AASG Service Schema Validation

Validation of getFeature Response documents against approved Schemas using Open source software (XML Explorer)

A product of the Arizona Geological Survey

2012-16-02

Contents

Section 1: Introduction

This document provides basic information regarding the protocols and software requirements for schema validation.

This document is primarily intended for Association of American State Geologists (AASG) subrecipients subcontracted by the Arizona Geological Survey (AZGS) under the Department of Energy (DOE) contract Number DE-EE0002850 for the National Geothermal Data System (NGDS) to perform the subcontract requirement to make "at risk" geothermal data available online to promote geothermal development throughout the Unites States.

Section 1.1: What is schema validation?

Schema validation is the process of validating a document against a schema. During schema validation, one determines whether or not data has been entered into a document according to the structure defined by a specific schema.

A schema is the framework for a data model or collection of attributes that describe something. For practical purposes, schemas dictate where and how data from a dataset should be entered into a document.

Schema validation is important for the AASG State Geothermal Data project because it adds another layer of quality control and quality assurance for the data we publish.

Section 1.2: NGDS Schemas

This document is not primarily intended to describe schema validation at an abstract level. Rather, it is written with an eye for validating AASG National Geothermal Data System (NGDS) web services against NGDS schemas.

NGDS schemas are used to model the data being digitized and shared by NGDS subrecipients. In other words: when NGDS subrecipients digitize and share their data, it is expected that they share their data as **web services** in accordance with NGDS schemas.

NGDS schemas were created according to the specifications of the United States Geoscience Information Network (USGIN).

Section 1.2.1: What is a web service?

Under the client-server relationship, a **web service** is a protocol for requesting data from a server; data requests and responses made in accordance with this protocol are standardized. In other words: by using standardized data requests and responses, client software can make requests for data regardless of server configuration.

Matters are complicated slightly by the fact that “hosting data as a service” is considered synonymous with “hosting data in accordance with a web service protocol.” Likewise, data hosted in accordance with a web service protocol is often referred to as a web service.

The Open Geospatial Consortium has provided specifications for several different flavors of web service that are relevant to geographic information systems, USGIN, and the State Geothermal Data project. These include web map services (WMS) and web feature services (WFS).

For more information, see the USGIN glossary entry describing web services.

Section 1.2.2: What is USGIN?

The **United States Geoscience Information Network (USGIN)** initiative is the product of a partnership between the Association of American State Geologists (AASG) and the United States Geological Survey (USGS) created to facilitate discovery of, and access to, geoscience information provided by state and federal geological surveys of the United States.

USGIN is:

- a conceptual framework for sharing geoscience data
- a distributed data-sharing network
- a collection of open-source applications, standards, procedures, and protocols for sharing geoscience data
- web-based, distributed, open-source, and interoperable

In other words, USGIN is a distributed data-sharing network that uses open-source software and existing World Wide Web infrastructure and browsers. Anyone can access the data shared over USGIN, and any data that is shared according to USGIN specifications is automatically a part of the network.

For more information about USGIN, see the USGIN website (http://usgin.org/).

Section 1.2.3: Where can I find NGDS schemas?

NGDS schemas can be found at: http://schemas.usgin.org/schemas/

Section 2: Schema Validation

Section 2.1: What you need:

1. A computer with an Internet connection
2. A **web service** to validate against an appropriate schema
3. An appropriate **schema**
    a. Schemas for XML documents are usually stored using the .xsd file extension
    b. As mentioned in Section 1.2.1, AASG schemas are located at : http://schemas.usgin.org/schemas/
4. A software package capable of XML schema validation
    a. This document provides instructions for schema validation using XML Explorer (available at: http://xmlexplorer.codeplex.com/)

Section 2.2: Downloading an appropriate schema

Your first step in the schema validation process will be to download the schema against which you will validate your web service.

1. In your web browser navigate to http://schemas.usgin.org/schemas/
2. Click the appropriate schema
    a. The appropriate schema will usually correspond to the type of features in your web service
        i. For example: if your web service contains WellHeaders, then you will use the WellHeader schema to validate the web service
    b. If you are unsure of the appropriate schema for your web service, contact the Arizona Geological Survey
3. After the schema appears in your web browser, click the **File** menu in your web browser and click **Save Page As…**
    a. Note that some modern web browsers hide the menu bar by default; you might need to press the Alt key on your keyboard to temporarily reveal the menu bar
4. Save the schema to a directory on your computer you intend to use for schema validation
    a. XML schema documents use the .xsd file extension

Section 2.3: Performing a GetFeature Request

To validate your web service against the schema you just downloaded, you will need to submit a **GetFeature** request to your desired web service. **To perform a GetFeature request, simply enter your GetFeature request into a web browser in the same manner as you would a standard URL.**

A getFeature request is a **WFS request** – that is, a request for data from a web feature service (WFS). A GetFeature request returns an XML representation of the attributes of features in the web service.

Note: large web services can have tens of thousands of features, so an unfiltered GetFeature request can be very demanding in terms of bandwidth and system resources. When entering GetFeature requests into a web browser, it is possible to filter the results of the request by appending conditions to the GetFeature request, but doing so is difficult because the filter syntax for GetFeature requests is complicated.

Section 2.3.1: GetFeature requests and XML Documents

As indicated in Section 2.2, a GetFeature request will return an **XML representation of the attributes of features in the web service**. This indicates the following:

- A GetFeature request returns an **XML document**
- The XML document provided in response to a GetFeature request is a representation of **features**
- A feature in a web service is a cartographic representation of a real-world object
- Features are described by **attributes**
- Attributes include data such as latitude and longitude coordinates and any other information relevant to a feature

For example: a web feature service might contain fifty features representing river systems in the United States. The attributes describing each feature might include the latitude and longitude coordinates of each river system, as well as information such as flow rates, seasonal volume, and depth at the river's deepest point.

An XML representation of the above example would list each feature in the web service, as well as associated attributes, within the structure of an XML document.

A GetFeature request submitted to the web feature service in the above example would return the XML document described above.

For more information about XML, visit the following locations on the USGIN website:

- XML
- Markup language
- Element
- XML Tutorial

Section 2.3.2: A sample GetFeature Request

The following is a sample GetFeature request:

http://services.azgs.az.gov/arcgis/services/aasggeothermal/CAWellHeaders/MapServer/WFSServer?service=WFS&request=GetFeature&typeName=Wellheader

If entered into a web browser, this request will return an XML document representing all **Wellheader** features in the web service (Figure 1).

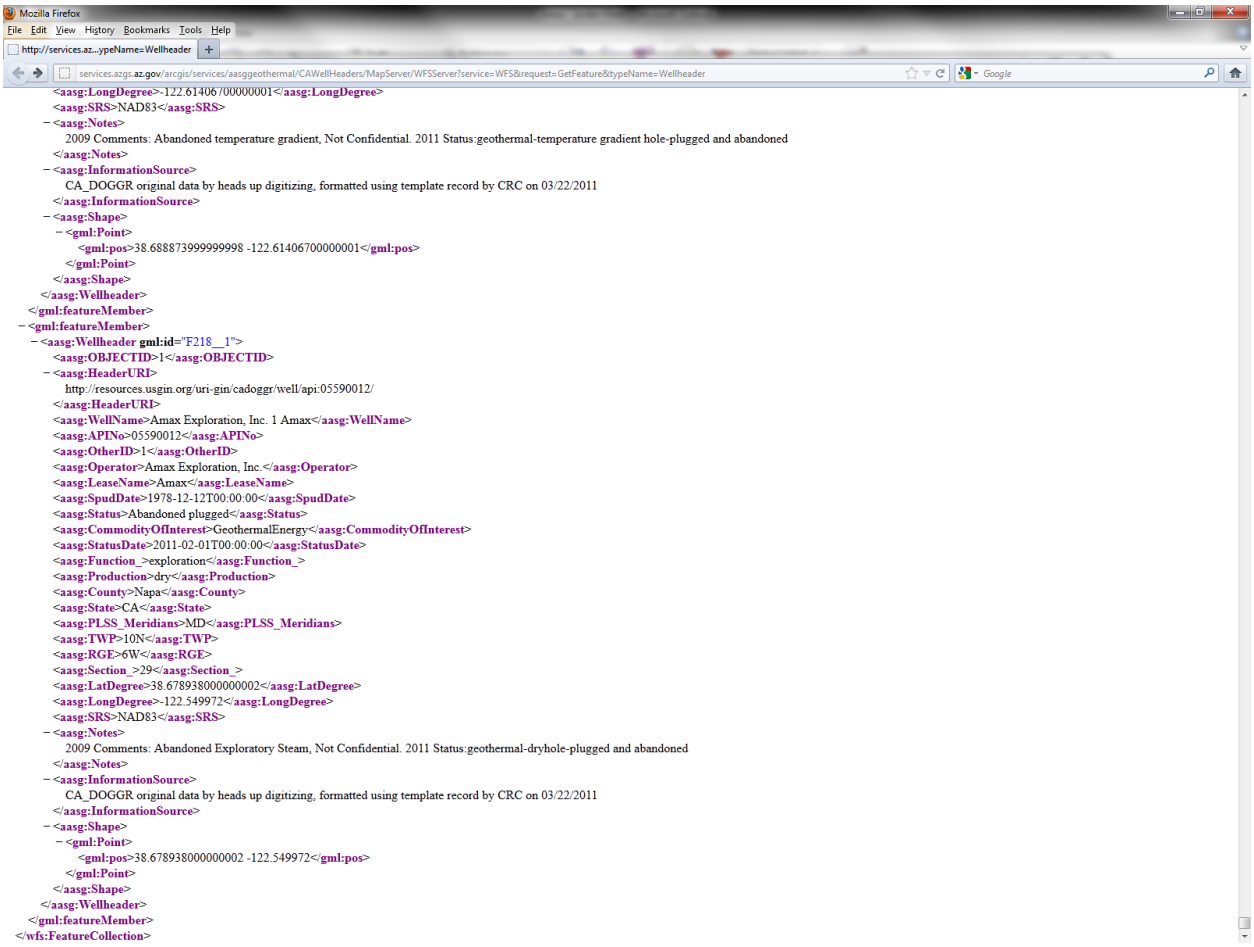

Figure 1

Section 2.3.3: Breaking down a GetFeature request

A GetFeature request can be broken down into two component parts: the **service endpoint**, and the **request** proper. These are demonstrated in the example below:

http://services.azgs.az.gov/arcgis/services/aasggeothermal/CAWellHeaders/MapServer/WFSServer?service=WFS&request=GetFeature&typeName=Wellheader

In this example, the **red** text constitutes the **service endpoint**; the **green** text constitutes the **request**.

The **service endpoint** is the web location that handles the service request. The service endpoint can be further subdivided into individual tokens, each of which can vary according to web service. A full explanation of the tokens in the service endpoint is beyond the scope of this document.

The **request** includes the desired **operation** and a collection of parameters that control the operation; some parameters are optional and others are required.

Section 2.4: Downloading the results of your GetFeature request

Having made a GetFeature request in your web browser, your next step is to download the XML response (hereafter referred to as the **GetFeature document**) to your computer and prepare to validate it. To prepare your GetFeature document for validation, you will need to modify it slightly after you download it.

1. Having performed a GetFeature request in your web browser, navigate to the **File** menu in your web browser and click **Save Page As…**
    a. This will save the GetFeature document to your computer as an XML document
    b. XML documents use the .xml file extension
    c. Note that some modern web browsers hide the menu bar by default; you might need to press the **Alt** key on your keyboard to reveal the menu bar
2. Save the **GetFeature document to the same directory on your computer to which you downloaded the schema in Section 2.2**
3. In Windows, navigate to the directory in which you just downloaded the **GetFeature document**
4. Open the **GetFeature document** you just downloaded in a text editor such as **Notepad** or **Wordpad**
    a. This can typically be accomplished by **right-clicking** the document and clicking **Open With…** in the context menu that appears
    b. Note that some text editors are easier to use for viewing XML documents than others – Notepad, for example, contains no features to make XML documents more human-readable. Consider trying different text editors to find the most optimal combination for viewing XML documents. Recommended text editors include:
        i. WordPad: basic text editor included with Windows
        ii. Notepad++: an actively maintained free-and-open-source text editor designed with XML support
5. Near the very top of the **GetFeature document**, highlight the **service URL** (Figure 2)

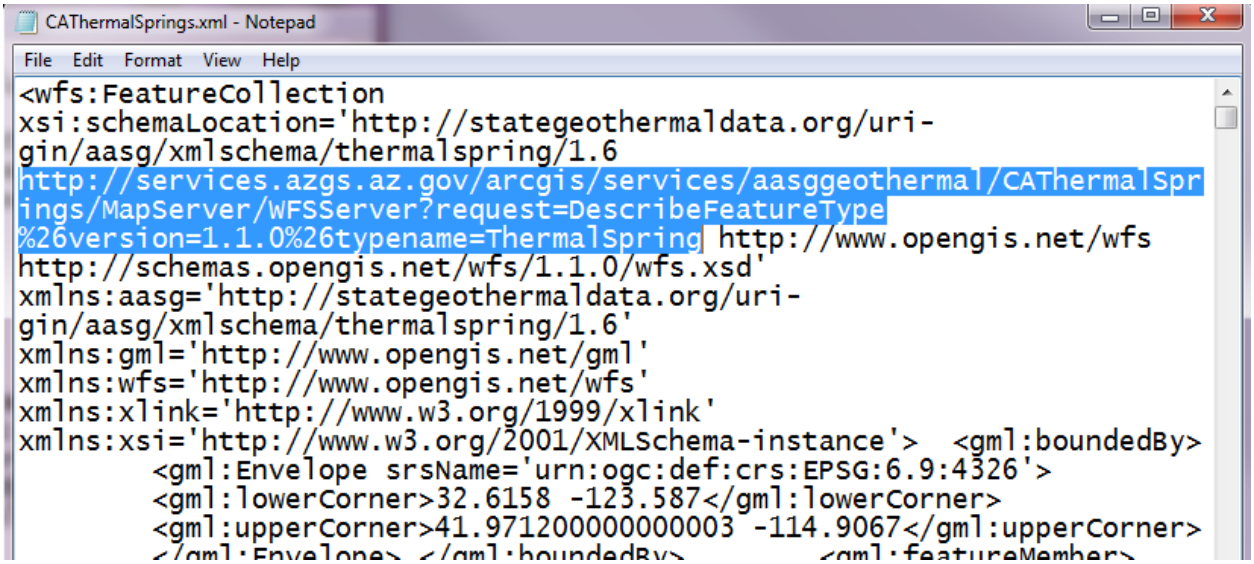

Figure 2

6. Replace the **service URL** with the exact name of the schema you just downloaded, complete with file extension (Figure 3).

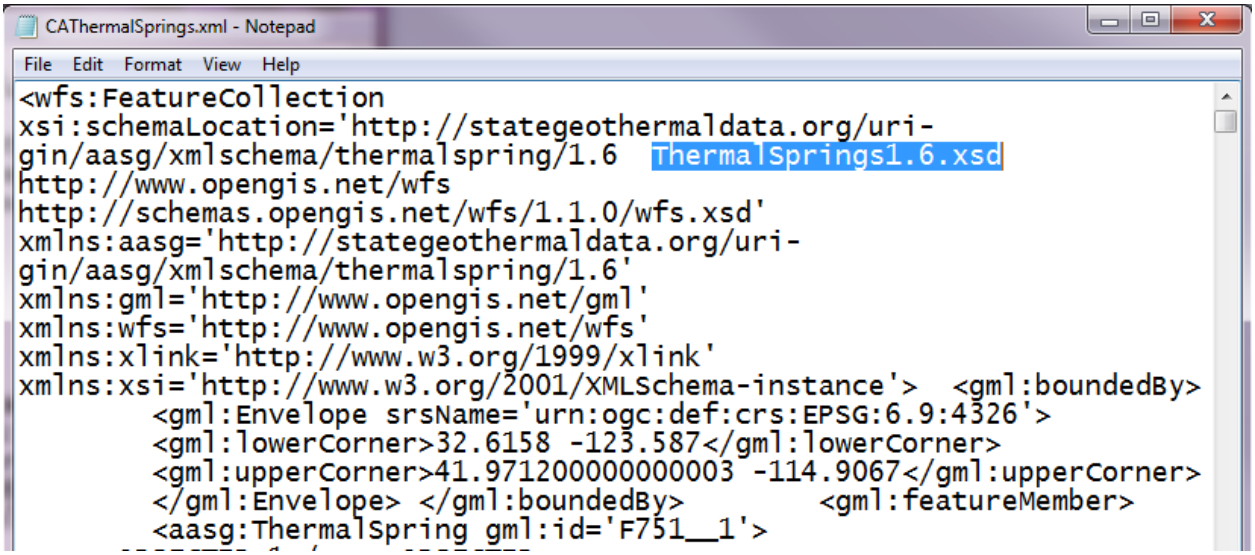

Figure 3

Section 2.5: XML Schema Validation with XML Explorer

Having downloaded your schema document and your **GetFeature document**, and having prepared your GetFeature document for validation, open the **GetFeature document** in XML Explorer.

1. Open XML Explorer
2. Click the **File** menu and click **Open**
3. Navigate to and open your **GetFeature document**
4. Click the **Errors** tab (Figure 4)
5. If the web service validates successfully against the designated schema, no errors should appear and your web service is valid

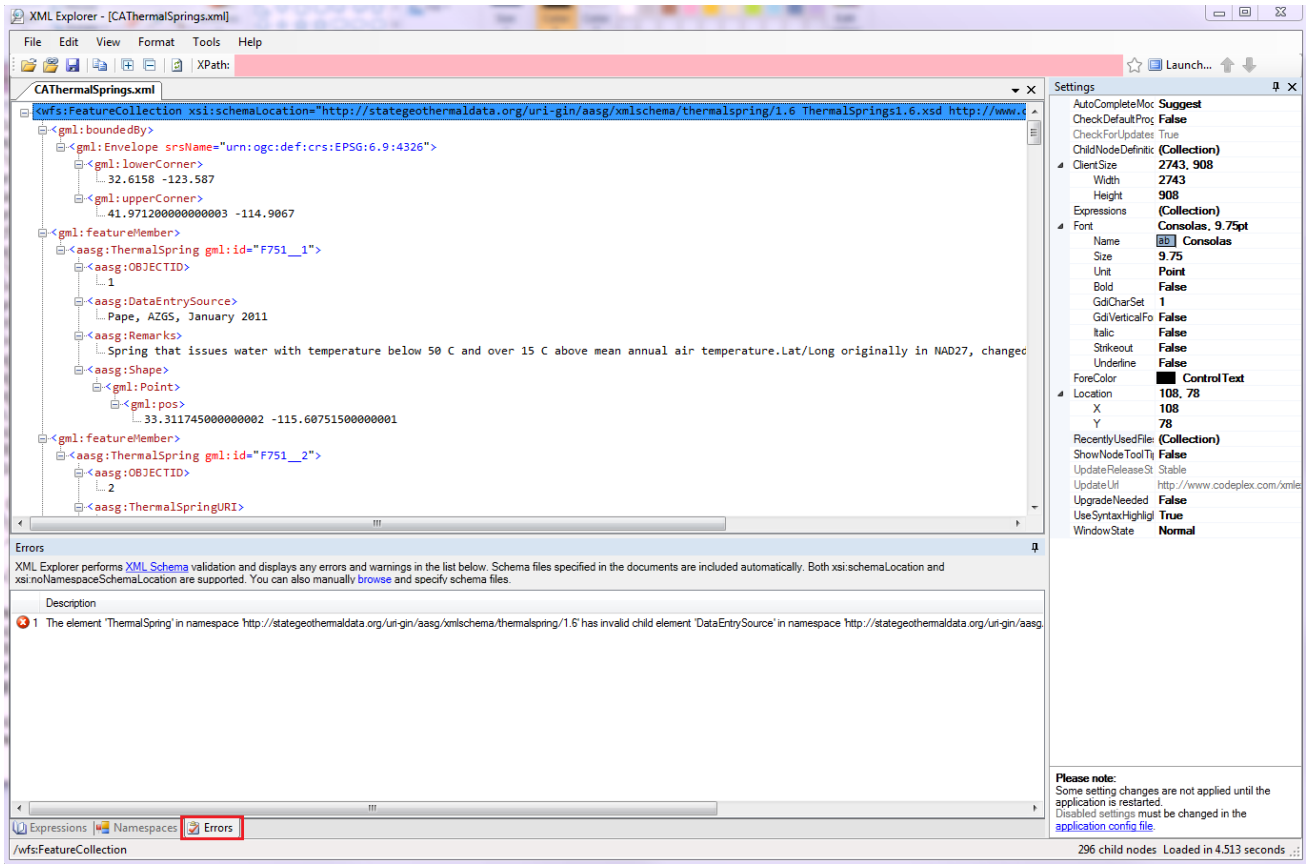

Figure 4

If errors appear, you will need to fix them within the web service itself. The process for doing so is outside the scope of this document.

Section 3.0: Troubleshooting

For questions on this process, the schemas, or other options for xml validation, please feel free to contact Celia Coleman at celia.coleman@azgs.az.gov.