Open Geospatial Consortium

Publication Date: 2014-07-16

Approval Date: 2014-06-14

Posted Date: 2014-05-16

Reference number of this document: OGC 14-048

Reference URL for this document: http://www.opengeospatial.net/doc/PER/testbed10/cci-hydro-model

Category: Public Engineering Report

Editor(s): Genong (Eugene) Yu, Liping Di

OGC® Testbed 10 Cross Community Interoperability (CCI) Hydro Model Interoperability Engineering Report

Warning

This document is not an OGC Standard. This document is an OGC Public Engineering Report created as a deliverable in an OGC Interoperability Initiative and is not an official position of the OGC membership. It is distributed for review and comment. It is subject to change without notice and may not be referred to as an OGC Standard. Further, any OGC Engineering Report should not be referenced as required or mandatory technology in procurements.

Document type:	OGC® Engineering Report
Document subtype:	NA
Document stage:	Approved for public release
Document language:	English

Abstract

This OGC® document gives guidelines for enabling interoperability among different hydro data models and services. The demonstration specifically gives out best practices for supporting interoperability among the National Hydrographic Network (NHN) of Canada, the National Hydrographic Dataset Plus (NHD+) of United States, and the OGC HY_Features model developed and proposed by the World Meteorological Organization (WMO). The discussed version of OGC HY_Features was adopted as the mediation bridge model to exchange information among heterogeneous hydrological models.

Keywords

ogcdoc, ogc documents, testbed10, cci, hydrology, modeling, hy_features, waterml, riverml, semantic mediation

Preface

This Engineering Report was created as a deliverable on Hydrological modeling for the OGC Testbed 10 (Testbed-10) initiative of the OGC Interoperability Program. Interoperability among heterogeneous hydro models is the focus of the effort. Best practices and recommendations are emerged for enabling interoperability among the hydrologic data models used by Hydrologic al Services of Canada and the United States using a common hydrologic feature model, namely HY_Features, developed in the joint WMO/OGC Hydrology Domain Working Group. The HY_Features model is currently an OGC Discussion Paper intended to be proposed to OGC as a specification for identifying and referencing hydrologic features. It was tested for the suitability to be a mediation model for interoperable discovery, access, and integration of heterogeneous hydrologic data and models.

License Agreement

Contents

Page

Figures

Page

Tables Page

OGC® Testbed 10 Cross Community Interoperability (CCI) Hydro Model Interoperability Engineering Report

1 Introduction

1.1 Scope

The scope of this Engineering Report is to capture advancements, best practices, standard adoption, specification change requests, and interoperability surrounding the Hydro Model Interoperability work. The work is a sub-task under the Cross-Community Interoperability (CCI) thread of the OGC Testbed 10 within the Interoperability Program. The effort focuses on demonstrating interoperability among hydrographic and hydrologic data sources. It bears the similar scope of CCI but a narrow subject on hydro models and services. The ultimate goal is to build on interoperability within hydro communities sharing geospatial data and advance semantic mediation approaches for data discovery, access, and use of heterogeneous hydro data models and heterogeneous hydro metadata models.

This OGC® document gives guidelines for enabling interoperability among different hydro data models and services. The demonstration specifically gives out best practices for supporting interoperability among the National Hydrographic Network (NHN) of Canada, the National Hydrographic Dataset Plus (NHD+) of United States, and the OGC HY_Features model developed and proposed by the World Meteorological Organization (WMO). The discussed version of OGC HY_Features was adopted as the mediation bridge model to exchange information among heterogeneous hydrological models.

This OGC® document is applicable to hydro model and service interoperability adopting emerging OGC specification – HY_Features.

1.2 Document contributor contact points

All questions regarding this document should be directed to the editor or the contributors:

Name	Organization
Genong (Eugene) Yu	George Mason University, gyu@gmu.edu
Liping Di	George Mason University, ldi@gmu.edu
Chenyu Hao	GIS.FCU, how@gis.tw
Lew Leinenweber	OGC, lleinenweber@opengeospatial.org
Irina Domblut	WMO-GRDC, Dornblut@bafg.de

Rob Atkinson	CSIRO, Rob.Atkinson@csiro.au

1.3 Future work

Hydrological data and model integration requires not only reconciliation of variations in geometric mapping but also semantic mapping of data structures and content values. Semantic mapping is a complex process, often performed in data preparation activities in different systems and processes in different ways. In the Testbed-10 a prototype mapping was designed and tested to link hydrological information models of different design but comparable semantics. From the experience of this test bed, the following further work is recommended.

- **Semantic support for mediation** – The prototype implementation was initially using xlink to achieve the mapping between different models. The major reasons are the existing knowledge and technology of the implementation organization and the easy use of xlink technology for the purpose of rapid development. The semantic mapping was tested using a prototype knowledgebase developed and implemented in OWL using Protégé editor. The knowledgebase was released for testing through a standard SPARQL engine. The demonstration has limited use of features and examples to mapping the hydrological features across models and therefore the use of semantic mapping is also limited on equivalences. In reality, semantic mapping does not stop on just equivalence of single properties, but may include more complex overlapping semantics of different groups of properties. The further requirements of semantic mapping in depth may call for a semantic mapping technology other than simple equivalence. Along the line of both geographic and semantic matching, the use of GeoSPARQL with the topological relationship defined in the HY_Features model may be further tested.

- **Standards based persistent hydrologic feature data access** – The demonstration in the Testbed-10 did not present the data in a standards based persistent data service, such as Web Feature Service (WFS). Non-standard access methods lead to a fragile and limited-scope solution, typically bound to a specific version of specific software. In order to serve hydrological features in a WFS, hydrology-specific application schemas are required. These should be grounded in the common concepts of the HY_Features model..

- **Semantic mapping framework conforms to ISO baselines** – The lack of a formal model of data products, or use of ad-hoc UML idioms, will lead to variability in implementing semantic mappings. This variability will in turn compromise the interoperability of the mapping specification, and of the data being described. This will perpetuate laborious, repetitive creation of ad-hoc semantic mappings on a case by case basis, typifying current low levels of data integration. This would also limit the interoperability and reusability of semantic mapping tools and frameworks. The purpose of frameworks is to support the semantic mapping process with either re-usable tools or pre-tested mappings of some components. It was recommended that such a mapping framework should conform to the ISO baselines, particularly the evolving ISO19150-2 standard[1], for encoding application schema in Resource Description Framework (RDF). This is important because it provides for a unique, resolvable identifier for each model element participating in semantic mappings. . It was also recommended that OGC take further steps to explore how to encapsulate the complexity of the problem by identification of common patterns, promulgating exemplar mappings supporting the development of appropriate tooling to make it a sustainable process. For example, the community would benefit from reusable or example semantic mappings based on the Observation and Measurements abstract model (ISO 19156) to typical implementation models such as NetCDF. The example use cases of this test bed are suggested as a test case for a future standardization activity for model mapping frameworks.

1.4 Forward

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. The Open Geospatial Consortium shall not be held responsible for identifying any or all such patent rights.

Recipients of this document are requested to submit, with their comments, notification of any relevant patent claims or other intellectual property rights of which they may be aware that might be infringed by any implementation of the standard set forth in this document, and to provide supporting documentation.

2 References

The following documents are referenced in this document. For dated references, subsequent amendments to, or revisions of, any of these publications do not apply. For undated references, the latest edition of the normative document referred to applies.

OGC 06-121r3, *OGC® Web Services Common Standard*

OGC 11-052r4, *OGC® GeoSPARQL - A Geographic Query Language for RDF Data*

OGC 11-039r3, *OGC® HY_Features: a Common Hydrological Feature Model Discussion Paper*

OGC 08-126, *The OGC® Abstract Specification Topic 5: Features*

OGC 99-110, *The OGC® Abstract Specification Topic 10: Feature Collections*

OGC 02-112, *The OGC® Abstract Specification Topic 12: The OGC Service Architecture*

OGC 99-114, *The OGC® Abstract Specification Topic 14: Semantics and Information Communities*

OGC 05-007r7, *OGC® Web Processing Service*

OGC 09-025r1, *OGC® Web Feature Service 2.0 Interface Standard*

ISO 19142:2010, *Geographic information -- Web Feature Service*

OGC 04-094, *OGC® Web Feature Service (WFS) Implementation Specification*

NOTE This OWS Common Standard contains a list of normative references that are also applicable to this Implementation Standard.

In addition to this document, this report includes several XML Schema Document files as specified in Annex A.

3 Terms and definitions

For the purposes of this report, the definitions specified in Clause 4 of the OWS Common Implementation Standard [OGC 06-121r3] and in OGC® Abstract Specification Topic 5: Features [OGC 08-126], Topic 10: Feature Collection [OGC 99-110], Topic 12: The OGC Service Architecture [OGC 02-112], and Topic 14: Semantics and Information Communities [OGC 99-114] shall apply. In addition, the following terms and definitions apply.

3.1 HY_Features

HY_Features is a hydrologic feature model developed in the joint WMO/OGC Hydrology Domain Working Group [OGC 11-039r3].

3.2 National Hydrographic Network

Its acronym is NHN. It is a hydro model developed and used in Canada.

3.3 National Hydrographic Dataset Plus

Its acronym is NHD+. It is a hydro model developed and used in United States.

NOTE NHD+ and NHD are used interchangeable in the context of this report.

4 Conventions

4.1 Abbreviated terms

CCI Cross-Community Interoperability

NHD+	National Hydrographic Dataset Plus
NHN	National Hydrographic Network
WFS	Web Feature Service
WPS	Web Processing Service

4.2 UML notation

Most diagrams that appear in this standard are presented using the Unified Modeling Language (UML) static structure diagram, as described in Subclause 5.2 of [OGC 06-121r3].

4.3 Used parts of other documents

This document uses significant parts of document [OGC 06-121r3]. To reduce the need to refer to that document, this document copies some of those parts with small modifications. To indicate those parts to readers of this document, the largely copied parts are shown with a light grey background (15%).

4.4 Data dictionary tables

The UML model data dictionary is specified herein in a series of tables. The contents of the columns in these tables are described in Table 1.

Table 1 — Contents of data dictionary tables

Column title	Column contents
Names (left column)	Two names for each included parameter or association (or data structure). The first name is the UML model attribute or association role name. The second name uses the XML encoding capitalization specified in Subclause 11.6.2 of [OGC 06-121r3]. The name capitalization rules used are specified in Subclause 11.6.2 of [OGC 06-121r3]. Some names in the tables may appear to contain spaces, but no names contain spaces.
Definition (second column)	Specifies the definition of this parameter (omitting un-necessary words such as “a”, “the”, and “is”). If the parameter value is the identifier of something, not a description or definition, the definition of this parameter should read something like “Identifier of TBD”.
Data type and value (third column) or Data type (if are no second items are included in rows of table)	Normally contains two items: The mandatory first item is often the data type used for this parameter, using data types appropriate in a UML model, in which this parameter is a named attribute of a UML class. Alternately, the first item can identify the data structure (or class) referenced by this association, and references a separate table used to specify the contents of that class (or data structure). The optional second item in the third column of each table should indicate the source of values for this parameter, the alternative values, or other value information, unless the values are quite clear from other listed information.
Multiplicity and use	Normally contains two items:

Column title	Column contents
(right or fourth column) or Multiplicity (if are no second items are included in rows of table)	The mandatory first item specifies the multiplicity and optionality of this parameter in this data structure, either “One (mandatory)”, “One or more (mandatory)”, “Zero or one (optional)”, or “Zero or more (optional)”. The second item in the right column of each table should specify how any multiplicity other than “One (mandatory)” shall be used. If that parameter is optional, under what condition(s) shall that parameter be included or not included? If that parameter can be repeated, for what is that parameter repeated?

When the data type used for this parameter, in the third column of such a table, is an enumeration or code list, all the values specified shall be listed, together with the meaning of each value. When this information is extensive, these values and meanings should be specified in a separate table that is referenced in the third column of this table row.

The data type of many parameters, in the third table column, is specified as “Character String type, not empty”. In the XML Schema Documents specified herein, these parameters are encoded with the xsd:string type, which does NOT require that these strings not be empty.

The contents of these data dictionary tables are normative, including any table footnotes.

4.5 Use of term NHD+ and NHD

For the convenience of diagram and description, NHD is used interchangeable with NHD+ to mean the hydro model currently used in USA. In text, NHD+ is used to represent the current hydro model used in US. In diagrams and some text, NHD is used to indicate the same model but for simplicity without using special “+” sign.

5 Hydro modeling overview

The specified Hydro Model Interoperability task addresses the interoperability among heterogeneous hydrographic and hydrologic data sources through adopting the HY_Features common hydrologic feature model. It is part of the Cross-Community Interoperability (CCI) Thread. The Testbed-10 seeks to specifically answer the following questions:

(1) How does the HY_Features model enable access to hydro data from multiple sources?

(2) How does the HY_Features model help the mediation between different hydro conceptual feature models?

(3) Would a HY_Features profile be required for each data source?

(4) What are the best practices to build a hydrographic and hydrologic data source mediation system based on the HY_Features model to support interoperability across jurisdictions, between various data models, and data structures?

In answering these questions, the following overall architecture has emerged during the Testbed-10 as shown in Figure 1.

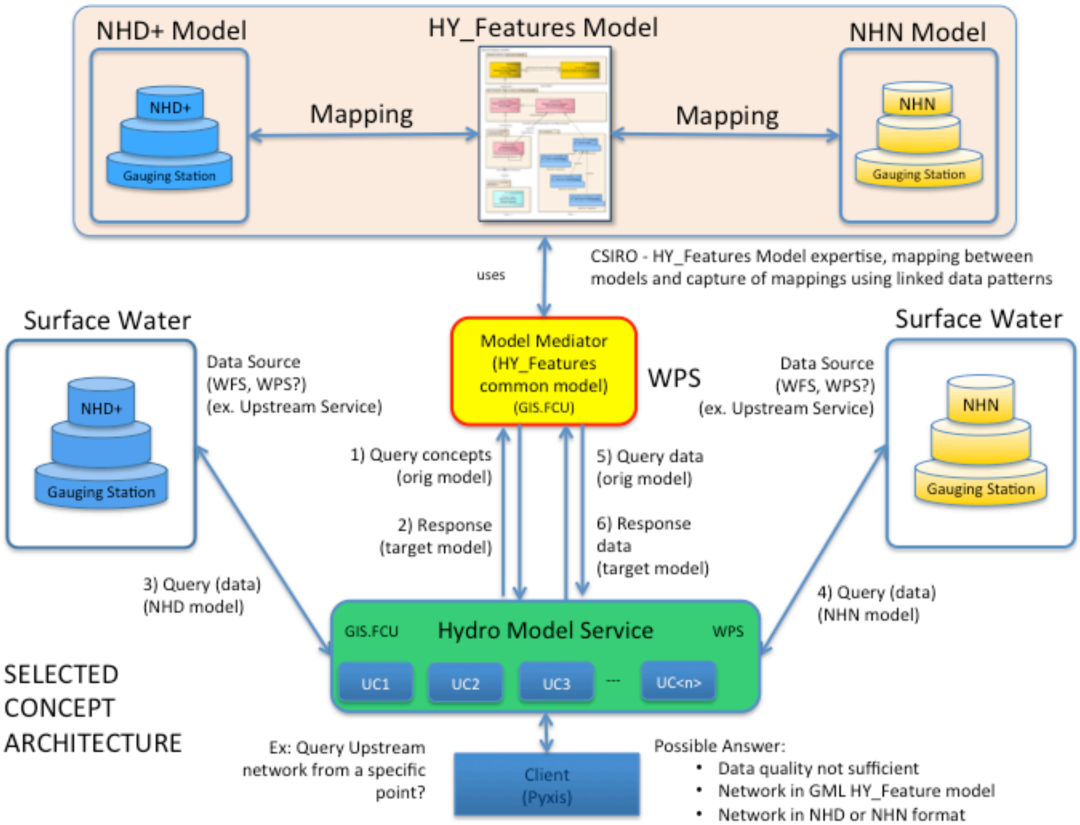

Figure 1 — Overall Architecture

6 Hydro model interoperability

6.1 Data and model resources

6.1.1 Data access services

6.1.1.1 Spatial feature data

Two data models are under investigation for semantically data integration by using a common hydrologic feature model – HY_Features. The two data models are NHD/NHD+ and NHN.

NHD+, or NHDPlus, is an integrated set of application-ready geospatial data sets that incorporates features from the National Hydrography Dataset (NHD), the National Elevation Dataset (NED), and the Watershed Boundary Dataset (WBD)[7]. The NHD+ Flow Lines, catchments, and watershed boundaries are available packaged with numerous other data products from: ftp://ftp.horizon-systems.com/NHDPlus/NHDPlusV21/Data/ .

Metadata for these are available from:

ftp://ftp.horizon-systems.com/NHDPlus/NHDPlusV21/Documentation/NHDPlusV2_Metadata.7z

NHD data access can be accomplished by five approaches online:

(1) Access data through an online order system – http://nhd.usgs.gov/data.html. The order would be usually gone through the packaging process and delivery in 1 or 2 hours.

(2) Direct access to sub-regions identified by 4-digit identification number. This can be accomplished by using NHD Viewer (http://viewer.nationalmap.gov/viewer/).

(3) Download the data by state from the ftp site - ftp://nhdftp.usgs.gov/DataSets/Staged/States/FileGDB

(4) Extract subset data of only Streamgages and Dams obtainable from ftp://nhdftp.usgs.gov/DataSets/National

(5) Watershed boundaries can be download from NRCS Geospatial Data Gateway (http://datagateway.nrcs.usda.gov/)

NHD+ data can be downloaded from http://www.horizon-systems.com/NHDPlus/V2NationalData.php by either http download or ftp download. The test dataset for the Red River can be downloaded at: ftp://ftp.horizon-systems.com/NHDPlus/NHDPlusV21/Data/NHDPlusSR/, where the NHDSnapshot (i.e. NHDSnapshot 04.7z) contains flowlines. The test dataset for the Great Lakes can be downloaded at: ftp://ftp.horizon-systems.com/NHDPlus/NHDPlusV21/Data/NHDPlusGL/.

The USGS National Water Census will be publishing WFS services for the NHD+ verison 2 Flowlines, catchments (One per flowline), and corresponding HUC12 watersheds (common aggregations of catchments). These services are not available at the time of this demonstration work. Therefore, the dataset are directly retrieved from the website through ftp or http download.

NHN data are also managed in a geodatabase. They can be accessed from the following website:

http://www.geobase.ca/geobase/en/find.do?produit=nhn

They are served online via three options to be accessed:

(1) By querying for a specific NHN Work Unit, for example, 05OB000, interactively through http://www.geobase.ca/geobase/en/browse.do?produit=nhn&decoupage=units&map=canada . They are available in ESRI shapefile, KML, GML, or ESRI File GeoDatabase(FGDB).

(2) By searching NHN Work Units at http://www.geobase.ca/geobase/en/find.do?produit=nhn

(3) By browsing the FTP directory at http://www.geobase.ca/geobase/en/download.do?produit=nhn&items=official/nhn_rhn/&protocol=ftp .

In the Testbed-10 demonstration project, the NHN data in ESRI shapefile format for the Red river was made available in a special pre-subset package. The data had been subset at the border between Canada and the United States, for the purpose of demonstrating cross-border data integration. The data is available through FTP at ftp://ftp.ctis.nrcan.gc.ca/pub/hydro/Denis/OWS10/RedRiver.zip.

6.1.1.2 Observation data

Gauge measurements or stream measurements can be retrieved from related website. For example Streamgages and Dams data can be extracted in NHD from ftp://nhdftp.usgs.gov/DataSets/National.

6.1.2 Information models

6.1.2.1 Conformance to ISO 19103 and ISO 19109

For information models to be interoperable a consistent UML style and data types are necessary. The ISO 19103 Conceptual Schema Language provides this, and the 19109 General Feature Model defines relevant Meta-classes. HY_Features, WaterML 2, and several other standards use this meta-standard. Conformance to this standard, and use of the ISO model library is necessary to generate GML and OWL using the automation tools available. Further information about the conformance to ISO standards in previous research can be found at https://www.seegrid.csiro.au/wiki/AppSchemas/WebHome

6.1.2.2 Semantic feature type catalog

The Spatial Identifier Reference Framework (SIRF) project, a project of the Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia, has developed a 'Semantic Feature Type Catalog' based on the OWL representations of information models. This is being prototyped and tested to support Testbed-10. Details can be found at https://wiki.csiro.au/display/SIRF/FeatureTypeCatalog.

The SPARQL endpoint for this is provided at: http://unsdi-dev.csiro.au/openrdf-sesame/repositories/ftc. Other API endpoints can be accessed at http://unsdi-dev.csiro.au/sissvoc/meta/.

To extend the functionality of the API, an application project should ideally provide a SPARQL CONSTRUCT query to generate a desired response, otherwise a description of the required function and its purpose. For easy visualization and display, a HTML interface to the API is required. The API could be extended to meet the specific requirements when the tools become completely available to developers through open source access in near future. The current HTML interface is enabled through a slightly hacked version of a very complex general purpose style sheet. Alternatives using Velocity are now available. Further details can be found at https://code.google.com/p/elda/. In the Testbed-10 demonstration, these APIs were not adopted and extended, considering their developing stage. However, the concepts and the framework did help in implementing the semantic mapping with available tools and resources in the Testbed-10 project. As their APIs evolves, it is perceived that they could be directly adapted and re-used for the future implementation of semantic mapping with the common HY_Features mediation model.

6.1.2.3 HY_Features common hydrological feature model

The HY_Features model[2] is available via a Feature Type Catalog Linked Data API using a temporary domain name provided within the SIRF project in the framework of Australian Hydrological Geospatial Fabric (AU-Geofabric). The URIs for concepts defined in the OWL model are used as the basis for a URI, which may be resolved to alternative representations. The SIRF project is ongoing. Efforts are currently working to hook up all the views to all the resources.

Concepts in the HY_Features model may be accessed using the URI pattern:

http://demo.sirf.net/def/schema/hydrology/{schemafilename}/{concept }

where {schemafilename} is the value of the tagged value "xsdDocument" (minus the .xsd extension) and {concept} is the name of the class or property, converted to standard conventions (UpperCamelCase? for classes, lowerCamelCase for properties).

For example, HY_Basin and outflowNode are given respectively at
http://demo.sirf.net/def/schema/hy_features/abstracthydrofeature/HY_Basin and
http://demo.sirf.net/def/schema/hy_features/abstracthydrofeature/outflowNode.

The current implementation allows requests to query each object. By default, the request for HTML for one of these objects will reach to a landing page listing all the available views and formats available. The OWL file can be accessed through API for application/rdf+xml. The following is an example request that will return the OWL file containing the relevant definitions, using these same URI patterns.

https://svnserv.csiro.au/svn/SIRF/public/def/schema/hy_features/abstracthydrofeature.rdf#HY_Basin

6.1.2.4 WaterML 2 model

The WaterML-2.0 model[3], [4] could be a potential bridging model for integration of heterogeneous hydrologic data. This work is in progress. In this demonstration, the WaterML-2.0 model was not adopted and examined, considering the time constraints on implementation and existing technology of the testing implementation organizations.

6.1.2.5 RiverML model

RiverML[5], [6] is an intended standard format for river geometry and flow data by extending WaterML 2.0[3], [4] with specialization of river geometry and hydraulics. This model is still under development. It was not be used in the OWS 10 test bed.

6.1.2.6 NHD+ model

The NHD+ model is provided by the USGS model definitions .xmi, .ead, and gml schemas that are available at this link: http://www.fgdc.gov/standards/projects/incits-l1-standards-projects/framework/models/.

6.1.2.7 NHN model

Each NHN data from the ftp site (ftp://ftp.ctis.nrcan.gc.ca/pub/hydro/Denis/OWS10/) contains a corresponding descriptions file in both English and French. The description files provide metadata information about the data, including identification, data quality, spatial data organization, entity and attribute, distribution, and metadata reference.

The NHN model is divided in 4 main packages[8]–[10]:

• Hydro network: The hydro network is made up of network linear element and network point element[8]. The network linear element includes network linear flow and water boundary entity. The network element represents hydro junctions.

• Hydrographic: The hydrographic features include man made hydrogrphic entity, hydrographic obstacle entity, island, water body, simple line watercourse, and coastline[8].

• Toponymy: Toponymy is an external package that defines hydro traversal, hydrographic entity, and named point, line, or polygon features[8].

• Hydro event: A hydro event can be a linear event or a point event[8].

For detailed catalogue distribution profile, further descriptions can be found on the following link[11], [12].

http://www.geobase.ca/doc/catalogue/GeoBase_NHN_Catalogue_1.0.1_EN.html

The descriptions cover feature classes, including bank, delimiter, external geometry event, external line event, eternal point event, flow property event, hydrographic obstacle entity, hydro junction, island, littoral, manmande hydrographic entity, manmade line event, manmande point event, named feature, network linear flow, obstracle line event, obstacle point event, single line watercourse, and waterbody[12].

6.2 Semantic mapping of hydrological models

The mapping between different hydrological models can be formalized through mediation by means of shared semantics. In the Testbed-10 a two-step mapping approach is applied using the HY_Features model in sense of a mediation service. At first, the hydrological models are mapped to common feature concepts of HY_Features. Secondly, the HY_Features model is “re-mapped” to the target models. Figure 2 shows the mapping approach intended to be used in this Testbed-10.

To achieve semantic interoperability among the NHD+ - and NHN data models, two separate mappings are required: (1) mapping of NHD+ features to the equivalent concepts of HY_Features, and (2) mapping NHN features to HY_Features equivalents.

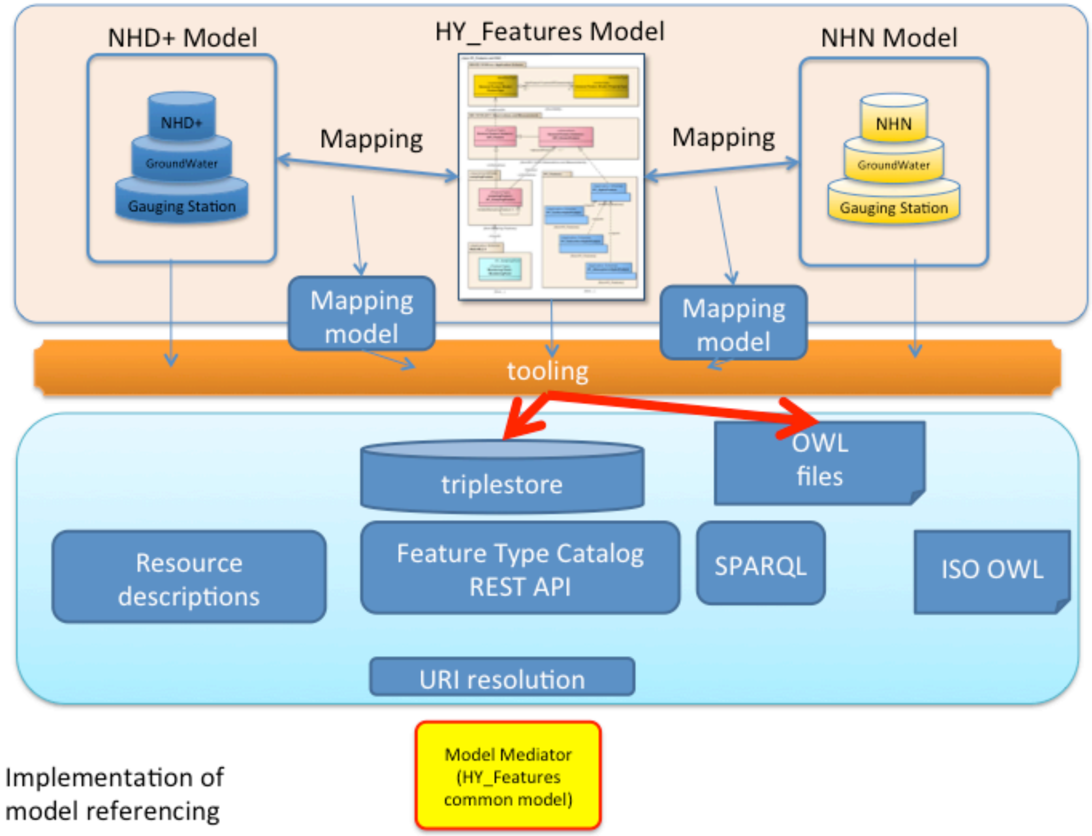

Figure 2 — Mapping framework

The semantic mapping is established through a mapping model which will allow a standard RDF/OWL encoding of data models conforming to ISO standards, particularly the emerging ISO19150-2 standard. Since each module of the HY_Features model is an ISO19109 application schema, which could be converted into RDF/OWL, the mapping to HY_Features common concepts will make NHD+ or NHN elements addressable and discoverable for automated web services following the Linked Data approach.

Figure 3 and Figure 4 show exemplarily how linear features of NHD+ or NHN source models to their HY_Features equivalents at the field/property level. For each source model, a similar mapping class is defined whose properties link the source and the target model. In the UML diagrams, entities shaded in “yellow” are mapping class and properties. Those in green are source model in source models – either in NHD+ or in NHN in the Testbed-10 project. Those in blue are target HY_Features model.

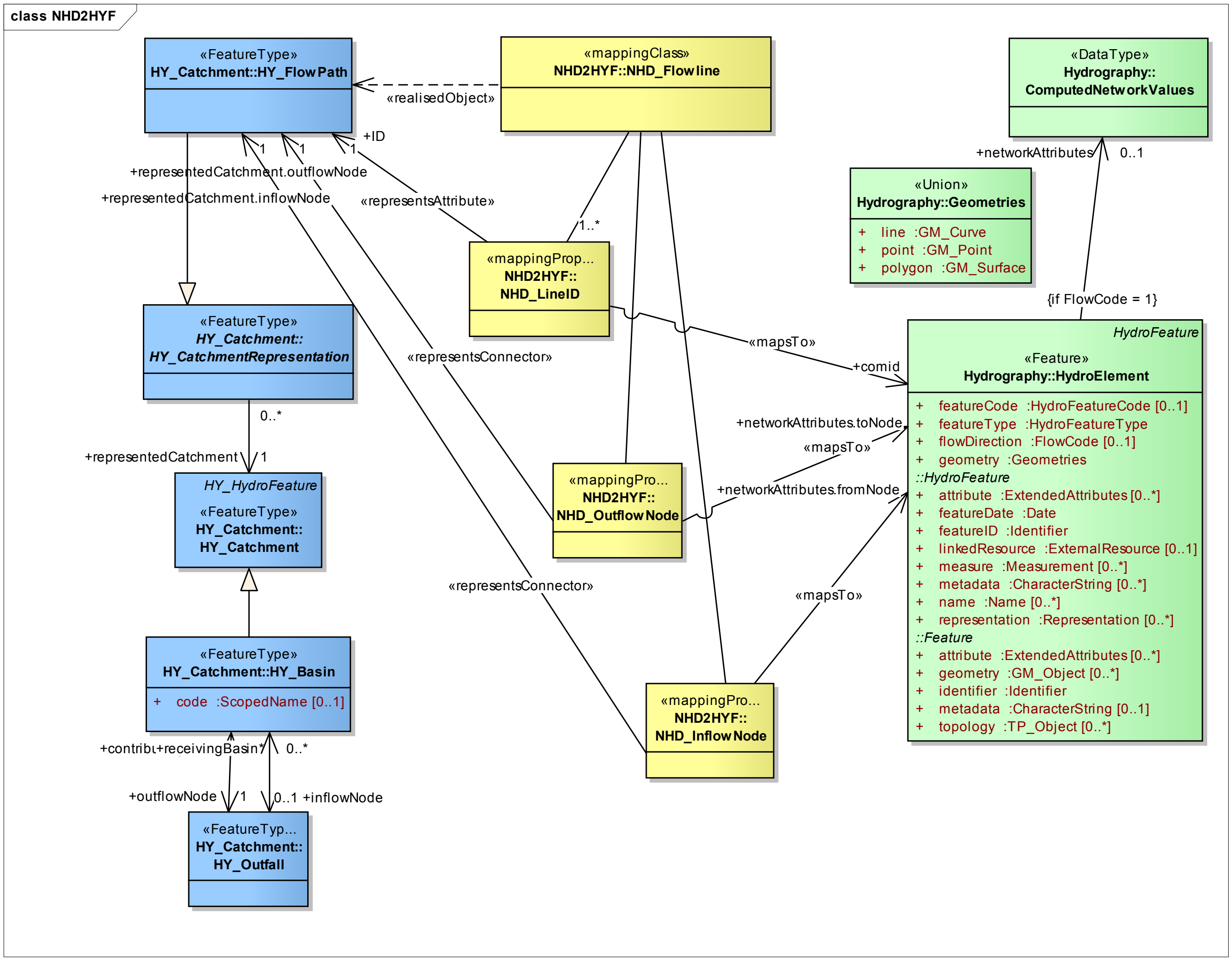

Figure 3 — NHD_Flowline - Mapping NHD_HydroElement to HY_FlowPath

Figure 3 shows the UML diagram expressing the mappings of a linear HydroElement in the NHD source model to HY_FlowPath in the HY_Features model. The following describe the three major property mappings of the mapping class:

1. NHD_InflowNode: This is the mappingProperty that maps the fromNode (relation) in NHD+ to the representedCatchment.inflowNode in HY_Features.

2. NHD_OutflowNode: This is the mappingProperty that maps the toNode (relation) in NHD+ to the representedCatchment.outflowNode in HY_Features.

3. NHD_LineID: This is the mappingProperty that maps the attribute comid in NHD+ to Flowpath.ID in HY_Features.

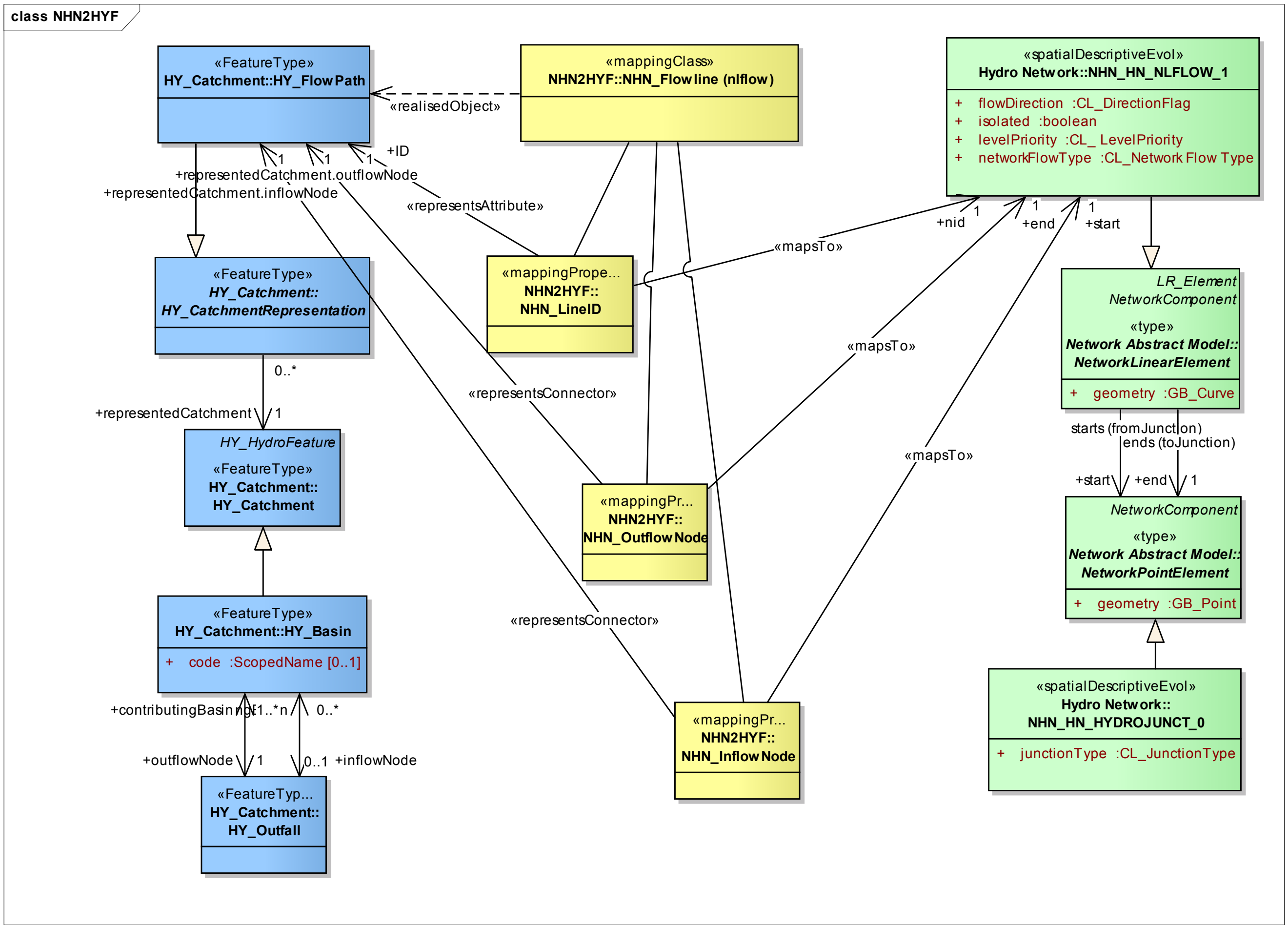

Figure 4 — NHN_Flowline - Mapping NHN_nlflow to HY_FlowPath

Figure 4 shows the UML diagram expressing the mappings of an nlflow feature in the NHN source models to HY_FlowPath in the HY_Features model. The following describe the three major property mappings of the mapping class:

1. NHN_InflowNode: This is the mappingProperty that maps the start node (fromJunction) relation in NHN to representedCatchment.inflowNode in HY_Features.

2. NHN_OutflowNode: This is the mappingProperty that maps the end node (toJunction) relation in NHN to representedCatchment.outflowNode in HY_Features.

3. NHN_LineID: This is the mappingProperty that maps the attribute nid in NHN to Flowpath.ID in HY_Features.

6.2.1 Semantic mapping implementation in the Testbed-10

In the hydro demonstration thread of the Testbed-10, three types of semantic constructs have been implemented to support the semantic mapping among NHD+, NHN, and HY_Features, as shown in Figure 5. They are data model structure (blue shaded section in Figure 5), conceptual model structure (red shaded section in Figure 5), and mapping model structure.

1. Data model structures: The data model structure includes elements for model, entity, and field. Here, model is used to define data model, like NHN and NHD+. Entity is used to define the entity or table belong to a model, like nlflow in NHN. The relation "belongModel" can be used to connect Model and Entity. Field is used to define the field or attribute of an entity, like to_junct in nlflow. The relation "belongEntity" can be used to connect Entity and Field. IDField is a Field, which is used to uniquely represent an identifier of an Entity.

2. Conceptual model structures: The conceptual model structure has ConceptualModel, ConceptualEntity, and ConceptualField. They are similar in definition as those for the data model structure, but they are used as an intermediate data structures to bridge the source model structure and the target model structure. In this project, the HY_Features model is used as the basis for constructing these conceptual model structures.

3. Mapping model structures: The mapping model structure has three classes, FieldMapping, MappingModel, and CustomizaField. The basic mapping approach is Field to Field mapping under this framework. Mapping relationships are described by properties, i.e. mappingField, targetField, sourceModel, and targetModel, to represent source field, target field, source model, and target model respectively. Under this framework, the class FieldMapping can use the property mappingField to connect with the Field, which is a class of the data model. The class FieldMapping can also use property targetField to connect with the ConceptualField, which is a class of the conceptual model structure.

CustomizedField is a class that belongs to ConceptualField and used to define mutual fields between data and conceptual model. For instance, if there is a field [len] representing length in NHN, and [length] in NHD+, then, a CustomizedField [Length] can be defined for mapping these fields in both data models. The MappingModel is a class used to indicate the mapping target model. The property sourceModel is used to connect MappingModel and the data model. The property targetModel is used to connect MappingModel and the conceptual model.

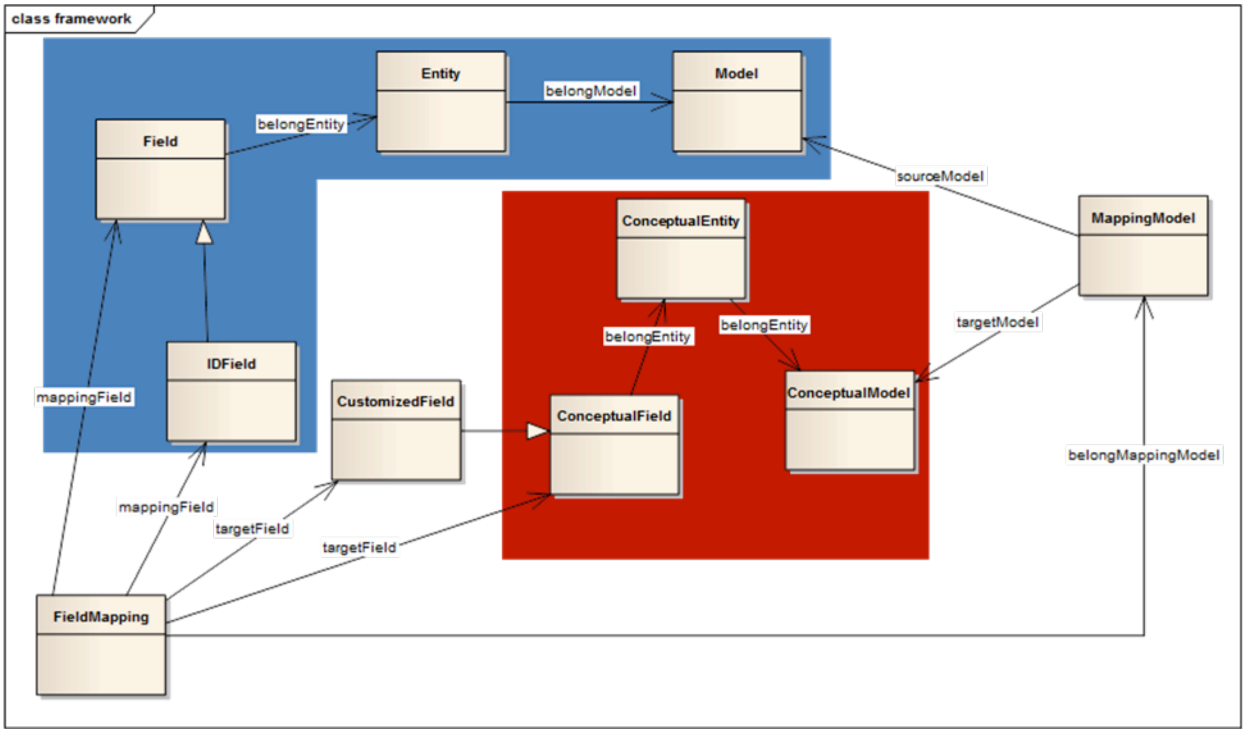

Figure 5 — Semantic Mapping Framework

6.2.2 NHD+ / NHN source models and HY_Features

The HY_Features model is used as the conceptual model structure under the mapping framework. Figure 6 shows the conceptual model relationship. The model has an entity HY_FlowPath and a field HY_FlowPath_ID.

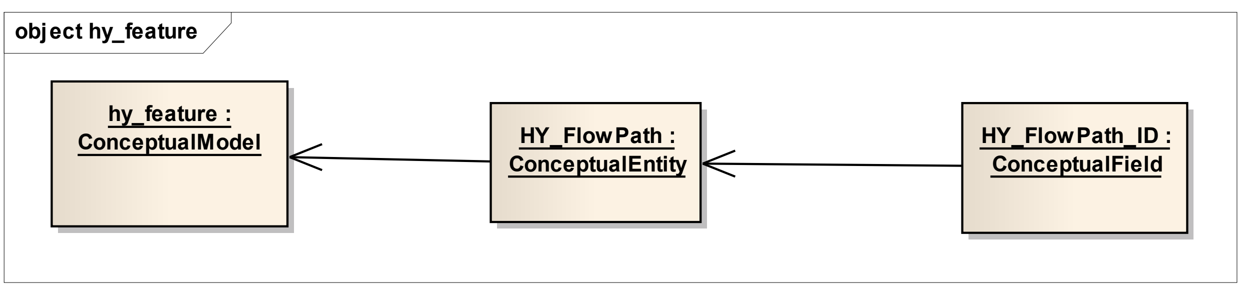

Figure 6 — Conceptual Model Structure: HY_Features

The NHD+ data model structure is shown in Figure 7. It has an entity flowline, and three fields comid, from_node, and to_node.

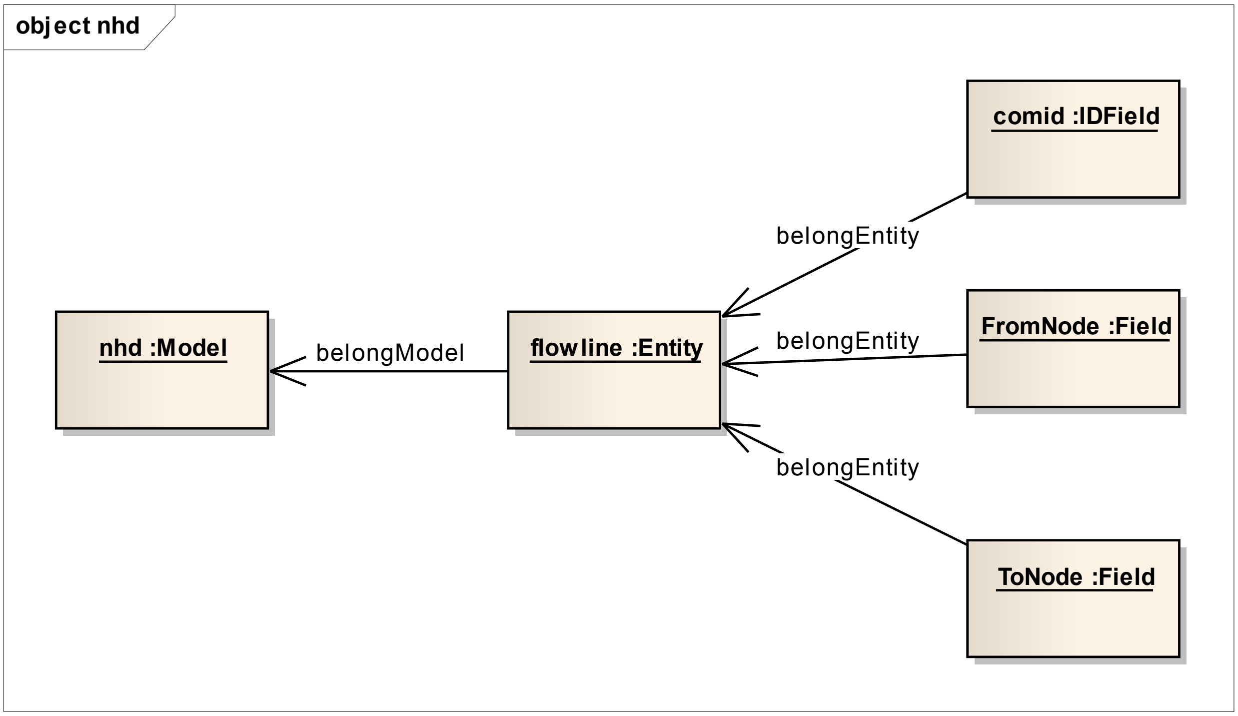

Figure 7 —Data Model Structure: NHD+

The NHN data model structure is shown in Figure 8. It has an entity nlflow, and three fields nid, from_junct, and to_junct.

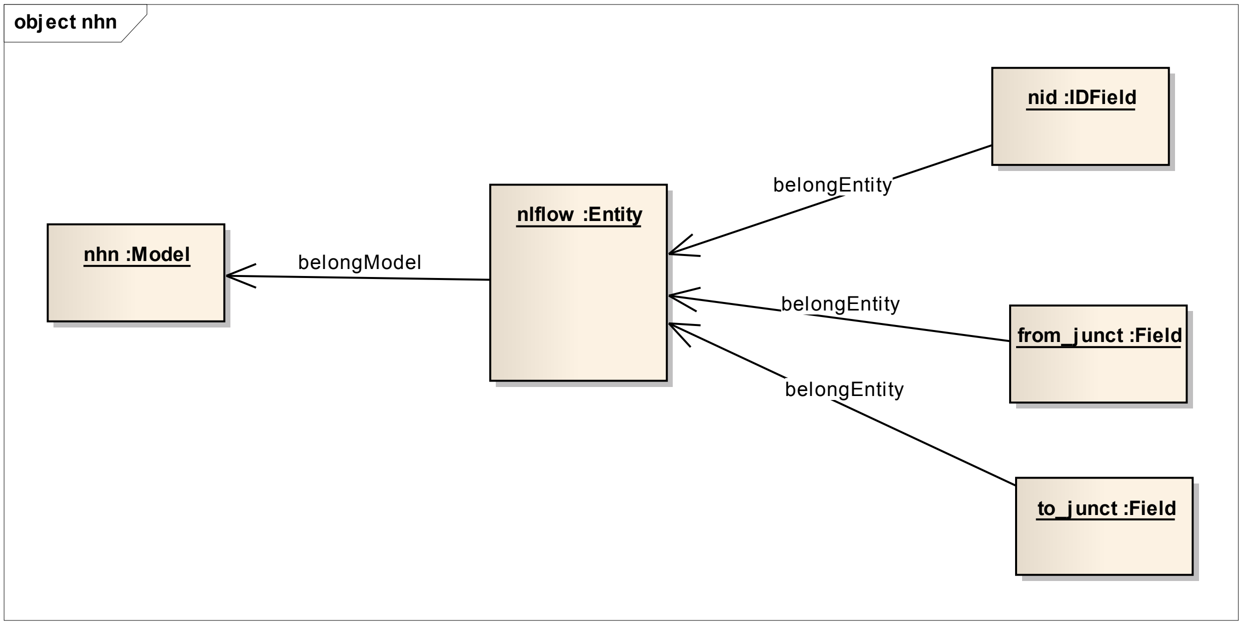

Figure 8 — Data Model Structure: NHN

6.3 Mediation system implementation

6.3.1 Overall system architecutre

To enable the realization of mapping between different hydro models through semantic mediation and demonstrate the capability of open standard and specifications in facilitating open access and interoperation, the Testbed-10 Hydro team implemented and deployed a mediation system in the Web environment. Figure 9 shows the overall architecture. The system consists of several Web services that were implemented following open standards and specifications in wrapping up existing legacy services and programs or internal Web services. The Hydro Model Services were published as standard OGC WPS processes, implemented following OGC Web Processing Service [OGC 05-007r7]. Several data services were reused either through database internet API or OGC data services – such as OGC Web Feature Service (WFS) [OGC 09-025r1, ISO 19142:2010, and OGC 04-094]. Figure 10 shows one typical interaction scenario between services of the mediation system.

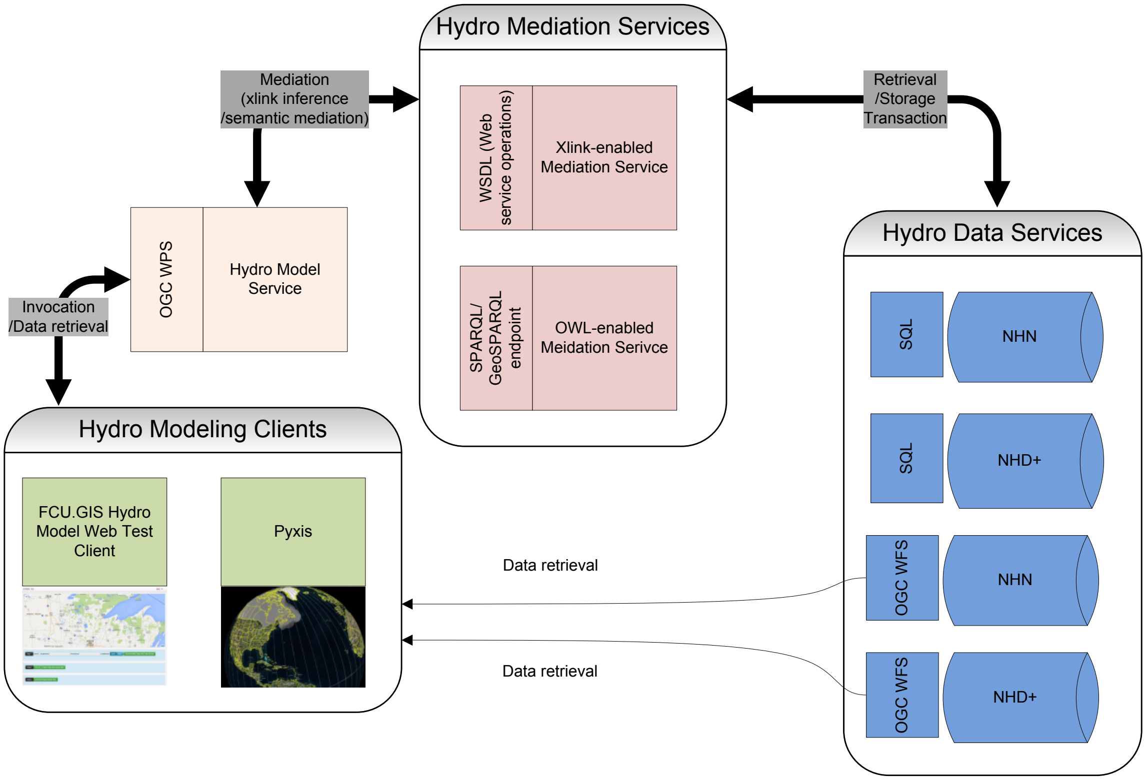

Figure 9 — Overall Architecture

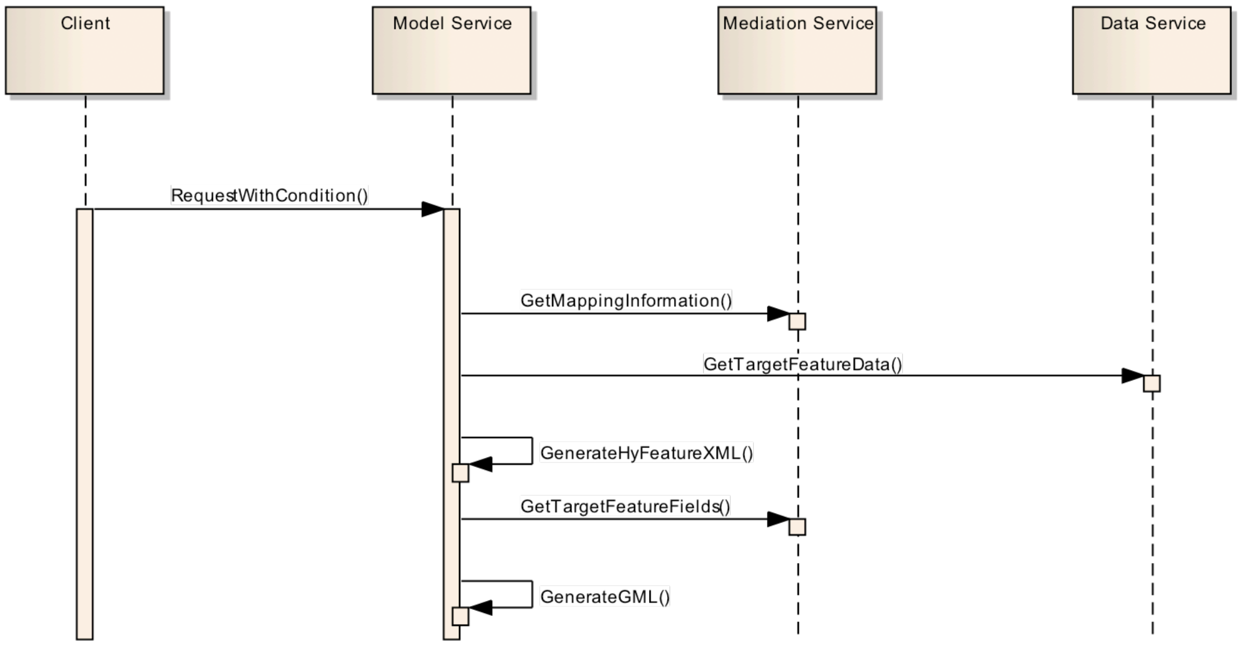

Figure 10 — Interactions between component services

6.3.2 Service components

6.3.2.1 Hydro model mediator

The hydro model mediator consists of Web services and components that interact with different hydro data models and response to requests from clients through the hydro model service. The current implementation leverages the existing Web services and databases to interact with back end data services. The backend data can exist in database, data store, and semantic data store (e.g. RDF data store, OWL data store). Their access may be enabled through traditional SQL database query, SPAQRL query, GeoSPARQL query, or WFS requests. As shown in Figure 9, the results could be stored persistently through transaction capability of OGC WFS and made available to clients if persistent storage is desired. Alternatively, the result can be directly sent back to client, especially when the process can be done synchronously within a short period of processing time.

6.3.2.2 Hydro model service

The Hydro Model Service was implemented as a Web Processing Service (WPS) based on 52North codebase.

6.3.3 Backend communication between service components

Figure 11 illustrates the communications between Model Mediator and Hydro Model Service at server side.

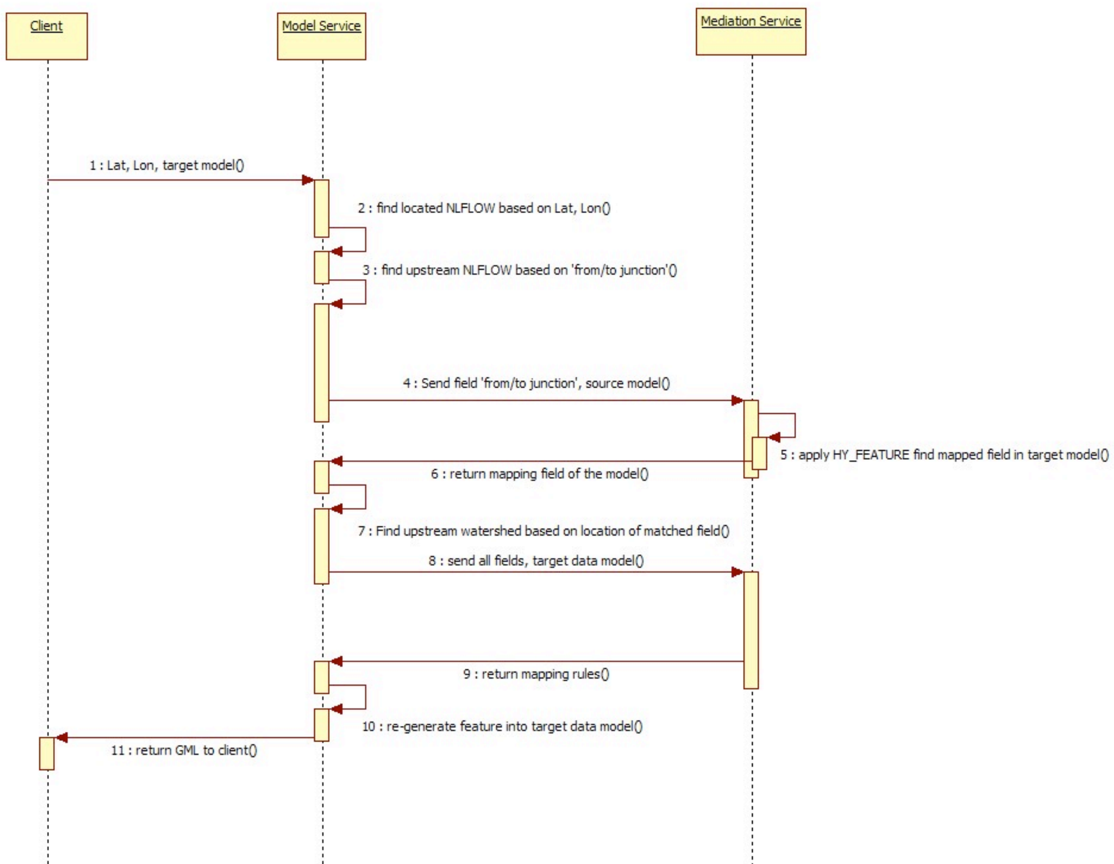

Figure 11 — Service Communication

This instance was assumed that initial POI located at *RedRiver*? of Canadian side.

1. Client sends a request with two parameters to model service, includes the POI(lat, lon) and data model that the user expects to get eventually.

2. After Model service has been triggered, the location of POI will be used to determine the located NLFLOW(also BANK and WATERBODY) feature.

3. Model service applies JUNCTION relationship(fields of NLFLOW attribute, matched juntion ID in both NLFLOW, isolated = 0) to find upstream NLFLOW features.

4. After all upstream NLFLOW has been processed out, Model Service will send junction fields to Mediation Service to get mapping field, class, feature in NHD.

5. Mediation Service uses HY_FEATURE to find the mapped fields of NHD

6. Mediation Service returns the NHD fields to model service.

7. Model service uses the location of upstream junction that nearest border as initial POI, and search all upstream watershed of the US side.

8. After Model service get entire upstream watershed from both sides, it sends a request to Mediation service with all fields and the target model.

9. Mediation service returns the mapping rules(not sure the implementation approach yet)

10. Model service re-generates the feature based on the mapping rules, and save as GML format with target data model.

6.3.4 Mapping implementation using xlink

Initially, xlink was used to quickly implement the relationships between classes of the HY_Features model. The reasons for adopting xlink are its ease of use, straightforwardness, and previous experiences of implementation team. A simple example is given below to illustrate the idea of using xlink in enabling the classes of the HY_Features model and their mapping to/from other models.

Using xlink to represent the relation between classess of the HY_Features model: The association [representedCatchment] of HY_FlowPath links to a HY_Basin, here, basin2. The association can be expressed as follows.

```
<HY_FlowPath>
  <representedCatchment xlink:href="#basin2"/>
</HY_FlowPath>
<HY_Basin id="basin2" >
  <inflowNode xlink:href="#outflow1"/>
  <outflowNode xlink:href="#outflow2"/>
</HY_Basin>
```

Using xlink to link objects among HY_Features, NHN, and NHD+: The following example illustrate how hydro feature “basin1” links between NHN and HY_Features through xlink. In this example, HY_Features “basin1” is linked to NHN object “nhn_hn_nlflow” uniquely identified by ID "d30e64b950f04319b".


```
<HY_Basin id="basin1"
xlink:href="http://geobase.ca/geobase/nhn/nhn_hn_nlflow#d30e64b950f04319b">
  <outflowNode xlink:href="#outflow1"/>
</HY_Basin>
```

Therefore, xlink can be used to establish a XML that links the real data model and the HY_Feature reference model. The Hydro Model Service can get the data through resolving this XML. The following is a complete HY_Features relationship in XML.

```
<HY_Features xmlns:xlink="http://www.w3.org/1999/xlink">
  <HY_FlowPath xlink:href="http://nhn/nlflow#2b1b3456762d">
    <representedCatchment xlink:href="#basin1"/>
   <id xlink:href="http://nhn/nlflow/nid"/>
    <fromnode xlink:href="http://nhn/nlflow/from_junction"/>
    <tonode xlink:href="http://nhn/nlflow/to_junction"/>
  </HY_FlowPath>
  <HY_FlowPath xlink:href="http://nhd/FlowlineVAA#2b1b9b62d">
    <representedCatchment xlink:href="#basin2"/>
    <id xlink:href="http://nhd/FlowlineVAA/ComId"/>
    <fromnode xlink:href="http://nhd/FlowlineVAA/FromNode"/>
    <tonode xlink:href="http://nhd/FlowlineVAA/ToNode"/>
  </HY_FlowPath>
  <HY_Basin id="basin1" xlink:href="http://nhn/waterbody#2b1b3456762d">
    <outflowNode xlink:href="#outflow1"/>
  </HY_Basin>
  <HY_Basin id="basin2" xlink:href=" http://nhd/waterbody#2b1b9b62d">
    <inflowNode xlink:href="#outflow2"/>
    <outflowNode xlink:href="#outflow3"/>
  </HY_Basin>
  <HY_Outfall id="outflow1" xlink:href="http://nhn/#aa9729ee053">
    <contributingBasin xlink="#basin1"/>
    <receivingBasin xlink="#basin2"/>
  </HY_Outfall>
  <HY_Outfall id="outflow2" xlink:href="http://nhd#1e7f53d2b29">
    <contributingBasin xlink="#basin2"/>
    <receivingBasin xlink="#basin3"/>
  </HY_Outfall>
  <HY_Outfall id="outflow3"
xlink:href="http://geobase.ca/geobase/nhn/nhn_hn_hydrojunct#4ed6574fc6fc">
    <contributingBasin xlink="#basin3"/>
  </HY_Outfall>
</HY_Features>
```


A flow line has been described with different class names, attributes and fields in NHN and NHD/NHD+ source models. In this pilot implementation, HY_Flowpath, a predefined class in HY_Features, was used to represent flow lines using the HY_Features model as a reference model. Attributes, fromNode and toNode, were created for mapping from_Junction/to_Junction in NHN and fromNode/toNode in NHD. Figure 12 shows exemplarily the mapping of nlflow in NHN, flowline in NHD, and PlusFlowline in NHD+ to HY_FlowPath in HY_Features.

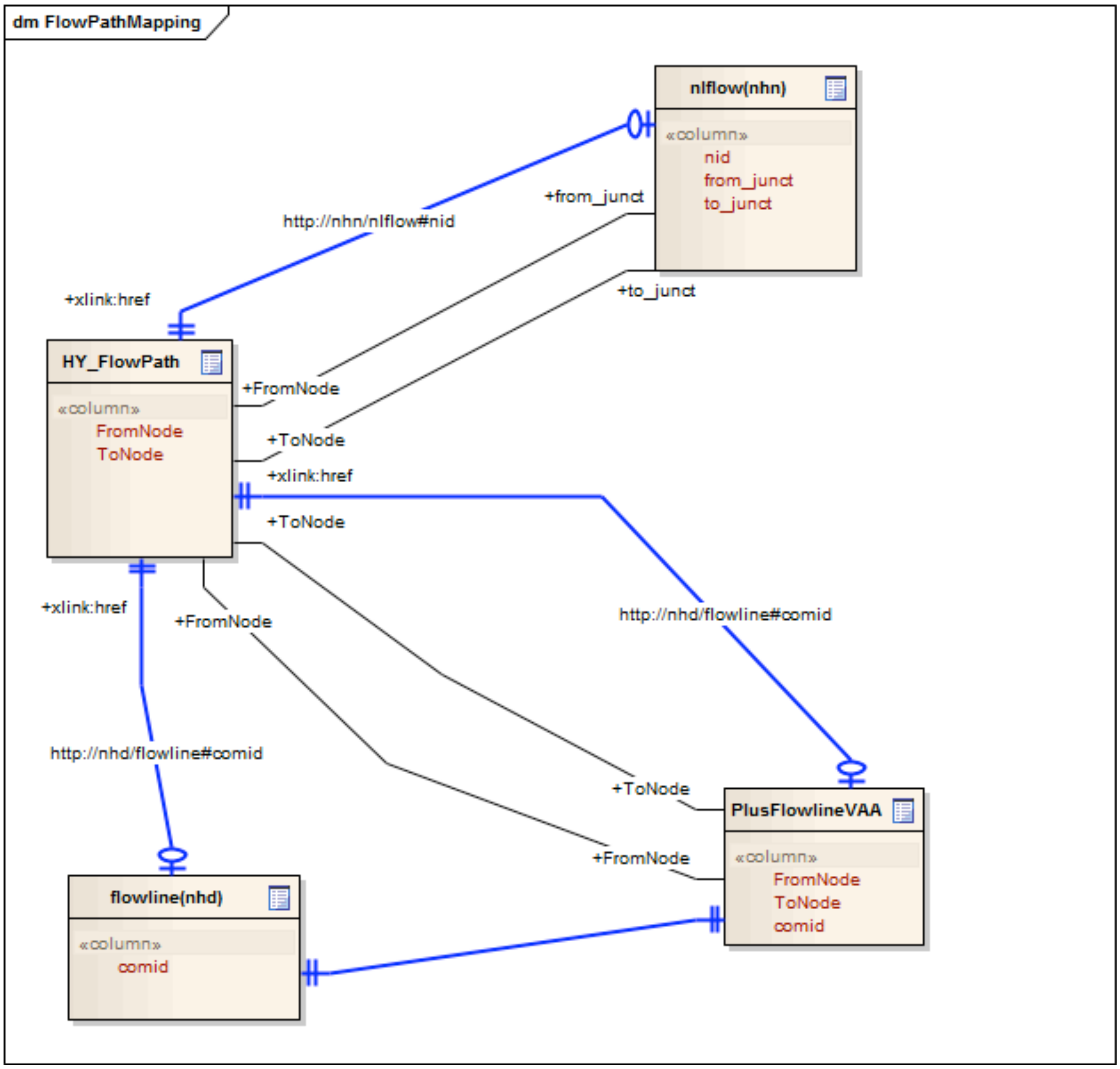

Figure 12 — Flow line mapping using Xlink

6.3.5 Semantic mapping implementation and example queries

The semantic mapping was implemented as a prototype OWL-based knowledgebase. The knowledgebase was made available at http://59.120.223.164:443/hy_feature_mapping/query (last accessed 7/16/2014). It is a standard SPARQL endpoint.

Using the SPARQL engine, we can query different matching models, entities, and fields. In the following example, we focus on flowpath in terms of HY_Features. Table 2 shows the query to find out what entities available in HY_Features model. Figure 13 shows the results returned from the prototype SPARQL engine.

Table 2. An example query to find the entities in HY_Features

```
prefix mapf: <http://modelmapping/framework#>
SELECT distinct ?entity
WHERE {
   ?entity mapf:belongModel  <http://hy_feature>
}
```

```
-------------------
| entity          |
===================
| hyf:HY_FlowPath |
-------------------
```

Figure 13. Result of the query in Table 2

The fields supported by the entity HY_FlowPath can be further found out through the query shown in Table 3. Figure 14 shows the fields returned from the prototype SPARQL engine.

Table 3. An example query to find the fields of HY_FlowPath in HY_Features model

```
prefix mapf: <http://modelmapping/framework#>
prefix hyf: <http://hy_feature#>
SELECT distinct ?hy_field
WHERE {
  ?hy_field mapf:belongEntity hyf:HY_FlowPath.
}
```

```
--------------------------------
| hy_field                     |
================================
| maph:HY_FlowPath_OutFlowNode |
| maph:HY_FlowPath_InFlowNode  |
| hyf_fp:HY_FlowPath_ID        |
--------------------------------
```

Figure 14. Result of the query in Table 3

Semantic mapping models for the entity can be found using the query shown in Table 4. Figure 15 shows the models that can be mapped to by the HY_FlowPath in HY_Features model. The SPARQL engine returns two matching models – NHD+ and NHN.

Table 4. An example query to find the models to be mapped from HY_FlowPath in HY_Features model

```
SELECT distinct ?model
WHERE {
  ?hy_field mapf:belongEntity hyf:HY_FlowPath.
  ?mapping mapf:targetField ?hy_field.
  ?mapping mapf:belongMappingModel ?mappingModel.
  ?mappingModel mapf:sourceModel ?model.
}
```

```
----------------
| model        |
================
| <http://nhd> |
| <http://nhn> |
----------------
```

Figure 15. Result of the query in Table 4

6.4 Use cases

6.4.1 Use case 1: Mapping hydrographic features from NHD+ to HY_Features

6.4.1.1 Description

The first use case is raised when user wants to convert a hydrographic feature in NHD+ database to a common feature in HY_Features. The following describes the general sequence of this use case.

1. Select a watershed or an area of interest (AOI) that lies in US where original data are stored and managed in NHD+ database or feature store.

2. The request of hydrographic features is sent to Hydro Model Service. The Hydro Model Service contact the NHD+ feature data store to retrieve the hydrographic features falling within AOI in original NHD+ model format.

3. The Hydro Model Service interacts with Hydro Mediation Service to semantically map hydrographical features in NHD+ to the counterparts in HY_Features.

4. The Hydro Model Service returns resulted feature collection in HY_Features to client in HY_Features model. The results can be direct feature stream or a feature collection stored and managed by a persistent feature data store such as WFS.

5. Provenance or lineage metadata is available to relate the identity of features in source models to HY_Features types.

Figure 16 shows the details in implementing the information mapping of a flowline feature in NHD+ to a HY_FlowPath feature in HY_Features. To support the mapping of different models through HY_Features for flow path, two fields of CustomizedField were added. These fields are HY_FlowPath_InFlowNode and HY_FlowPath_OutFlowNode to represent flow nodes to be mapped. The mapping started with establishing a MappingNHD as a MappingModel class to represent NHD+ features. The sourceModel of MappingNHD will be set as NHD and its targetModel is set as HY_Features. A FieldMapping class, mapping_flowline_ToNode is defined to link the CustomizedField – HY_FlowPath_OutFlowNode to fields in NHD+ model. The mappingField of mapping_flowline_ToNode is set to ToNode and its targetField is set to HY_FlowPath_OutFlowNode. Similarly, a FieldMapping class, mapping_flowline_FromNode is to map FromNode.

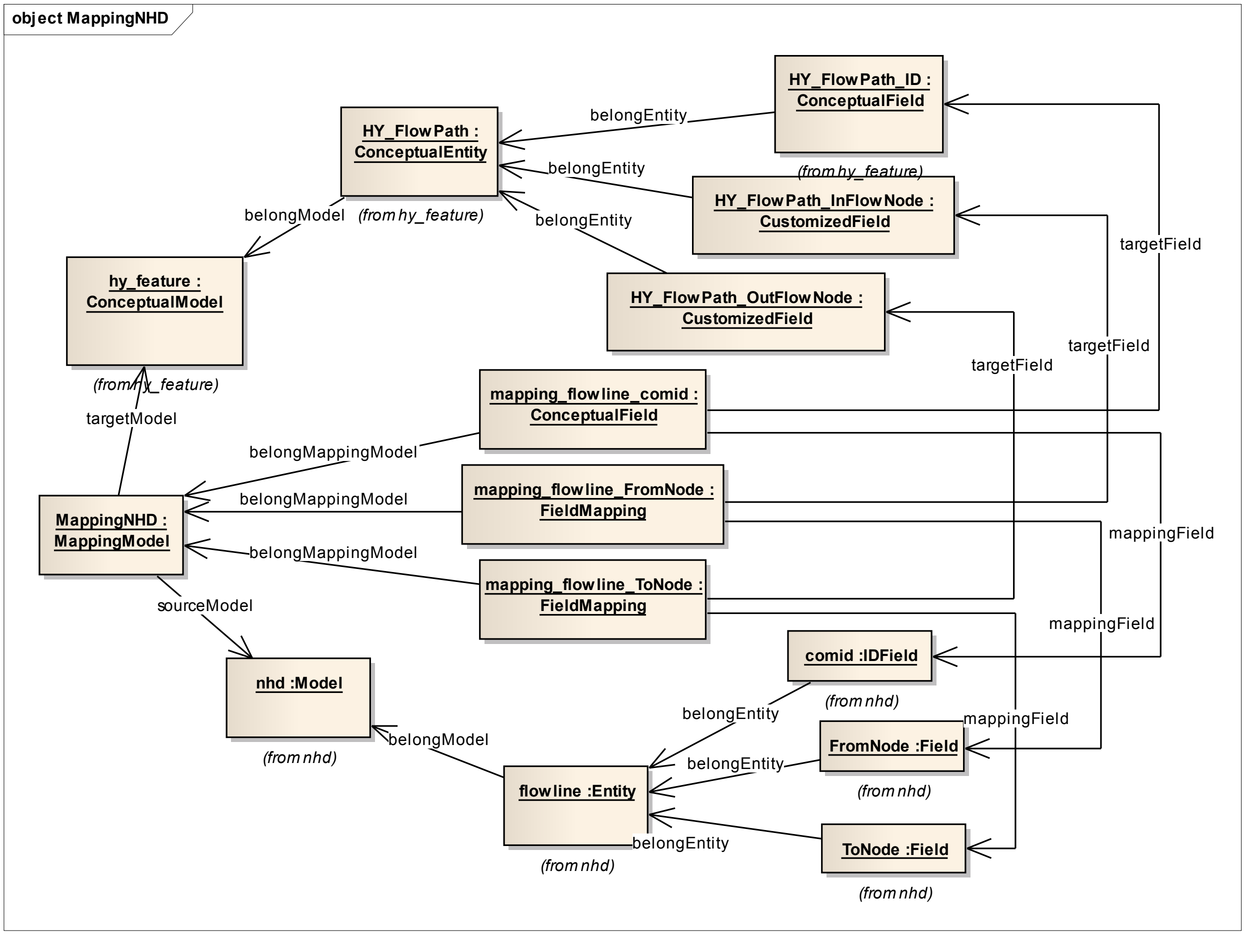

Figure 16 — Use case 1: mapping flow line from NHD+ to HY_Features

6.4.1.2 Demonstration

Demonstration endpoint: http://59.120.223.164/hydro/ (last accessed 7/16/2014). Figure 17 and Figure 18 show the two major steps to use Hydro Model Service and Hydro Mediation Service to map hydrographic features in NHD+ to those in HY_Features. As shown in Figure 17, an AOI covering between 97.237ºW, 48.984ºN, and 97.220ºW, 48.992ºN. The AOI lies in US. It covers a small area across Red River between North Dakota and Minnesota, close to the border of Canada.

Figure 18 shows the results in HY_Features returned from the Hydro Model Service.

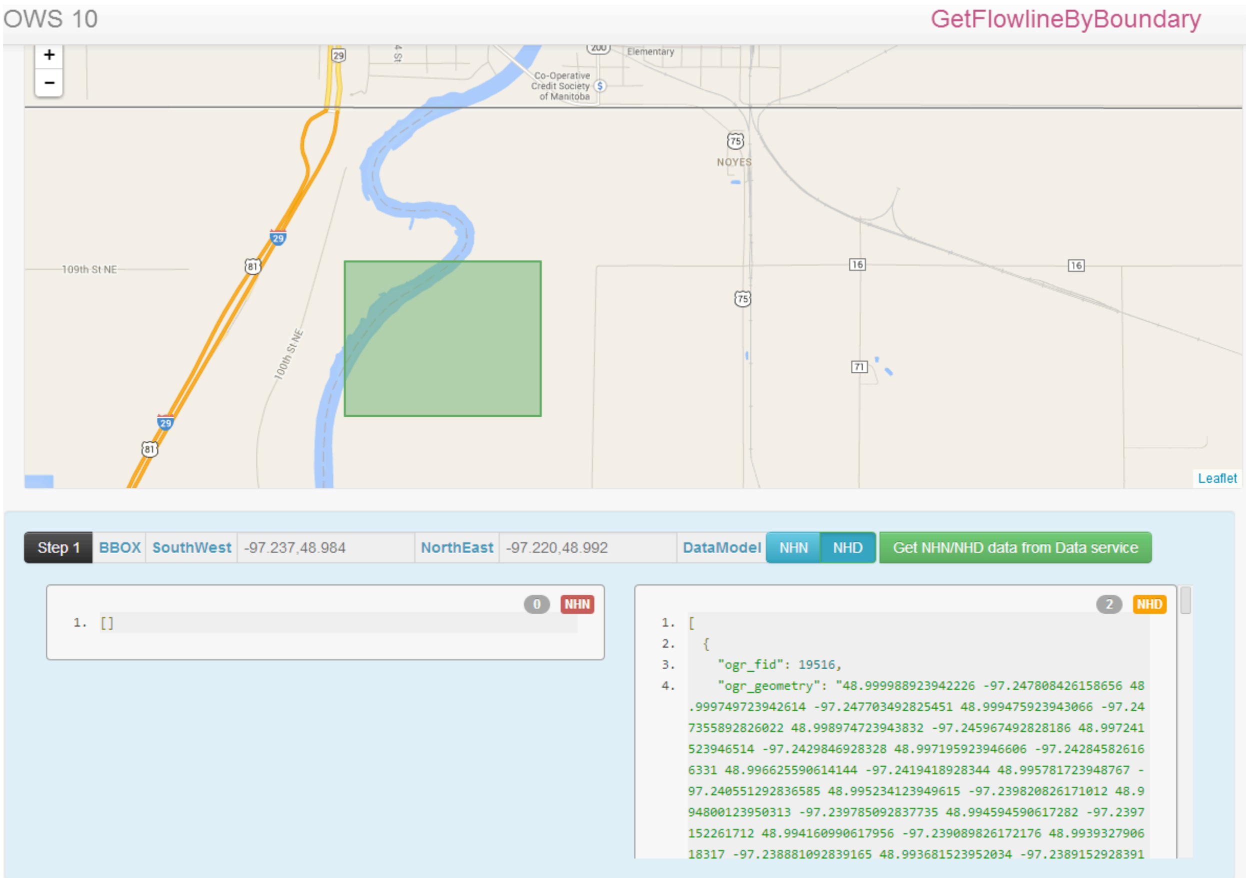

Figure 17 — Use case 1 step 1: Define an area of interest in US

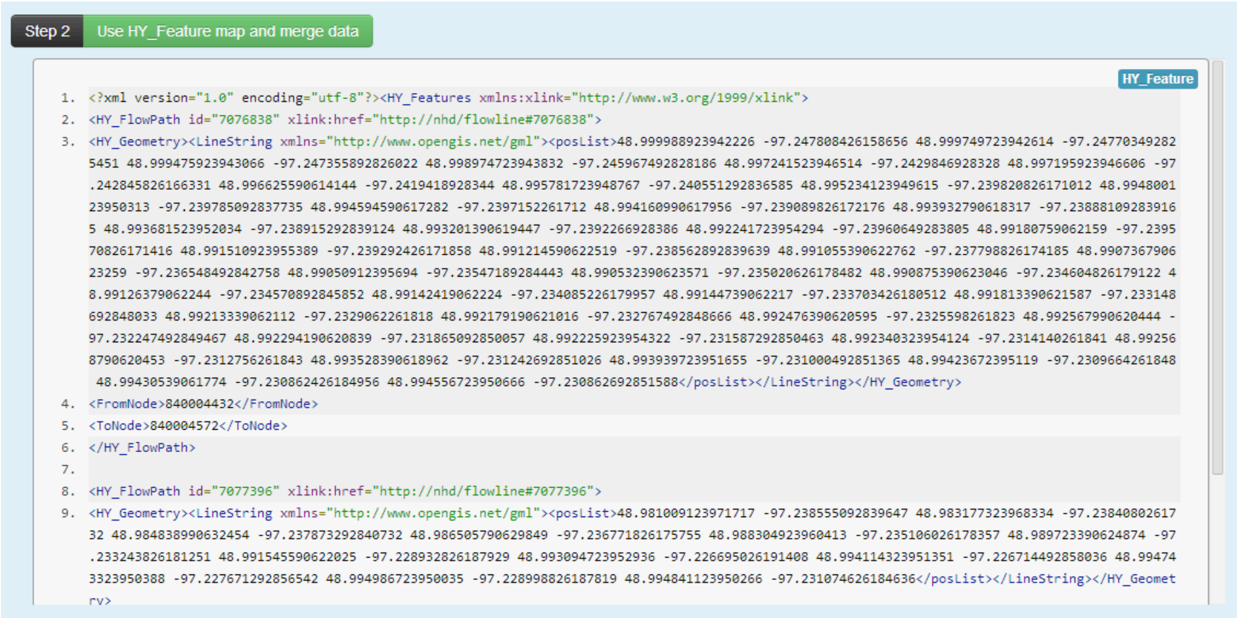

Figure 18 — Use case 1 step 2: Retrieve results in HY_Features

Semantically, the match model can be queried through the SPARQL engine described in Section 6.3.5. Table 5 shows the example query to find the mapping models, entities and fields between HY_FlowPath in HY_Features to those in NHD+. Figure 19 the results returned from the SPARQL engine given in Section 6.3.5.

Table 5. An example query to find the mapping of HY_FlowPath to NHD+

```
SELECT distinct ?hy_field ?entity ?source_field
WHERE {
  ?mapping rdf:type mapf:FieldMapping.
  ?mapping mapf:targetField ?hy_field.
  ?mapping mapf:mappingField ?source_field.
  ?hy_field mapf:belongEntity hyf:HY_FlowPath.
  ?source_field mapf:belongEntity ?entity.
  ?entity mapf:belongModel <http://nhd>.
}
```

```
--------------------------------------------------------------------------
| hy_field                      | entity                  | source_field      |
==========================================================================
| hyf_fp:HY_FlowPath_ID         | <http://nhd/flowline>   | flowline:comid    |
| maph:HY_FlowPath_InFlowNode   | <http://nhd/flowline>   | flowline:FromNode |
| maph:HY_FlowPath_OutFlowNode  | <http://nhd/flowline>   | flowline:ToNode   |
--------------------------------------------------------------------------
```

Figure 19. Result of the query in Table 5

6.4.2 Use case 2: Mapping hydrographic features from NHN to HY_Features

6.4.2.1 Description

The first use case is raised when user wants to convert a hydrographic feature in NHN database to common feature in HY_Features. The following describes the general sequence of this use case.

1. Select a watershed or an area of interest (AOI) that lies in Canada where original data are stored and managed in NHN database or feature store.
2. The request of hydrographic features is sent to Hydro Model Service. The Hydro Model Service contacts the NHN feature data store to retrieve the hydrographic features falling within AOI in original NHN model format.
3. The Hydro Model Service interacts with Hydro Mediation Service to semantically map hydrographical features in NHN to the counterparts in HY_Features.
4. The Hydro Model Service returns resulted feature collection in HY_Features to client in HY_Features model. The results can be direct feature stream or a feature collection stored and managed by a persistent feature data store such as WFS.
5. Provenance or lineage metadata is available to relate the identity of features in source models to HY_Features types.

Figure 20 shows the details in implementing the information mapping of nlflow in NHN model to Flowpath in HY_Features model. To support the mapping of different models through HY_Features for flow path, two fields of CustomizedField were added. They are HY_FlowPath_InFlowNode and HY_FlowPath_OutFlowNode to represent flow nodes to be mapped. The mapping started with establishing a MappingNHN as a MappingModel class to represent NHN features. The souceModel of MappingNHN will be set as NHN and its targetModel is set as HY_Features. A FieldMapping class, mapping_nlflow_from_junct is defined to link the CustomizedField – HY_FlowPath_InFlowNode to fields in NHN model. The mappingField of mapping_nlflow_from_junct is set to from_junct and its targetField is set to HY_FlowPath_InFlowNode. Similarly, a FieldMapping class, mapping_nlflow_to_junct is to map to_junct.

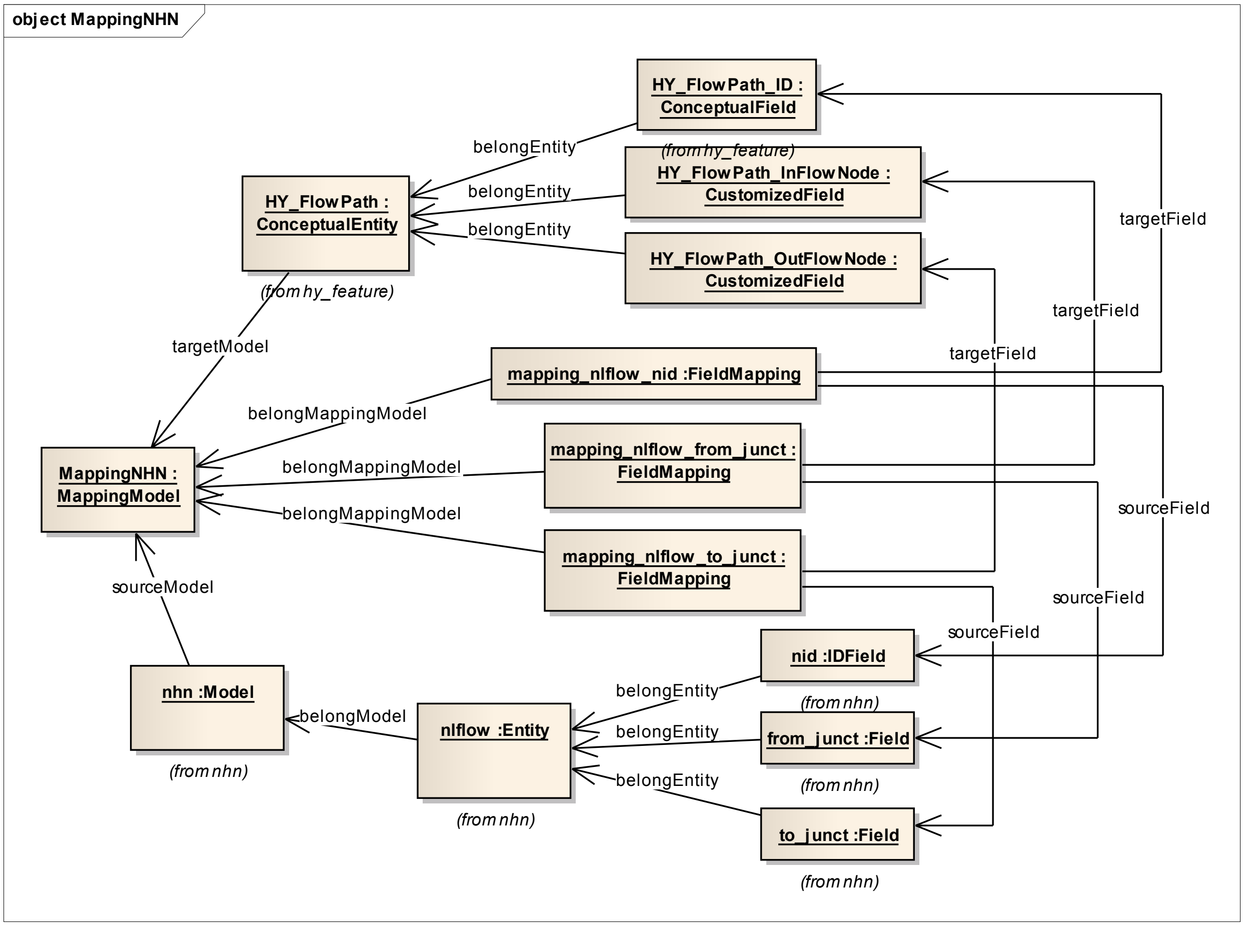

Figure 20 — Use case 2: Mapping flow line from NHN to HY_Features

6.4.2.2 Demonstration

Demonstration endpoint: http://59.120.223.164/hydro/ (last accessed 7/16/2014). Figure 21 and Figure 22 show the two major steps to use Hydro Model Service and Hydro Mediation Service to map hydrographic features in NHN to those in HY_Features. As shown in Figure 21, an AOI covering between 97.363ºW, 49.332ºN, and 97.300ºW, 49.366ºN. The AOI lies in Canada. It covers a small area across Red River in Manitoba, Canada.

Figure 22 shows the results in HY_Features returned from the Hydro Model Service.

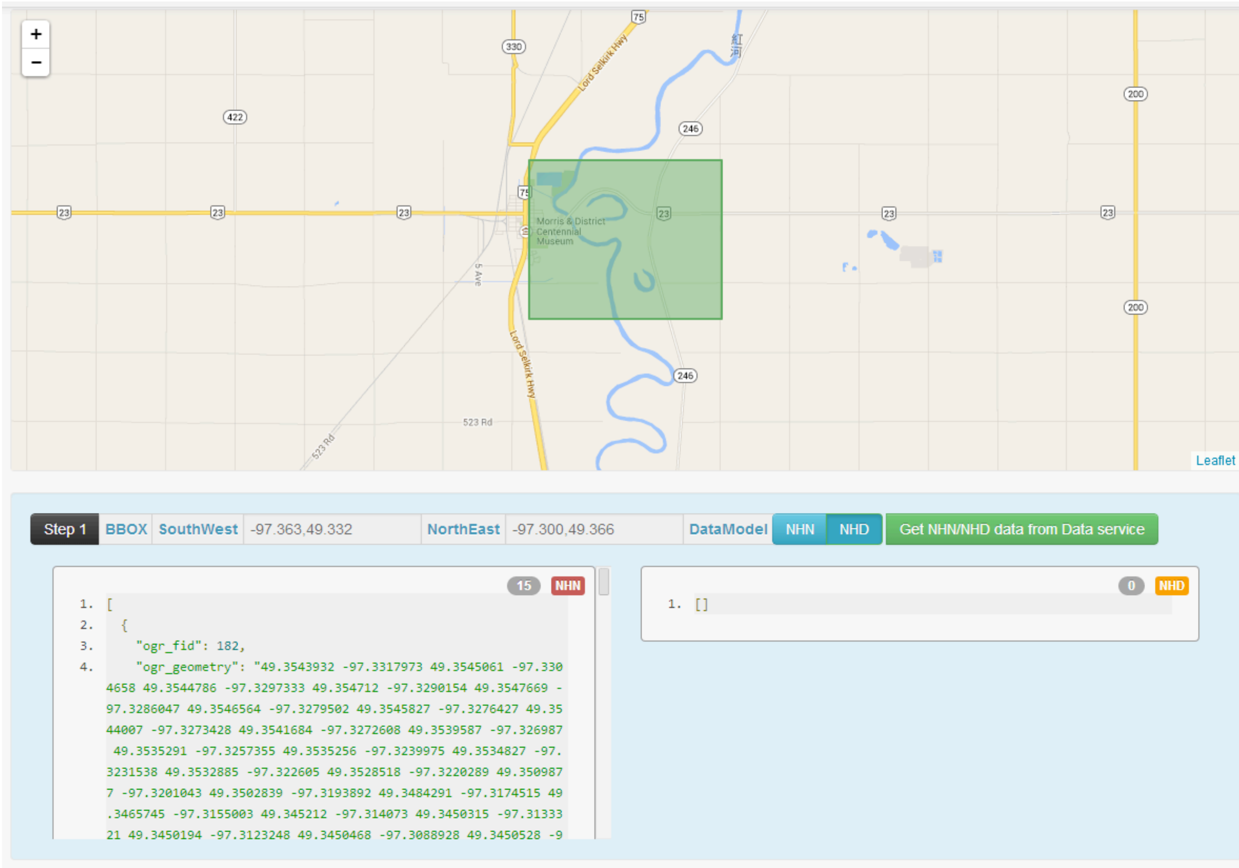

Figure 21 — Use case 2 step 1: Define an area of interest In Canada

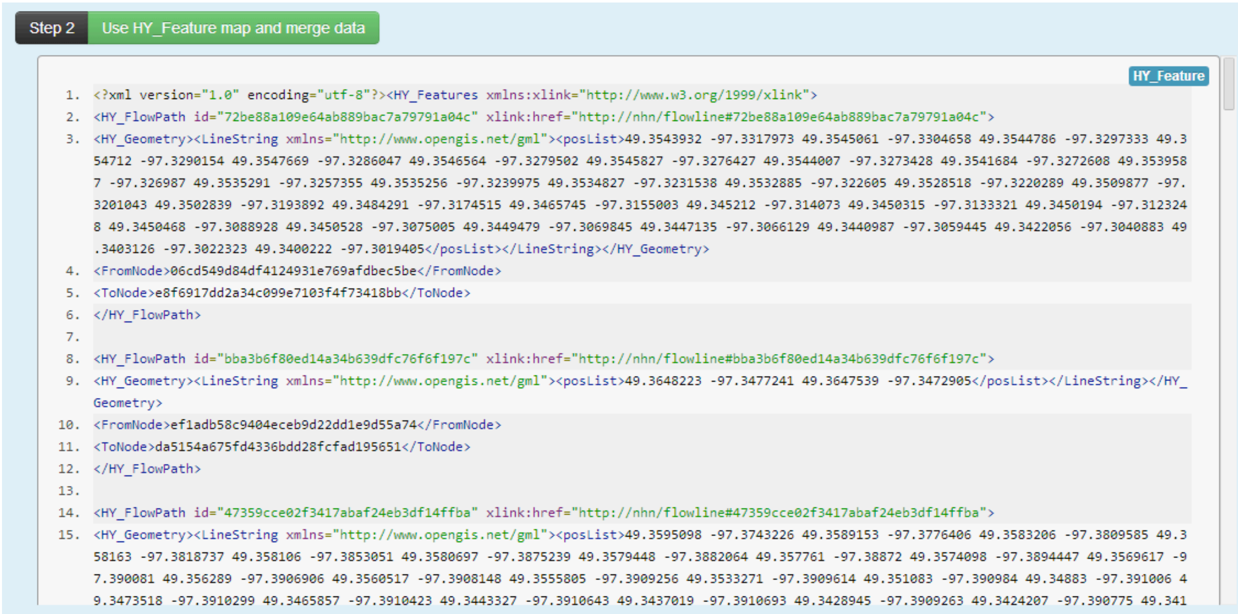

Figure 22 — Use case 2 step 2: Retrieve results in HY_Features

Similar to the semantic matching for conversion from NHD+ to HY_Features, the match model can also be queried through the SPARQL engine described in Section 6.3.5. Table 6 shows the example query to find the mapping models, entities and fields between HY_FlowPath in HY_Features to those in NHN. Figure 23 the results returned from the SPARQL engine given in Section 6.3.5.

Table 6. An example query to find the mapping of HY_FlowPath to NHN

```
SELECT distinct ?hy_field ?entity ?source_field
WHERE {
  ?mapping rdf:type mapf:FieldMapping.
  ?mapping mapf:targetField ?hy_field.
  ?mapping mapf:mappingField ?source_field.
  ?hy_field mapf:belongEntity hyf:HY_FlowPath.
  ?source_field mapf:belongEntity ?entity.
  ?entity mapf:belongModel <http://nhn>.
```

```
-------------------------------------------------------------------
| hy_field                       | entity     | source_field       |
===================================================================
| hyf_fp:HY_FlowPath_ID          | nhn:nlflow | nlflow:nid         |
| maph:HY_FlowPath_InFlowNode    | nhn:nlflow | nlflow:from_junct  |
| maph:HY_FlowPath_OutFlowNode   | nhn:nlflow | nlflow:to_junct    |
-------------------------------------------------------------------
```

Figure 23. Result of the query in Table 6

6.4.3 Use case 3: Hydrographic mapping between NHD+ and NHN

6.4.3.1 Description

This use case applies to clients who want to map one format or hydro model to another using HY_Features model as the mediation model. In this demonstration implementation setting, there are two cases corresponding to the use of mapping functions demonstrated in Use Case 1 in Section 6.4.1 and Use Case 2 in Section 6.4.2. Here are the details for two sub-cases:

1. Use Sub-Case 1: First, we apply Use Case 1 (See Section 6.4.1) to map hydrographic features in NHD+ model to those in HY_Features model. Second, we apply Use Case 2 (Refer to Section 6.4.2) in reverse mapping direction to map the features in HY_Features to those in NHN model. Figure 24 shows the complete sequence of interactions. The results are presented to clients in NHN model.

2. Use Sub-Case 2: First, we apply Use Case 2 (See Section 6.4.2) to map hydrographic features in NHN model to those in HY_Features model. Second, we apply Use Case 1 (Refer to Section 6.4.1) in reverse mapping direction to map the features in HY_Features to those in NHD+ model. Figure 25 shows the complete sequence of interactions. The results are presented to clients in NHD+ model.

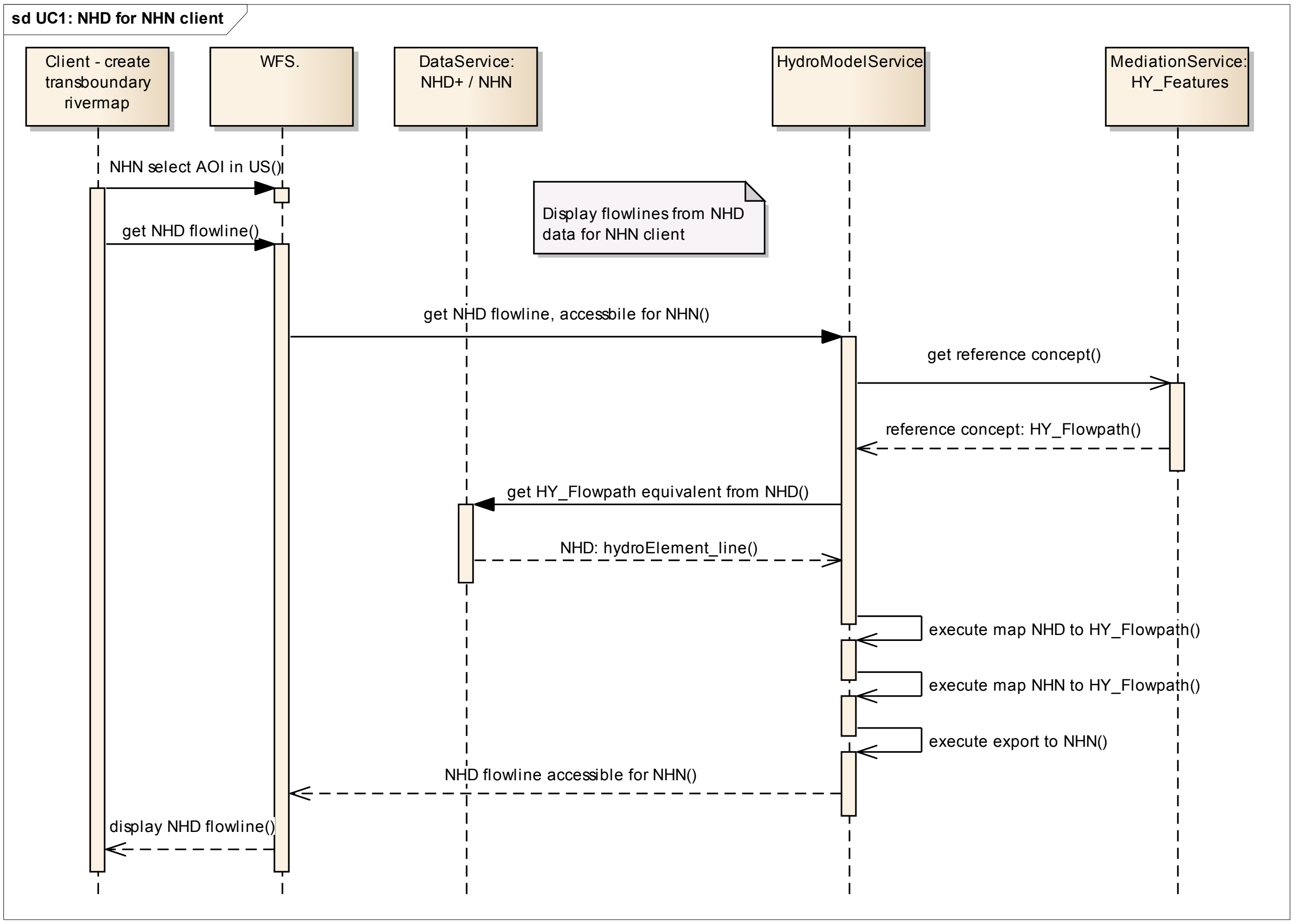

Figure 24 — Sequence graph of Use Case 3.1: Map NHN to NHD+

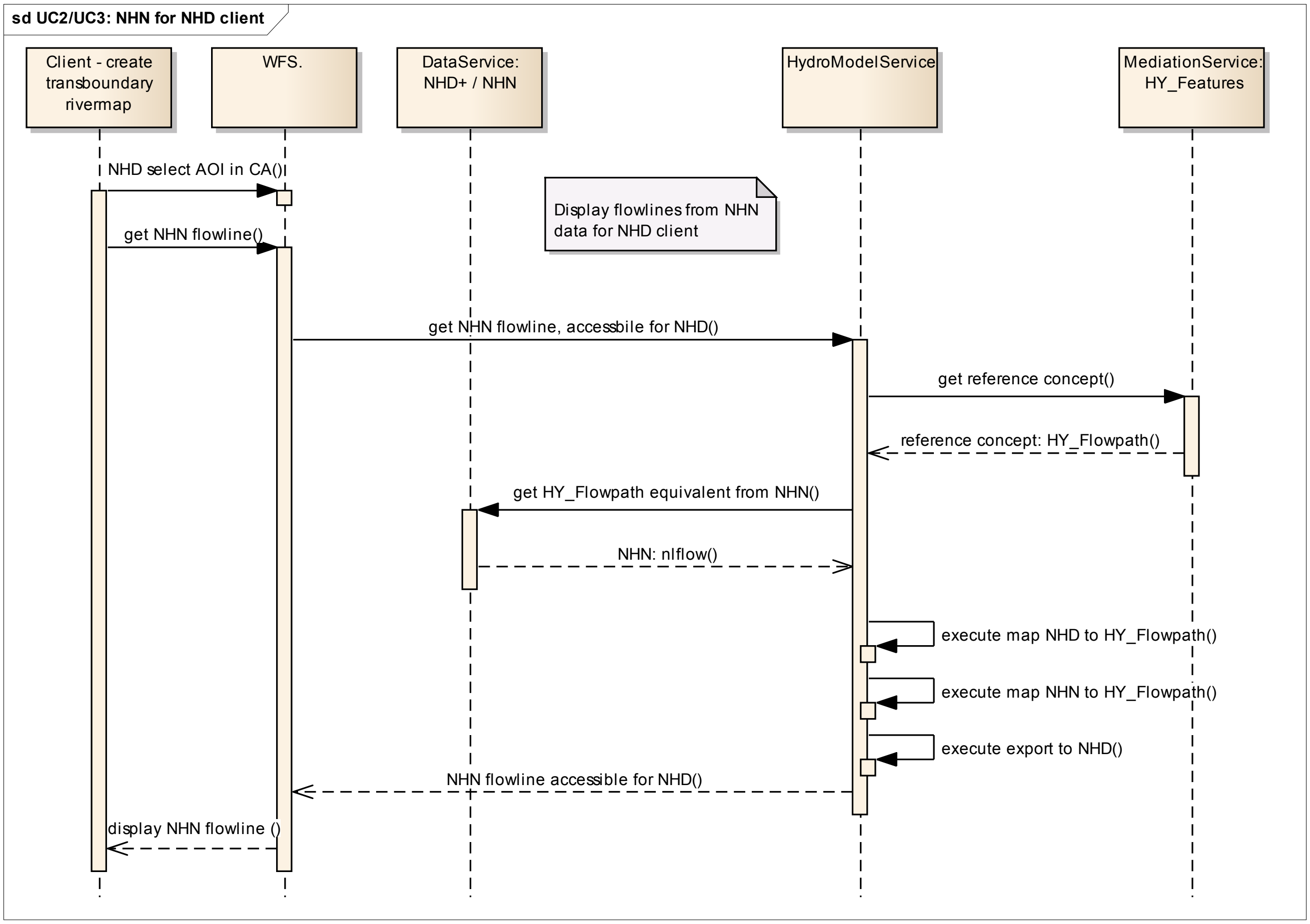

Figure 25 — Sequence graph of Use Case 3.2: Map NHD+ to NHN

6.4.3.2 Demonstration

Demonstration endpoint: http://59.120.223.164/hydro/ (last accessed 7/16/2014). The use case can be demonstrated by taking one step further from those demonstrated in Use Case 1 or Use Case 2 to convert features from one model to another. For example, we can continue on the demonstration case in Section 6.4.2.2 and take one step further to map those features expressed in HY_Features model to those in NHD+. The results are in NHD+ as shown in Figure 26. Similar demonstration can be done with the AOI in Section 6.4.1.2 to map NHD+ to NHN.


```
<?xml version="1.0" encoding="utf-8"?><ogr:FeatureCollection xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:ogr="http://nhd/flowline" xmlns:gml="http://www.opengis.net/gml/3.2">
<gml:featureMember>
<ogr:Feature gml:id="72be88a109e64ab889bac7a79791a04c">
<ogr:geometryProperty><LineString xmlns="http://www.opengis.net/gml" gml:id="72be88a109e64ab889bac7a79791a04c.geom"><posList>49.3543932 -97.3317973 49.3545061 -97.3304658 49.3544786 -97.3297333 49.354712 -97.3290154 49.3547669 -97.3286047 49.3546564 -97.3279502 49.3545827 -97.3276427 49.3544007 -97.3273428 49.3541684 -97.3272608 49.3539587 -97.326987 49.3535291 -97.3257355 49.3535256 -97.3239975 49.3534827 -97.3231538 49.3532885 -97.322605 49.3528518 -97.3220289 49.3509877 -97.3201043 49.3502839 -97.3193892 49.3484291 -97.3174515 49.3465745 -97.3155003 49.345212 -97.314073 49.3450315 -97.3133321 49.3450194 -97.3123248 49.3450468 -97.3088928 49.3450528 -97.3075005 49.3449479 -97.3069845 49.3447135 -97.3066129 49.3440987 -97.3059445 49.3422056 -97.3040883 49.3403126 -97.3022323 49.3400222 -97.3019405</posList></LineString></ogr:geometryProperty>

<ogr:comid/>
<ogr:enabled/>
<ogr:fcode/>
<ogr:fdate/>
<ogr:flowdir/>
<ogr:FromNode>06cd549d84df4124931e769afdbec5be</ogr:FromNode>
<ogr:ftype/>
<ogr:gnis_id/>
<ogr:gnis_name/>
<ogr:gnis_nbr/>
<ogr:lengthkm/>
<ogr:ogr_geometry/>
```


Figure 26 — Use Case 3.1 demo: Results in NHD+

6.4.4 Use case 4: Finding upstream service

6.4.4.1 Description

The use case represents the situation when a user wants to find all the hydrographic features following an upstream tracing. Figure 27 shows the typical sequence of finding upstream service. The following describe the general sequence briefly.

1. The user chooses a point, either in US or in Canada and submits a request to the Hydro Model Service. To simplify description, we assume that we choose one point in Canada as the first starting point. We also decide to retrieve the results in NHN model. (Note: In the following steps, we will take some assumption to walk through one logic branch while ignoring another fork, to shorten the description. In all cases, the walkthrough for another fork works similarly.)

2. The Hydro Model Service search for hydrographic features within a searching threshold (e.g. 1 kilometer buffer). The searching request can be either broadcasted to all data store services or specifically sent to the data store if there is a prior knowledge through which given location determines data store or model. For simple description, let us assume that we adopt the broadcasting approach.

3. Of the two receiving data stores – NHD+ and NHN, at the starting point, only does the NHN data store send back hydrographic features. The Hydro Model Service adds it to the feature collection to be returned.

4. The Hydro Model Service determines the upstream endpoint of the found hydrographic features and broadcasts search request for the endpoints to all data stores again.

5. All data stores return the hydro features connecting to the endpoints. If the feature is returned from NHN data store in NHN model, the Hydro Model Service adds it to the feature collection to be returned.

6. If the feature is returned from NHD+ data store in NHD+ model, the Hydro Model Service sends the feature in NHD+ model to the Hydro Mediation Service for mapping it to feature in NHN model. The Hydro Model Service adds the converted feature in NHN model to the feature collection to be returned.

7. The Hydro Model Service repeats step 4-6 until there is no more connecting hydro feature found along the upstream trace.

8. The Hydro Model Service returns the results in NHN model.

The above sequence can have the following variations:

(1) The user chooses a point in Canada and decides to get results in NHN model (This is what the above sequence describes.)

(2) The user chooses a point in US and decides to get results in NHN model.

(3) The user chooses a point in Canada and decides to get results in HY_Features model.

(4) The user chooses a point in US and decides to get results in HY_Features model.

(5) The user chooses a point in Canada and decides to get results in NHD+ model.

(6) The user chooses a point in US and decides to get results in NHD+ model.

To all the above variations, we should be able to find the corresponding variations of sequence following the way we describe the sequence for variation (1).

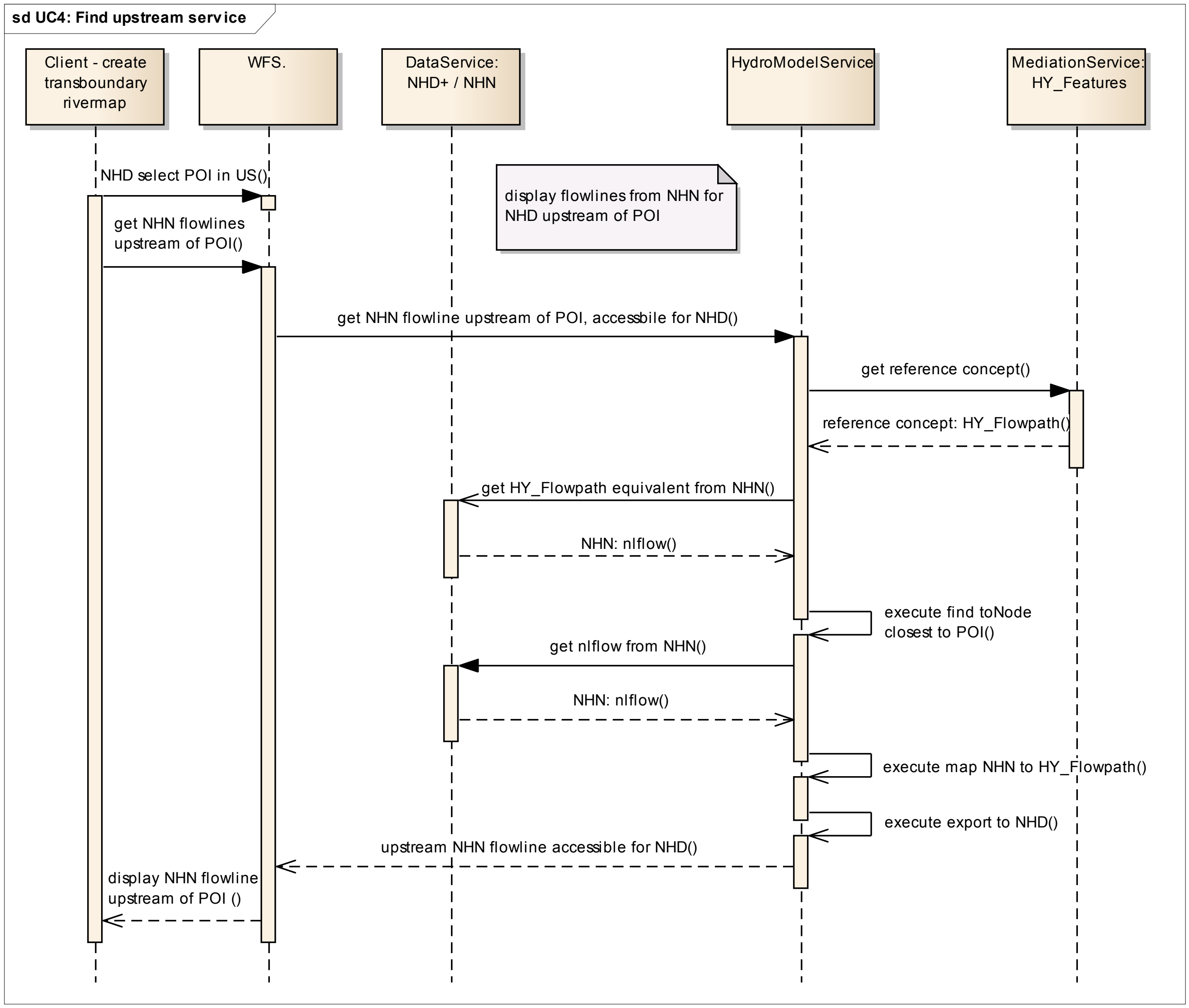

Figure 27 — Use case 4: Finding upstream service

6.4.4.2 Demonstration

Demonstration endpoint: http://59.120.223.164/hydro/ (last accessed 7/16/2014). In this demonstration, the user chooses a start point at 97.189ºW, 49.038ºN, a point located in Ginew, Manitoba R0A 2R0, Canada. Figure 28 shows the starting point, the initial request and response from Hydro Mediation Service. The following brief the major steps.

1. The hydrographic feature falling within the threshold is Joe River, a branch to Red River. The search begins to find upstream of Joe River. See Figure 28 for illustration of location and initial interaction between the Hydro Model Service and the Hydro Mediation Service.

2. The first several found sections of Joe River are features in NHN model, in Canada.

3. At the border between US and Canada, the connected upstream of Joe River is found and search continues. The Hydro Model Service interacts with the Hydro Mediation Service map found features in NHD+ model in US to corresponding features in NHN model.

4. The search end at the upstream source of Joe River in Minnesota, US. The complete feature collection in NHN model is returned to the user by the Hydro Model Service. Figure 29 shows the found feature collection of Joe Rive, covering both sections in Manitoba, Canada and Minnesota, US.

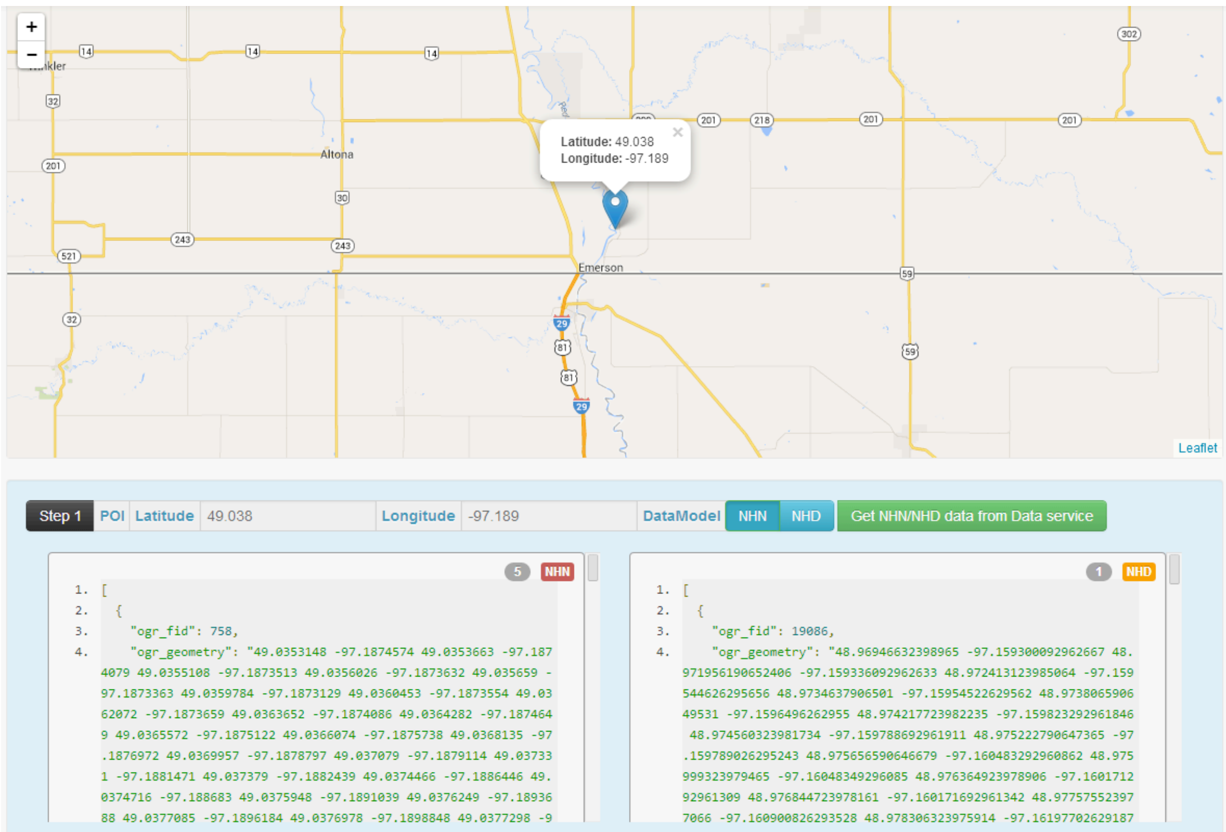

Figure 28 — Use case 4 demo: Starting point and requests

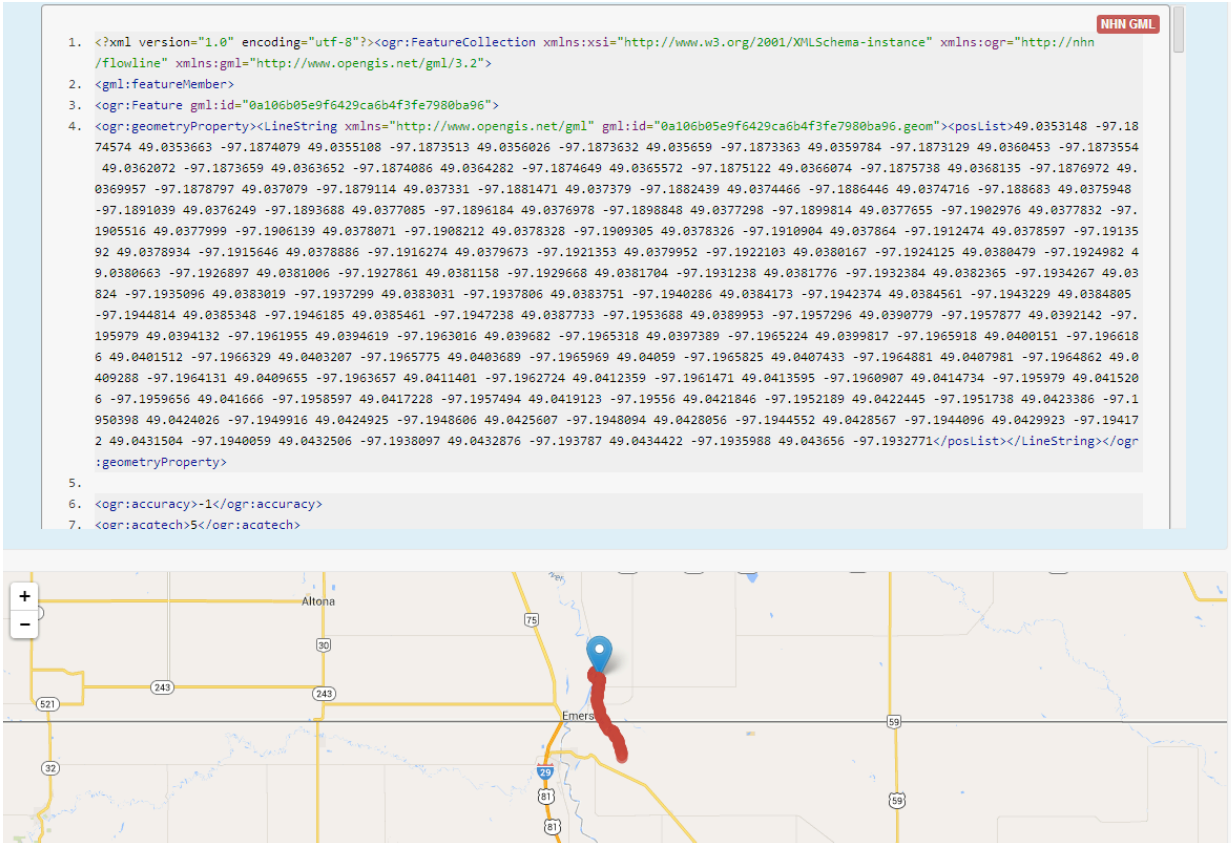

Figure 29 — Use case 4 demo: Result feature collection in NHN model

6.4.5 Use case 5: Heterogeneous hydrological model integration mapping (NHN/NHD+)

6.4.5.1 Description

Use case 5 is the integrated mapping of heterogeneous hydrographic models into one target hydrographic model. Figure 30 shows one typical sequence of converting features under heterogeneous hydro models into one target hydro model. The following describe the general sequence.

1. The user selects an area of interest (AOI). The AOI covers an area that has hydrographic features in different hydrographic models. In this scenario, a border area across US and Canada that covers hydrographic features in NHD+ model and NHN model respectively. The user sends in the request for retrieving all hydrographic features intersected with the AOI and requires the results returned in NHD+ model.

2. The Hydro Model Service receives the request from the user. It broadcasts the search for hydrographic features intersected with the AOI to all data stores, including NHD+ data store and NHN data store.

3. The Hydro Model Service retrieves all hydrographic features intersected with the AOI from NHD+ data store in NHD+ model. The returned features in NHD+ model are sent to the Hydro Mediation Service to be mapped into features in HY_Features. It adds all the returned features in HY_Features model into the intermediate feature collection in HY_Features model.

4. The Hydro Model Service retrieves all hydrographic features intersected with the AOI from NHN data store in NHN model. The returned features in NHN model are sent to the Hydro Mediation Service to be mapped into features in HY_Features. It merges all the returned features in HY_Features model into the intermediate feature collection in HY_Features model.

5. The merged features in HY_Features model are then sent to the Hydro Mediation Service to be mapped to NHD+ model. The mapped feature collection is returned to the Hydro Model Service.

6. The Hydro Model Service returns all the merged and converted feature collection in NHD+ model that the AOI intersects with.

The above sequence can have the following variations:

(1) The user chooses an AOI across borders of Canada and US. The user decides to require result hydro features in NHD+ model (This is what the above sequence describes.)

(2) The user chooses an AOI across borders of Canada and US. The user decides to require result hydro features in NHN model

(3) The user chooses an AOI across borders of Canada and US. The user decides to require result hydro features in HY_Features model.

To all the above variations, we should be able to find the corresponding variations of sequence following the way we describe the sequence for variation (1).

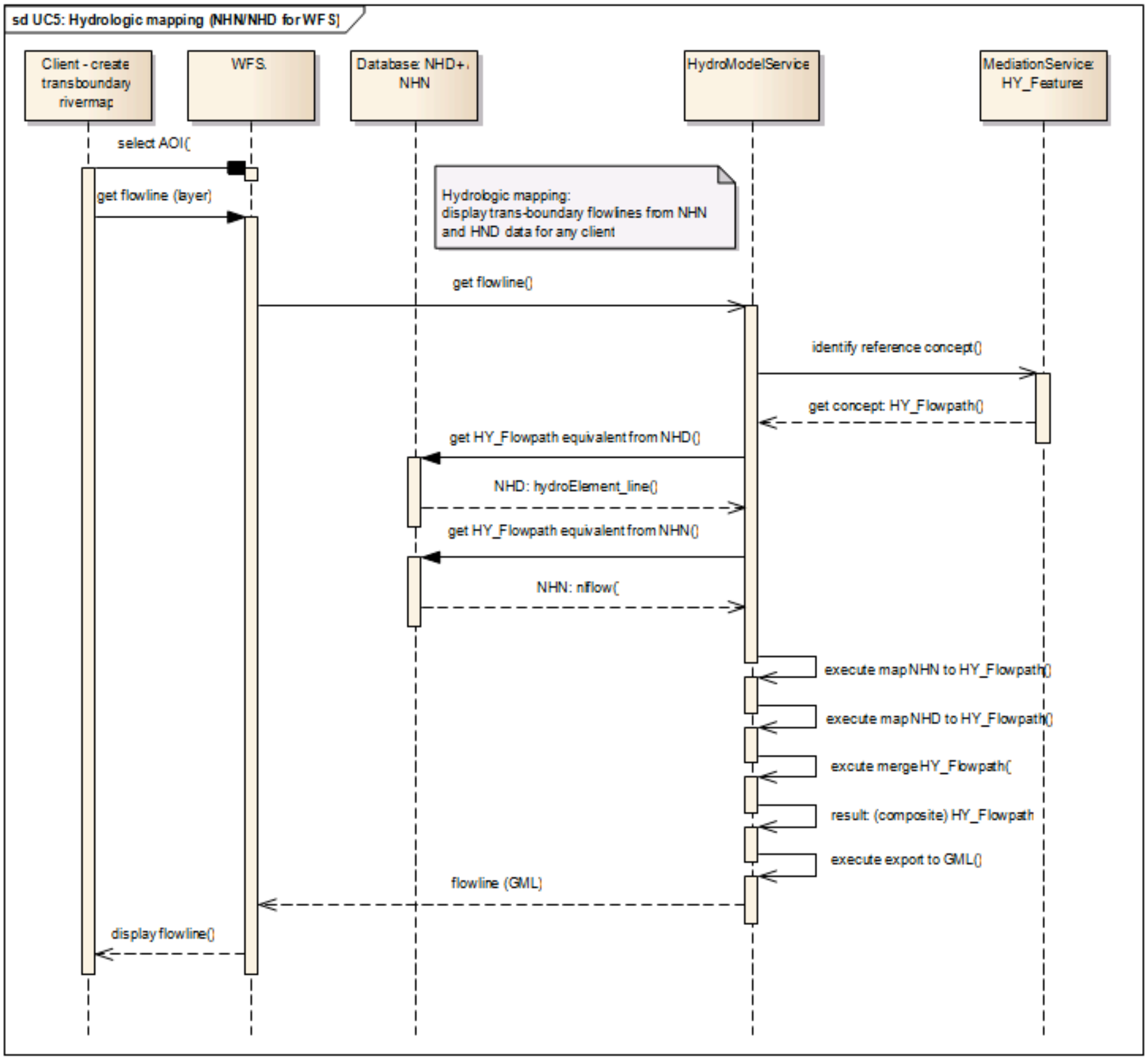

Figure 30 — Use Case 5: Hydrographic Mapping of Heterogeneous Hydro Models

There was one problem rised for crossing model data integration. There is data displacement due to scale and other resolution. The endpoint of a flow path may not exactly match that of the corresponding path. To enable the continous connection of flow paths, the project adopted what the sponsor suggested to establish a framework to allow manual definitioin of relationships along borders. Figure 31 shows the overall framework to set the equivalent models between two models. Figure 32 shows how to establhis the equivalent relationship between NHD+ and NHN models.

Extra information is required to connect flowlines from two different data models across borders, especially when there is difference in scale, name, and/or feature types. The information about the equivalence of two instances from two different hydrological models needs to be fomulated and provided to the mapping mediator. Figure 31 shows the framework used in the demonstration. A class CrossModelData is used to define the equivalence. In semantical representations, four property predicates are defined to capture the equivalence relationhip: crossModel relates a CrossModelData to Model, crossField to the field of the model, dataField to information field, and dataEquivalent to associations of two CrossModelData instances.

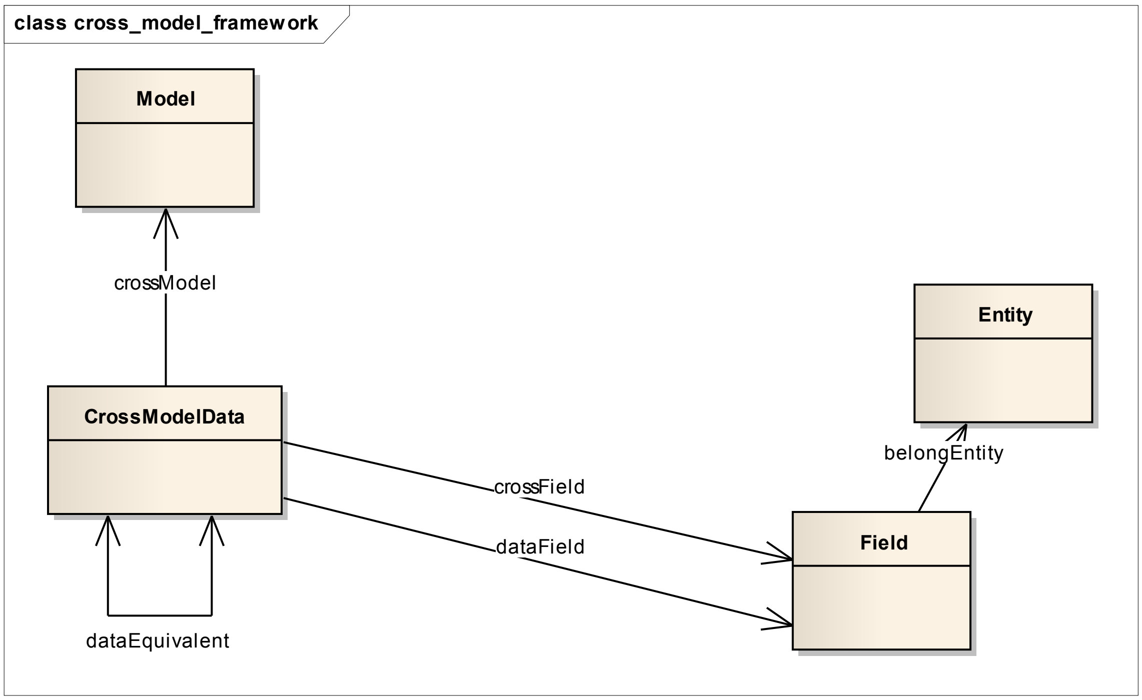

Figure 31 — Use Case 5: Cross Model Framework

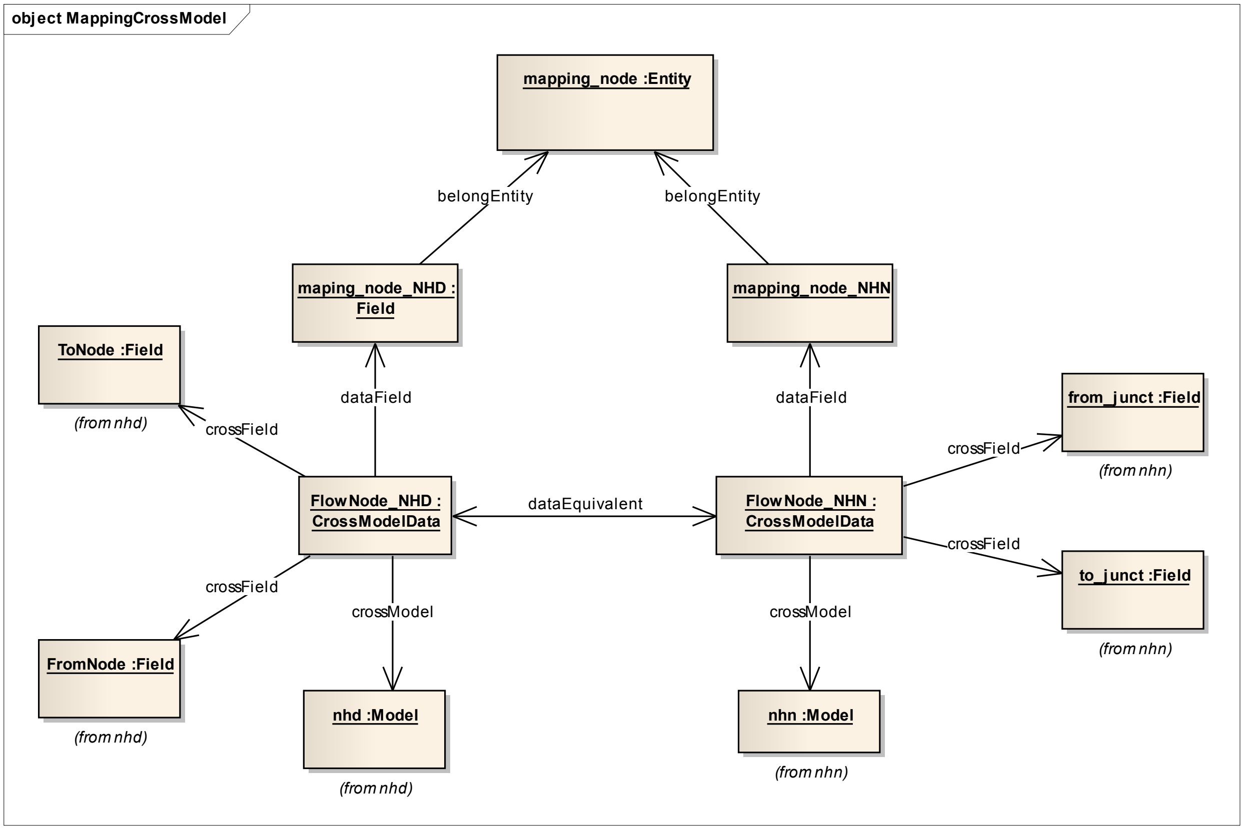

Figure 32 — Use Case 5: Cross Model Data Matching

6.4.5.2 Demonstration

Demonstration endpoint: http://59.120.223.164/hydro/ (last accessed 7/16/2014). In this demonstration, the user an AOI with bounding box 97.683ºW, 48.778ºN, and 96.639ºW, 49.182ºN. The AOI crosses the border between US and Canada, covering partial Cavalier, North Dakota 58220, US and Niverville, Manitoba R0A, Canada. Figure 33 shows the AOI and initial interactions between the Hydro Model Service and the Hydro Mediation Service. This demonstration matches the sequence described in Section 6.4.5.1. The following brief the major steps and their intermediate results.

1. The Hydro Model Service received the AOI. It interacts with data services to retrieve any hydrographic features intersecting with the AOI. The results are gathered from both NHD+ data store and NHN data store as shown in Figure 33.

2. The Hydro Model Service interacts with the Hydro Mediation Service to map linear features from both NHD+ model and NHN model to HY_Features model. All features are merged into a flow line feature collection in HY_Features model as shown in Figure 34.

3. The Hydro Model Service interacts with the Hydro Mediatiion Service to map the merged flow line feature collection in HY_Features to the target NHD+ model. Figure 35 shows the returned, merged feature collection in NHD+ model.

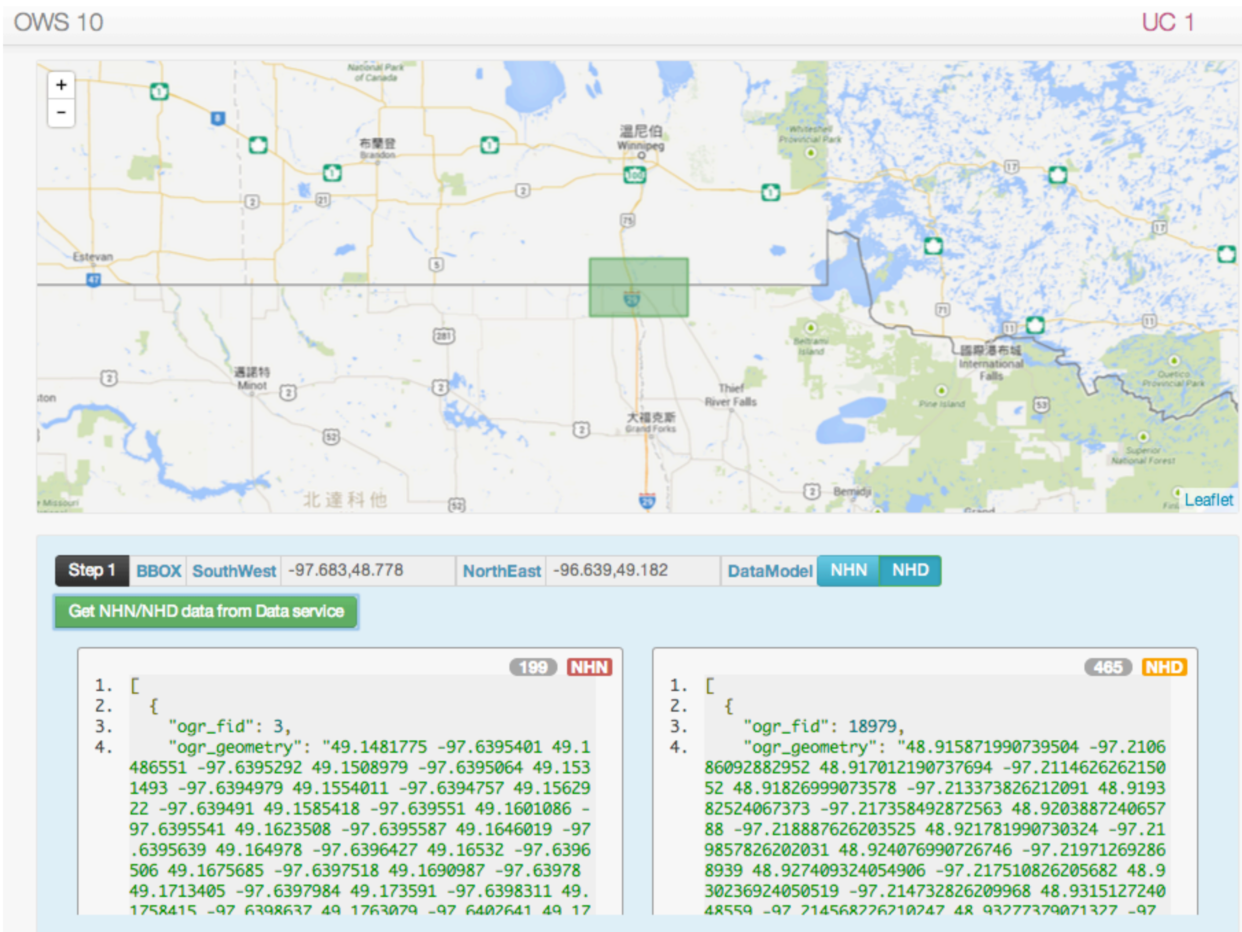

Figure 33 — Use case 5 demo: Selecting a cross-border AOI

Step 2 Use HY_Feature map and merge data

HY_Feature

```xml
<?xml version="1.0" encoding="utf-8"?><HY_Features xmlns:xlink="http://www.w3.org/1999/xlink">
<HY_FlowPath id="fe2b5d584b164a00a39085935dab2500" xlink:href="http://nhn/flowline#fe2b5d584b164a00a39085935dab2500">
<HY_Geometry><LineString xmlns="http://www.opengis.net/gml"><posList>49.1481775 -97.6395401 49.1486551 -97.6395292 49.1508979 -97.6395064 49.1531493 -97.6394979 49.1554011 -97.6394757 49.1562922 -97.639491 49.1585418 -97.639551 49.1601086 -97.6395541 49.1623508 -97.6395587 49.1646019 -97.6395639 49.164978 -97.6396427 49.16532 -97.6396506 49.1675685 -97.6397518 49.1690987 -97.63978 49.1713405 -97.6397984 49.173591 -97.6398311 49.1758415 -97.6398637 49.1763079 -97.6402641 49.176412 -97.6404081 49.1767892 -97.6406409 49.177032 -97.6407908 49.1771137 -97.6410982</posList></LineString></HY_Geometry>
<FromNode>5861580abcd0452dace1ba4da8f66632</FromNode>
<ToNode>a459ef2bdef04b13acacaaf32796d2a7</ToNode>
</HY_FlowPath>

<HY_FlowPath id="6583108717734dc7811789f002165a63" xlink:href="http://nhn/flowline#6583108717734dc7811789f002165a63">
<HY_Geometry><LineString xmlns="http://www.opengis.net/gml"><posList>49.0646908 -97.3497687 49.0644278 -97.3494799 49.0642853 -97.3494171 49.0640799 -97.3497206 49.0639081 -97.349752 49.0637763 -97.3496213 49.0636744 -97.3493691 49.0636697 -97.348821 49.0641574 -97.3483827 49.0645459 -97.3483223 49.064945 -97.3478243 49.0650788 -97.3478729 49.0652696 -97.3481713 49.0656438 -97.3490689 49.0659305 -97.3498658 49.0661046 -97.3501084 49.0662597 -97.3503911</posList></LineString></HY_Geometry>
<FromNode>66f3bd4d3cac401ba6525a964dd64430</FromNode>
<ToNode>3df46f2c73be4d9e8b2287ed5c0ca2c4</ToNode>
</HY_FlowPath>
```

Figure 34 — Use case 5 demo: Merged features in HY_Features model

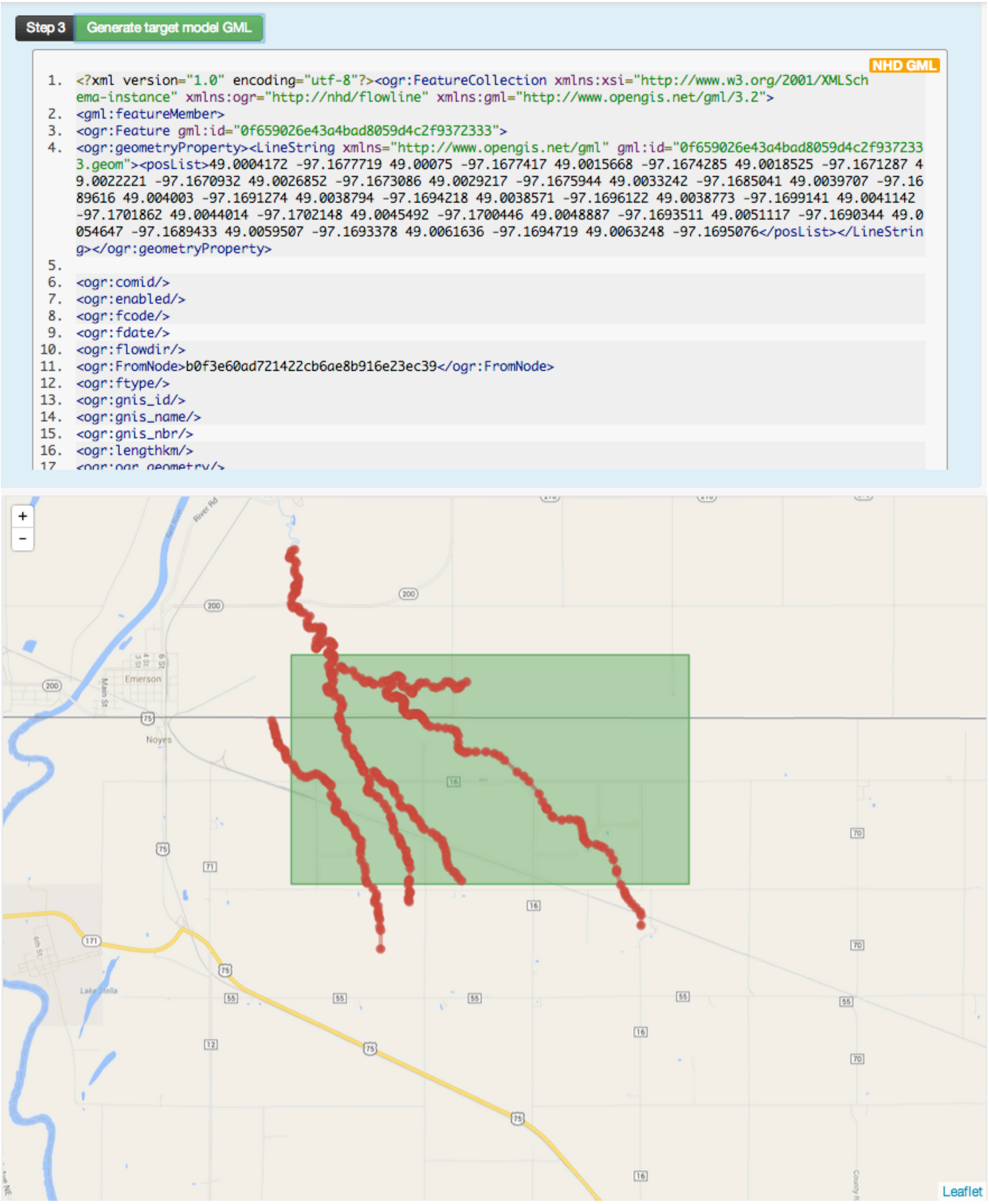

Figure 35 — Use case 5 demo: Merged features in NHD+ model

Internally, the semantic SPARQL engine was used to query the possible models to connect the equivalent nodes. Table 7 lists one example query to look for cross model data euqivalent. Figure 36 shows the result returned from the SPARQL engine described in 6.3.5. In this example, according to the result, we can use the value of to_junct to search the data in the field [mapping_node_NHN] of the entity [mapping_node], and get the value from the field [mapping_node_NHD] which contains the value to be converted to in NHD model.

Table 7. An example query to match cross-model data

```
SELECT distinct ?entity ?sourceField ?targetField
WHERE {
   ?cross mapf:crossField nlflow:to_junct.
   {
     {?cross mapf:dataEquivalent ?t}
     union
     {?cross ^mapf:dataEquivalent ?t}
   }
   ?cross mapf:dataField ?sourceField.
   ?t mapf:dataField ?targetField.
   ?sourceField mapf:belongEntity ?entity
}
```

```
-------------------------------------------------------------------
| entity            | sourceField         | targetField         |
===================================================================
| nd:mapping_node   | nd:mapping_node_NHN | nd:mapping_node_NHD |
-------------------------------------------------------------------
```

Figure 36. Result of the query in Table 7

7 Best Practices

7.1 Access hydrologic from multiple data sources

Hydrologic features are subject of common discourse. Depending on application and scale they are multiply represented in a wide range of typical datasets and data products. The most common form is the geometric representation in GIS layers of polygons, polylines and points. Scaling up or down, as well as the aggregation in alternative hierarchies will lead to multiple representations of the same hydrologic feature. Displaying data from different sources, as requested in trans-boundary and cross-domain applications, requires compatible formats. Data integration also requires the compatibility of the concepts which the displayed features represent.

The *HY_Features* model provides commonly agreed concepts of the hydrologic features and their fundamental relationships. Providing a re-usable core of common concepts, it allows the reference of specific features realized in existing implementations to the general concept they represent. This enables the access to different data sources describing a shared hydrologic feature by mapping to the equivalent concept in the *HY_Features* model.

In the Testbed-10 Hydro thread, the *HY_Features* model is used to establish the link between the National Hydrographic Network (NHN) of Canada and the National Hydrographic Dataset Plus (NHD+) of the United States. Particularly, the general *HY_Flowpath* concept (of the *HY_Features* model) is referenced by the NHN and the NHD+ models, to merge the linear features from both sources and display exemplarily a combined *flowline* representing the Red River crossing the USA and Canada border. In addition, the data obtained from one national database are provided in a format which is accessible for the other, e.g. NHN data accessible for a NHD+ client. Provided in the GML format, the data stored in the national databases may become available for any other client or service such as WFS or WMS.

Since *HY_Flowpath* carries a relationship to the represented catchment (*representedCatchment*), each particular flowline can be linked with other representations of this catchment, e.g. a watershed polygon or a network of monitoring points, provided that these representations are mapped to the *HY_Features* equivalents. A prototype type knowledgebase in OWL was implemented and tested with such equivalent semantics. The knowledgebase was released for testing through a standard SPARQL engine. As demonstrated for the linear features stored in the NHN and the NHD+ datasets, other applications may map their intended hydrologic features to the *HY_Features* equivalent concepts to provide a common basis for a shared referencing of hydrologic features in data infrastructures as well as to make information systems interoperable “on the fly”.

7.2 Mediation between hydrologic models

The shared referencing of hydrologic features in data infrastructures may be challenging if different semantics and identifiers are used in the source applications for representation. Provided that common concepts are accessible, a mediation service may use these as the target reference to invoke a particular information or data service in order to provide the source information to a requesting service. Being generally applicable and independent from scales, the *HY_Features* model provides commonly agreed target references which will enable an intended Hydro Mediation Service to invoke a national information service whose implemented feature concepts are mapped to its equivalents in the *HY_Features* model.

A semantic mapping of application-specific features to their common equivalents provided, the *Hydro Model Service* will be able to manage different representations of an identified hydrologic feature and “negotiate” which information system or service should be invoked to get the most appropriate response to a particular request.

For example, a WFS will send a request for the linear representation of a particular basin to the intended *HY_Features* mediation service. Invoking a range of import services using a target reference *HY_Catchment*, which is referenced by a national "Hydrographic Unit", the mediation service will get a response from those services, whose "Flowline" concept is mapped to the common *HY_Flowpath* concept, which is one of multiple representations associated with the represented catchment. Applying a Linked Data approach, where *HY_Catchment – HY_Flowpath* – "Flowline" may form the *Subject-Predicate-Object Triple*, the mediation service will be able to handle different resources to find all "Flowlines" that represent the "Hydrographic Unit" of interest.

The *Hydro Model Service* developed within the Testbed-10 may be further developed towards the interface/s required to invoke the import systems as well as to enable the requesting service to access the information or data in the "expected" format. Implemented as a WPS, the *Hydro Model Service* is capable of including the required *mapping* of different source models to the common *HY_Features* model, either as a process on its own "executed" with *input*s each set by the target and the sources, and a pre-defined output required as input for the next process, or as a preset *input* required by a particular process. To this, the *Hydro Model Service* may be further developed to offer a range of preset (registered) mappings in order to discover the most appropriate resource, and furthermore to set mappings between formalized concepts, i.e. between application schemas built according to ISO 19109, and the *HY_Features* model.

ISO/TS 19150-1:2012 defines the framework for semantic interoperability of geographic information. This framework defines a high level model of the components required to handle semantics in the ISO geographic information standards with the use of ontologies. The rules provided therein for converting UML schemas to OWL, needs to be examined for their applicability for a formalized model mapping onto the HY_Features in terms of an upper level ontology.

7.3 Mediation power of the HY_Features model

The HY_Features common model was used as the conceptual bridge between NHD+ and NHN models. The demonstration prototype shows that HY_Features can achieve the integration of both datasets without information loss. The detailed example on flow path demonstrated that HY_Features provides corresponding components at the levels of model, entity, and field to NHN hydrological model and NHD+ hydrological model.

The proposed mapping framework and the specific "field to field mapping" of NHN and NHD+ to the general HY_Flowpath concept demonstrated the mapping of features of the same geometry, exemplarily flow lines, by referencing a common concept, here HY_Flowpath in HY_Features. The applicability of this approach to map features of other geometry, especially of different geometries using the pre-defined relationships of HY_Features is left to future test beds.

Matching multiple and disparate representations of hydrologic features may call for a semantic mapping technology beyond the simple equivalence of meaning, e.g. contextual relationships. The HY_Features model supports this by definition of fundamental relationships of features caused and determined by hydrologic processes.

7.4 Best practices for a hydro model mediation system

Conversion of models, entities, and fields from heterogeneous sources involves many aspects of data integration. When multiple models (more than two models) are involved, the mapping between models can take different approaches. Typically, there are three types of approaches:

1. The direct one-to-one mapping, where each element of the source model is transformed to the target model directly, usually case by case. The obvious advantages are its capability to convert form sources to target model with minimal information loss given each one-to-one mapping is well constructed. The major drawback is its requirements of extensive knowledge to all the models and its laborious work load in creating individual mappings.

2. Use a well-known top common model to represent the common features from all data sources. This approach requires each source model to be mapped to the top model. This approach has two limitations. For one, the detail of mapping from source to target depends on the details of top common model, i.e. the more comprehensive the top common model, the more detail the mapping can achieve. Secondly, this approach requires knowledge about the top model, what is often inconvenient for users who usually have a strong preference in working on their own familiar model.

3. Using a bridging model for mediation between different models. The bridge common model, ideally, is derived by finding the most detailed, common reference model for all models under study. The relevant parts of each model are mapped to the bridge common model. This mapping is used to link the corresponding elements of the source and the target model. The advantage is the easier implementation and less mappings to be designed. The identification of the common reference model among different source models can be done semantically. When the knowledgebase is properly designed, the common model can be identified among all source models under study almost automatically.

In the test bed, the third approach is adopted, but, taking for granted, the HY_Features common hydrologic feature model was directly used as the bridge common model to achieve mappings between different hydrological models. The HY_Features model was intentionally developed as the common reference model for all hydrologic features using agreed international standards for terminologies and definitions. From this demonstration, it was observed that HY_Features is able to support the conversion of hydrological features between NHN and NHD+ without information loss.

Furthermore, it was confirmed that the modular structure of HY_Features supports the selected mapping of particular model elements of interest. Thus, there was no need to build a complete HY_Features profile for each source model. With respect to a shared referencing of features, whenever identified to be hydrologic ones, the persistent storage of reference concepts would be helpful as well as an offering of pre-defined mappings when deploying an (application-specific) mediation service.

Bibliography

[1] ISO/TC 211, ISO/DIS 19150-2 - Geographic information -- Ontology -- Part 2: Rules for developing ontologies in the Web Ontology Language (OWL). ISO, 2014.
[2] I. Dornblut and R. Atkinson, "OGC HY_Features: a Common Hydrologic Feature Model," Open Geospatial Consortium Inc., Wayland, MA, USA, OGC® Implementation Specification – Application Profile OGC 11-039r3, Feb. 2014.
[3] ogc, "waterml schema," 2014. [Online]. Available: http://schemas.opengis.net/waterml/. [Accessed: 01-Apr-2014].
[4] P. Taylor, "OGC® WaterML 2.0: Part 1- Timeseries," Open Geospatial Consortium Inc., Wayland, MA, USA, OGC® Implementation Standard - Corrigendum OGC 10-126r4, Feb. 2014.
[5] S. R. Jackson, D. K. Arctur, and D. R. Maidment, "RiverML: A River Geometry and Observation Exchange Language," presented at the OGC Hydro DWG Workshop Quebec City 2013, 2013.
[6] CRWR, "CRWR Tools RiverML," 2014. [Online]. Available: http://tools.crwr.utexas.edu/RiverML/. [Accessed: 01-Apr-2014].
[7] L. McKay, T. Bondelid, T. Dewald, A. Rea, C. Johnston, and R. Moore, "NHDPlus Version 2: User Guide (Data Model Version 2.1)," United States Environmental Protection Agency, user's manual EPA Contract CM130105CT0027, Feb. 2014.
[8] GeoBase, National Hydro Network, Canada, Level 1, Data Model, Edition 1.0, 1st ed. Sherbrooke, QC Canada: Natural Resources Canada Geomatics Canada, Centre for Topographic Information, 2004.
[9] GeoBase, National Hydro Network, Canada, Level 1, Product Specifications, Edition 1.0, 1st ed. Sherbrooke, QC Canada: Natural Resources Canada Geomatics Canada, Centre for Topographic Information, 2004.
[10] GeoBase, National Hydro Network, Canada, Level 1, Linear Referencing System (LRS), Data Catalogue, Edition 1.0, 1st ed. Sherbrooke, QC Canada: Natural Resources Canada Geomatics Canada, Centre for Topographic Information, 2004.
[11] GeoBase, National Hydro Network, Data Product Specifications, Distribution Profile, Edition 1.1, 1st ed. Sherbrooke, QC Canada: Natural Resources Canada Geomatics Canada, Centre for Topographic Information, 2010.
[12] GeoBase, "National Hydro Network, Feature Catalogue Distribution Profile, Edition 1.0.1," 01-Jun-2008. [Online]. Available: http://www.geobase.ca/doc/catalogue/GeoBase_NHN_Catalogue_1.0.1_EN.html. [Accessed: 03-May-2014].

7.5 Revision history

Date	Release	Editor	Primary clauses modified	Description
2014-02-18	0.0	Eugene Yu		Initial outline
2014-03-07	0.1	Eugene Yu		Filled with content and contributions from Chenyu How and Irina Domblut
2014-03-11	0.2	Eugene Yu		Added descriptions for five use cases and their variations with contributions from Chenyu How and Irina Domblut
2014-03-18	0.3	Eugene Yu		Added descriptions for the semantic mapping framework with contributions from Chenyu How and Irina Domblut
2014-03-31	0.4	Eugene Yu		Drop controlled vocabulary section. Add references and descriptions.
2014-04-07	0.5	Eugene Yu		Add example mapping diagrams of NHD+ to HY_Features and NHN to HY_Features. Add description of cross-model-framework to deal with feature inconsistence of different models.
2014-04-21	0.6	Eugene Yu		Completed version, except client description. Newly added: semantic mapping examples/SPARQL queries; future work; mediation power of HY_Features model; mediation model and recommendations.
2014-05-02	0.7	Eugene Yu		Added document number. Made several minor revision.
2014-05-07	0.8	Eugene Yu		Added one more future work as suggested by Irina Dornblut and Rob Atkinson. Incorporated all the changes suggested by Irina Dornblut.
2014-05-14	0.9	Eugene Yu		Incorporated all the suggested changes suggested by Rob Atkinson.

