005), this has been accomplished by defining the attribute schema of input features in a bespoke (but customizable) manner. This allows the specification of a process that operates (for example) on a collection of features with an attribute ‘geometry’ of type ‘Polygon’ and an attribute ‘class’ of type ‘String’. Thus, the complexity of making input data

---

6 Such as: http://schemas.opengis.net/gml/2.1.2/feature.xsd
compatible with the requirements of the service is relocated from the server to the
client. A client application seeking to use that specific process must first convert local
data in order to be compatible with the process requirements. Otherwise a developer
of a process would have to be cautious about the diversity of different data inputs
making it much more difficult to implement a new process. With this approach, an
existing algorithm can be published more easily. Overall, this approach enables a
higher degree of interoperability and will lead finally to a meaningful and less error-
prone process interaction.

The following tasks have been added to the WPS extension draft:

- The XML schemata describing the complex input data (Feature,
  FeatureCollection) in DescribeProcess reference an xsd document generated
  on-the-fly by the service instead of referencing standard XML schemata like
  the http://schemas.opengis.net/gml/2.1.2/feature.xsd.
- The required attribute schema of a service is defined internally in the service
  and transformed into an XML schema upon request.
- For simplicity, instead of creating every schema on-the-fly a set of standard
  schemata will be defined. Examples would be:
  o Feature with one geometry of any type.
  o Feature with one geometry of type Polygon.
  o Feature with one geometry of type LineString.
  o …
  o Feature with one geometry of type Point or Polygon
  o …
  o Feature with one geometry of type Point and one attribute “name” of
type String
  o …
- Features and FeatureCollections can be described by optional metadata (e.g.
  semantic keywords from local or global classification of geographic features).
  Further work in the research community will be needed.

5 Discussion & Conclusion

A platform which allows the exchange and comparison of generalisation functionality
is an interesting prospect for the generalisation community. This platform will serve
as a basis for more complex generalisation processes and the sharing of knowledge
about generalisation. Additionally, such a platform might be the first step towards
enabling truly automated generalisation processing by the means of Web Services. In
the past, considerable research has been carried out in this context (Section 2). In
order to build a more suitable platform (in particular, one that also takes current
standardization efforts towards processing services into account) a working group has
been founded. This working group focuses on two different aspects:

- Increasing the commitment towards Web Services within the generalisation
  community (Section 3)
- Enabling a more standards-based and more interoperable platform based on an
  extension of the OGC WPS standard (Section 4).

Regarding the first aspect, the approach chosen is to draft a Memorandum of
Understanding that will be used to solicit and secure the commitment of the key
players in generalisation research, software development, and map production. An
early draft of the MOU exists; a more complete version is planned to be submitted to the 2008 ICA Workshop in Montpellier for discussion and review.

Also in context of the WPS extension, some issues are still under discussion. The applied operator classification (Foerster et al. 2007a) is an initial start and it is intended to keep the classification open for further changes. In the future, all participants will be able to suggest further operators, which will be discussed and finally be added to the existing classification. The namespaces approach for the used identifiers of operators and data types also remains to be discussed. The working group concluded that using the URL of the current ICA commission website as a namespace is not the ideal solution (one possibility could be to establish a designated URL such as [www.ica.org](http://www.ica.org)).

Overall, the working group still has many issues to resolve in order to establish a fully interoperable research platform, which will be beneficial to software suppliers, data producers (NMAs) and researchers, but also for applications from other domains. But we are confident that important first steps have been accomplished with the activities reported in this paper.

**Acknowledgements**

We want to thank all the participants of the working group, who provided valuable input for this project:

- Bulent Cetinkaya (General Command of Mapping, Turkey)
- Julien Gaffuri (IGN, France)
- Tapani Sarjakoski (Finnish Geodetic Institute, Finland)
- Vincent Smith (Intergraph Inc., US)
- Anand Suchith (University of Nottingham)
- Dan Warner (1Spatial, UK).

Additionally, we want acknowledge the work of Moritz Wittensoeldner, who implemented the concepts described in this paper.

**References**

1Spatial. (2007): Radius Clarity - a rule based environment for automated generalisation. [http://www.1Spatial.co.uk/products/radius_clarity/](http://www.1Spatial.co.uk/products/radius_clarity/).


