Evaluation Tests and Results

Andriana Gkaniatsou

July 11, 2013

1 Introduction

In the following sections we describe some evaluation tests we conducted and we present the results we obtained. For each evaluation test we assume that we are the owners of a specific dataset and that we receive queries that we need to provide answers. Our goal is to show that our system is capable to provide the exact or associated answers for information provided by the query.

Each query corresponds to the schema of a real dataset [1]. As a first step each query represents an exact dataset schema (including the dataset name). Then, depending the results we twist the predicate name.

1.1 SPSM Limitations/Bugs

As we will show in the following sections there are some SPSM bugs that affect the capability of our system to produce results. If SPSM does not match two predicate names, it does not produce matches for the corresponding arguments. Thus, for testing purposes in some experiments we twist the predicate name.

2 SEPA datasets

We assume that we are the owners of the SEPA datasets, and that any of the incoming queries is formed by a different agency. Each query represents the dataset schema of an agency. The SEPA datasets that we used are the following:

- bathingWaterLocations(year, locationCode, localAuthority, locationDescription, sepaTeam, sepaRegion, hydrocode, lat, ngr, catchment, long)
- waterBodyPressures(envImprovementNext, atGepNextPeriod, atGepCurClassYear, atGepSysdate, pressureId, waterBodyId, locationCode, industrySectorCode, assessmentCategory, assessmentParameter, swmiSector, industrySector, dataSource, comments, atGepPeriodAfterNext, isPrimary, atGepCurrentPeriod, pressureType, envImprovementAfterNext, envImprovementCurrent, purposeCode, purpose, eiCurrentFailReason, identifiedDate, eiCurrentConfidence, eiNextConfidence, hmawbDesignationInd, eiAfterConfidence, affectsGroundwater, eiNextFailReason, activityCode, activity, swAsset, protectedAreaId, eiAfterFailReason, sourceCode, source)

[1] These datasets are publicly available from different agencies. Some of them are available only for only browsing, while other for downloading as well.

- waterBodyMeasures(atGepPeriodAfterNext, atGepCurClassYear, waterBodyId, pressureId, deadline, comments, secondaryMeasure, primaryMeasure, owner, atGepSysdate, atGepNextPeriod, measureFixedDate, agreed, projected, atGepCurrentPeriod, dataSource, measureId, locationCode, protectedAreaId, onTrack, notOnTrackReason)
- surfaceWaterBodies(genre, lengthKm,a ssociatedGroundwaterId, sbdCode, dataSource, isHeavilyModified, geologyTypology, currentClassificationYear, associatedGroundwater, statusRiskAssessment, riverBasinDistrict, category, rbdCode, isArtificial, riverName, long, altitudeTypology, catchmentId, riverNumber, classificationCertainty, waterBodyId, currentOverallClassification, classificationCertaintyBand, waterBodyName, subBasinDistrict, ecoRegion, isLessThanGood, wiseCode, catchment, targetClass, lat, team, sizeTypology, noDetRiskAssessment, areaSqKm)
- bathingWaters(year, description, classDescription, bwLocation, bathingWaterId)
- classifications(classificationYear, parameter) and parameter is (certainty, certaintyBand, classificationParameter, contributingLocationCode, lessThanGood, reportingParameter, reportingParameterCode, status, statusDescription)

2.1 Test 1: European Statistics, Fresh Water Resources dataset

Incoming query is formed based on the freshWaterResources dataset:
freshWaterResources(resource, measure, geo, timePeriod)

2.1.1 Test 1.1

Plural: on The query is the dataset schema:
`freshWaterResources( resource, measure, geo, timePeriod)`

Results: The narrow-down process returns as possible matching candidates the following datasets:

- waterBodyPressures
- waterBodyMeasures
- surfaceWaterBodies
- bathingWaterLocations

However, SPSMS fails to match any of these dataset names with the one provided by the query.

Plural: off The narrow-down process returns an extra matching candidate:

- bathingWater

However SPSM did not return any results.

Timings:

- real 1m52.214s
- user 0m44.714s
- sys 0m7.701s

2.1.2 Test 1.2

We change the predicate name of the query into `freshWater`:

```
freshWater( resource, measure, geo, timePeriod).
```

Results: The narrow-down process returned the same candidates as before. However, SPSMS did not returned any matching results.

2.1.3 Test 1.3

We change the predicate name of the query with into `water`:

```
water( resource, measure, geo, timePeriod)
```

Plural: On, Results: The narrow-down process returned the following possible candidates:

- waterBodyPressures
- waterBodyMeasures
- surfaceWaterBodies
- bathingWaterLocations

The results we obtained from SPSM are the following:

Match 1.

water $\mapsto$ waterBodyPressures
water(timePeriod) $\mapsto$ waterBodyPressures(identifiedDate)
water(geo) $\mapsto$ waterBodyPressures(waterBodyId)
water(measure) $\mapsto$ waterBodyPressures(assessmentCategory)
water(resource) $\mapsto$ waterBodyPressures(source)

Match 2.

water$\mapsto$ bathingWaters [2]

We applied the substitutions provided by SPSM and we deleted the arguments that do not appear to have any mappings. A sample of the answers we retrieved is presented in A.1)

[2]We discard this match as there are no argument mappings

Plural: Off, Results: The narrow-down process retuned one more matching candidate:

- bathingWater

SPSM retuned the same results as above.

Timings:

- real 2m10.208s
- user 0m49.725s
- sys 0m8.023s

2.2 Test 2: European Statistics, Environmental Protection Expenditure dataset

Incoming query is formed based on the Environmental Protection Expenditure dataset:

environmentProtectionRegion(dimension, unit, environmentDomain, environmentExposure, geo, timePeriod).

2.2.1 Test 2.1

The query represents is the same with the agency's dataset schema:

`environmentalProtectionExpenditutureRegions( dimension, unit, environmentDomain, environmentExposure, geo, timePeriod )`

Plural: On, Results: The narrow-down process did not returned any possible candidates.

Plural: Off, Results: The narrow-down process retuned bathingWaterLocation as a matching candidate. SPSM does not return any results.

Timings:

- real 0m33.804s
- user 0m15.963s
- sys 0m1.040s

2.2.2 Test 2.2

We change the predicate of the query name into `environmentProtectionRegion`:

`environmentProtectionRegion( dimension, unit, environmentDomain, environmentExposure, geo, timePeriod )`.

Plural: On, Results: The narrow-down process did not returned any possible candidates.

Plural: Off, Results: The narrow-down process returns bathingWaterLocation as a matching candidate. SPSM does not return any results.

Timings:

- real 0m27.942s
- user 0m16.596s
- sys 0m1.044s

2.2.3 Test 2.2

We change the predicate name of the query with *locations*[3]:

```
locations( dimension, unit, environmentDomain, environmentExposure, geo, timePeriod ).
```

Plural: On, Results: The narrow down process returned the following matching candidates:

- bathingWaterLocations

However, SPSM did not return any results [4].

Plural: Off, Results: The narrow-down process returns bathingWaterLocation as a matching candidate. This time, SPSM also returns results:

Match 1.
location $\mapsto$ bathingWaterLocation
location(unit) $\mapsto$ bathingWaterLocation(localAuthority)

A sample of the returned answers is presented in A.2.

Timings:

- real 1m6.662s
- user 0m19.747s
- sys 0m1.282s

2.2.4 Test 2.3

We replace the predicate name with *water*:

```
water( dimension, unit, environmentDomain, environmentExposure, geo, timePeriod ).
```

[3]The narro-down process uses Wordnet and SUMO. None of these two ontologies contain arguments in plural. So, the narrow-down process will not match area with regions, but will do in case of area and region.

[4]This is due to a bug in SPSM algorithm

Plural: On, Results: The narrow down process returned the following matching candidates:

- waterBodyPressures
- waterBodyMeasures
- surfaceWaterBodies
- bathingWaterLocations

The SPSM results are the following:

Match 1.

water ↦ waterBodyPressures
water(geo) ↦ waterBodyPressures(waterBodyId)
water(environmentExposure) ↦ waterBodyPressures(activity)
water(environmentDomain) ↦ waterBodyPressures(affectsGroundwater)
water(unit)↦ waterBodyPressures(dataSource)
water(dimension) ↦ waterBodyPressures(industrySector)

Match 2.

water ↦ bathingWaters [5].

Match 3.

water ↦ waterBodyMeasures water(geo) ↦ waterBodyMeasures(waterBodyId)
water(unit) ↦ waterBodyMeasures(secondaryMeasure)

We applied the substitutions provided by SPSM and we deleted the arguments that do not appear to have any mappings. A sample of the results we obtained s presented in A.3.

Plural: Off, Results: The narrow-down process returns one more matching candidate: bathingWater. SPSM returned the same results as before.

Timings:

- real 2m47.152s
- user 1m0.255s
- sys 0m9.011s

2.3 Test 3: Scottish Protection Agency: Protected Areas dataset

Incoming query is formed based on the Scottish Protection Agency, Protected Areas dataset: spa_protectedAreas(area, perimeter, paCode, siteName, siteHa, status, eurCode)

2.3.1 Test 3.1

The query is the same as the dataset schema:

```
spa_protectedAreas( area, perimeter, paCode, siteName, siteHa, status,
eurCode )
```

[5] We discarded this match since there are no matching arguments.

Plural: On, Results: The narrowing-down process did not return any matching candidates.

Plural: Off, Results: The narrow-down process did not return any results.
Timings:

- real 0m28.893s
- user 0m13.259s
- sys 0m3.721s

2.3.2 Test 3.2

We replace the predicate name of the query with *areas*:
`protectedAreas( area, perimeter, paCode, siteName, siteHa, status, eurCode )`

Plural: On, Results: The narrowing-down process did not return any results [6].

Plural: Off, Results: The narrow-down process did not return any results.
Timings: Same as before.

2.3.3 Test 3.3

We replace the query predicate name with the generic *water*:
`water( area, perimeter, paCode, siteName, siteHa, status, eurCode )`

Plural: On, Results: The narrow-down process returned the following results:

- waterBodyPressures
- waterBodyMeasures
- surfaceWaterBodies
- bathingWaterLocations
- waterBodyPressures
- waterBodyMeasures
- surfaceWaterBodies
- bathingWaterLocations

[6] Although there exist a dataset with similar name: bathingWaterLocations, due to the plural of location, the narrowing-down process cannot find any wordnet or sumo related word. If we replace locations with location, then, we get results.

SPSM returned the following mappings:

Match 1.

water ↦ bathingWaters

Match 2.

water ↦ waterBodyPressures water(eurCode) ↦ waterBodyPressures(locationCode)
water(status) ↦ waterBodyPressures(pressureType)
water(paCode) ↦ waterBodyPressures(purposeCode)
water(perimeter) ↦ waterBodyPressures(affectsGroundwater)
water(area) ↦ waterBodyPressures(waterBodyId)

Match 3. water ↦ waterBodyMeasures])
water(eurCode) ↦ waterBodyMeasures(locationCode)
water(area) ↦ waterBodyMeasures(waterBodyId)

A sample of the answers we obtained is presented in A.4.

Plural: Off, Results The narrow-down process returned one additional matching candidate: bathingWater. SPSM returned the same results as before.

Timings:

- real 2m57.830s
- user 1m7.376s
- sys 0m10.141s

2.3.4 Test 3.4

We replace the query predicate name with the generic *locations*:

```
locations( area, perimeter, paCode, siteName, siteHa, status, eurCode
)
```

Plural: On, Results The narrow-process returned bathingWaterLocations as a matching candidate. SPSM returned the following matches:

Match 1.

location ↦ bathingWaterLocation
location(eurCode) ↦ bathingWaterLocation(locationCode).

A sample of the answers we obtained is presented in A.5

2.4 Test 4: The UK Environmental Agency: Bathing Waters dataset

The Uk Environmental Agency offers an online API to query and retrieve information from the bathing waters dataset. The schema of this dataset is the following:

uk_BathingWaters(sampleClassification, prefLabel, long, lat, northing, easting, latestComplianceAssessment, type, country, district, envelope, latestBathingWaterProfile, sedimentTypesPresent, uriSet, regionalOrganization, yearDesignated, latestSampleAssessment, eubWidNotation, waterQualityImpactedByHeavyRain, zoneOfInfluence, samplingPoint, complianceClassification, primaryTopic, extendedMetadataVersion, definition, label).

2.4.1 Test 4.1

The query we form is the same as the bathing waters dataset schema:
`uk_BathingWaters(sampleClassification, prefLabel, long, lat, northing, easting, latestComplianceAssessment, type, country, district, envelope, latestBathingWaterProfile, sedimentTypesPresent, uriSet, regionalOrganization, yearDesignated, latestSampleAssessment, eubWidNotation, waterQualityImpactedByHeavyRain, zoneOfInfluence, samplingPoint, complianceClassification, primaryTopic, extendedMetadataVersion, definition, label)`

Plural: On, Results: The narrow-down process returned the following matching candidates:

- bathingWaters
- bathingWaterLocations
- waterBodyMeasures
- waterBodyPressures

SPSM retuned the following matches:
Match 1.
uk_BathingWaters ↦ waterBodyMeasures
uk_BathingWaters(label)↦ waterBodyMeasures(waterBodyId)
uk_BathingWaters(primaryTopic)↦ waterBodyMeasures(secondaryMeasure).

Match 2.
uk_BathingWaters ↦ waterBodyPressures
uk_BathingWaters(label)↦ waterBodyPressures(waterBodyId)
uk_BathingWaters(type) ↦ waterBodyPressures(pressureType)
uk_BathingWaters(prefLabel) ↦ waterBodyPressures(activity)
uk_BathingWaters(sampleClassification) ↦ waterBodyPressures(affectsGroundwater)
Match 3.
uk_BathingWaters ↦ bathingWaters.
uk_BathingWaters(label) ↦ bathingWaters(description)

Please see A.6 for some of the answers we got.

Plural: On, Results: Both narrow-down process and SPSM retuned the same results.
Timings:

- real 3m55.407s
- user 1m25.673s
- sys 0m18.614s

3 European Statistics Statistics, Protected Areas dataset

During the following evaluation tests we assume that we are the owners of the EU Statistics ProtectedAreas dataset. The schema of this dataset is the following:

protectedAreas(obsValue, timePeriod, freq, dataSet, indic_md, geo, obsStatus, structure, seeAlso, comment)

3.1 Test 1: Scottish Protection Agency

Incoming query is formed based on the Scottish Protection Agency, Protected Areas dataset:

```
spa_protectedAreas( area, perimeter, paCode, siteName, siteHa, status, eurCode )
```

Plural: On, Results Narrowing - down process suggested the protectedAreas dataset as a possible candidate. The results we obtained from SPSM are the following:

Match 1:

spa_protectedAreas $\mapsto$ protectedAreas
spa_protectedAreas(status) $\mapsto$ protectedAreas(obsStatus)
spa_protectedAreas(area) $\mapsto$ protectedAreas(structure)

A sample of the answers this query returned is presented in B.1

Plural: Off, Results We get the same results as before.

Timings:

- real 1m20.555s
- user 0m16.934s
- sys 0m1.090s

3.2 Test 2: European Statistics, Environmental Protection Expenditure dataset

Incoming query is formed based on the Environmental Protection Expenditure dataset: environmentProtectionRegions(dimension, unit, environmentDomain, environmentExposure, geo, timePeriod)

3.2.1 Test 2.1

The query is the same as the dataset schema:

```
environmentProtectionRegions( dimension, unit, environmentDomain, environmentExposure, geo, timePeriod )
```

Plural: On, Results The narrow-down process retuned protectedAreas dataset as a matching candidate. However SPSM did not return any results.

Plural: Off, Results The results are the same as before.
Timings:

- real 0m32.891s
- user 0m9.840s
- sys 0m0.155s

3.2.2 Test 2.2

We change the predicate name of the query into: ProtectedRegions. The query we test is the following:

```
ProtectedRegions( dimension, unit, environmentDomain, environmentExposure,
geo, timePeriod )
```

Plural: On, Results The narrow down process returned protectedAreas as a matching candidate. SPSM produced the following results:
Match:
protectedRegions $\mapsto$ protectedAreas
protectedRegions(geo) $\mapsto$ protectedAreas(geo)
protectedRegions(environmentDomain) $\mapsto$ protectedAreas(timePeriod)
A sample of the results we obtained is demonstrated in B.2.

Plural: Off, Results Timings:

- real 0m36.244s
- user 0m20.941s
- sys 0m1.183s

3.3 Test 3: UK Environmental Agency: Bathing Waters Dataset

Incomming query is formed based on the UK Environmental Agency's Bathing Waters dataset. The schema of that dataset is the following:
uk_BathingWaters(sampleClassification, prefLabel, long, lat, northing, easting, latestComplianceAssessment, type, country, district, envelope, latestBathingWaterProfile, sedimentTypesPresent, uriSet, regionalOrganization, yearDesignated, latestSampleAssessment, eubWidNotation, waterQualityImpactedByHeavyRain, zoneOfInfluence, samplingPoint, complianceClassification, primaryTopic, extendedMetadataVersion, definition, label)

3.3.1 Test 3.1

The query we form is the same as the bathing waters dataset schema:

```
uk_BathingWaters( sampleClassification, prefLabel, long, lat, northing,
easting, latestComplianceAssessment, type, country, district, envelope,
latestBathingWaterProfile, sedimentTypesPresent, uriSet, regionalOrganization,
yearDesignated, latestSampleAssessment, eubWidNotation, waterQualityImpactedByHeavyRain,
zoneOfInfluence, samplingPoint, complianceClassification, primaryTopic,
extendedMetadataVersion, definition, label)
```

Plural: On, Results: The narrow-down process did not return any results.

Plural: Off, Results: The results are the same as above.
Timings:

- real 0m16.666s
- user 0m9.912s
- sys 0m0.145s

3.3.2 Test 3.2

We change the predicate name of the query from uk_bathingWaters to uk_protectedBathingWaters. The query that we send is the following:

```
uk_protectedBathingWaters( sampleClassification, prefLabel, long,
lat, northing, easting, latestComplianceAssessment, type, country,
district, envelope, latestBathingWaterProfile, sedimentTypesPresent,
uriSet, regionalOrganization, yearDesignated, latestSampleAssessment,
eubWidNotation, waterQualityImpactedByHeavyRain, zoneOfInfluence, samplingPoint,
complianceClassification, primaryTopic, extendedMetadataVersion, definition,
label)
```

Plural: On, Results The narrow down process returned protectedAreas as possible candidate, but SPSM did not produced any matching results.

Plural: Off, Results The results and timings are the same as above.

3.3.3 Test 3.3

We change the predicate name of the query from uk_bathingWaters to uk_protectedAreas. The query that we send is the following:

```
uk_protectedAreas( sampleClassification, prefLabel, long, lat, northing,
easting, latestComplianceAssessment, type, country, district, envelope,
latestBathingWaterProfile, sedimentTypesPresent, uriSet, regionalOrganization,
yearDesignated, latestSampleAssessment, eubWidNotation, waterQualityImpactedByHeavyRain,
zoneOfInfluence, samplingPoint, complianceClassification, primaryTopic,
extendedMetadataVersion, definition, label)
```

Plural: On, Results: The narrow down process returned protectedAreas as matching candidates. SPSM produced the following matching results:
Match 1.
uk_ProtectedAreas $\mapsto$ protectedAreas
uk_ProtectedAreas(yearDesignated) $\mapsto$ protectedAreas(timePeriod)

A sample of the answers this query returned is demonstrated in B.3.

Plural: Off, Results: Both the narrow-down process and SPSM produced the same results as before.

4 European Statistics, Protected Areas and Environmental Protection Expenditure dataset

During the following evaluation tests we assume that we are the owners of both EU Statistics protectedAreas and environmentProtectionRegions datasets. The schemamata of these datasets are the following:

protectedAreas(obsValue, timePeriod, freq, dataSet, indic_md, geo, obsStatus, structure, seeAlso, comment)

environmentProtectionRegions(dimension, unit, environmentDomain, environmentExposure, geo, timePeriod)

4.1 Test 1: Scottish Protection Agency: Protected Areas dataset

The query that we form is the schema of the Scottish Protection Agency, Protected Areas dataset:

`spa_protectedAreas( area, perimeter, paCode, siteName, siteHa, status, eurCode )`

Plural: On, Results: The narrow down process returned both environmentProtectionRegions and protectedAreas as matching candidates. SPSM produced the following matching results:

Match 1:
spa_protectedAreas $\mapsto$ protectedAreas
spa_protectedAreas(status) $\mapsto$ protectedAreas(obsStatus)
spa_protectedAreas(area) $\mapsto$ protectedAreas(structure)

A sample of the answers this query returned is demonstrated in B.1

Plural: Off, Results: Both narrow-down process and SPSM produced the same results as before.

Timings:

- real 0m46.684s
- user 0m35.146s
- sys 0m2.154s

5 European Statistics, Transport Network dataset

For the following tests we use the Transport Network dataset. The schema of this dataset is the following:

`transportNetwork(obsValue, timePeriod, freq, dataset, unit, geo, obsStatus, structure, seeAlso, comment)`

5.1 Test 1: NAPTAN, Stop Availability dataset

The icomming query is formed based on the NAPTAN Stop Availability schema:

stopAvailability(atcoCode, startDate, endDate, availabilityStatus, note, note-

Lang, transferStopAtcoCode, creationDateTime, modificationDateTime, revisionNumber, modification)

5.1.1 Test 1

The query that we send is exactly the same as the stopAvailability dataset schema:

```
stopAvailability(atcoCode, startDate, endDate, availabilityStatus,
note, noteLang, transferStopAtcoCode, creationDateTime, modificationDateTime,
revisionNumber, modification)
```

Plural:On, Results: The narrow down process did not return any results.

Plural:Off, Results: The narrow-down process and SPSM produced the same results as before.

Timings:

- real 0m17.179s
- user 0m12.492s
- sys 0m0.170s

5.1.2 Test 1.1

We replace the predicate name of the query with transportStopAvailability. The query that we send is the following:

```
transportStopAvailability(atcoCode, startDate, endDate, availabilityStatus,
note, noteLang, transferStopAtcoCode, creationDateTime, modificationDateTime,
revisionNumber, modification)
```

Plural:On, Results: The narrow-down process retuned the transportNetwork dataset as a possible matching candidate. However, SPSM did not return any results.

Plural:Off, Results: The narrow-down process and SPSM produced the same results as before.

Timings:

- real 0m30.403s
- user 0m26.874s
- sys 0m1.808s

5.1.3 Test 1.2

We substitute the predicate name of the query with transport:

```
transport(atcoCode, startDate, endDate, availabilityStatus, note,
noteLang, transferStopAtcoCode, creationDateTime, modificationDateTime,
revisionNumber, modification)
```

Plural: On, Results: The narrow-down process returned transportNetwork as a matching candidate. SPSM produced the following results:

Match 1.

transport $\mapsto$ transportNetwork
transport(note) $\mapsto$ transportNetwork(comment)

A sample of the answers we retrieved is presented in C

Plural: Off, Results: The narrow-down process and SPSM produced the same results as before.

Timings:

- real 0m45.144s
- user 0m34.398s
- sys 0m2.173s

6 All datasets

For this experiment we assume that we are the owners of all datasets that we have previously tested. The datasets and the corresponding schemata are:

- bathingWaterLocations(year, locationCode, localAuthority, locationDescription, sepaTeam, sepaRegion, hydrocode, lat, ngr, catchment, long)
- waterBodyPressures(envImprovementNext, atGepNextPeriod, atGepCurClassYear, atGepSysdate, pressureId, waterBodyId, locationCode, industrySectorCode, assessmentCategory, assessmentParameter, swmiSector, industrySector, dataSource, comments, atGepPeriodAfterNext, isPrimary, atGepCurrentPeriod, pressureType, envImprovementAfterNext, envImprovementCurrent, purposeCode, purpose, eiCurrentFailReason, identifiedDate, eiCurrentConfidence, eiNextConfidence, hmawbDesignationInd, eiAfterConfidence, affectsGroundwater, eiNextFailReason, activityCode, activity, swAsset, protectedAreaId, eiAfterFailReason, sourceCode, source)
- waterBodyMeasures(atGepPeriodAfterNext, atGepCurClassYear, waterBodyId, pressureId, deadline, comments, secondaryMeasure, primaryMeasure, owner, atGepSysdate, atGepNextPeriod, measureFixedDate, agreed, projected, atGepCurrentPeriod, dataSource, measureId, locationCode, protectedAreaId, onTrack, notOnTrackReason)
- surfaceWaterBodies(genre, lengthKm,a ssociatedGroundwaterId, sbdCode, dataSource, isHeavilyModified, geologyTypology, currentClassificationYear, associatedGroundwater, statusRiskAssessment, riverBasinDistrict, category, rbdCode, isArtificial, riverName, long, altitudeTypology, catchmentId, riverNumber, classificationCertainty, waterBodyId, currentOverallClassification, classificationCertaintyBand, waterBodyName, subBasinDistrict, ecoRegion, isLessThanGood, wiseCode, catchment, targetClass, lat, team, sizeTypology, noDetRiskAssessment, areaSqKm)
- bathingWaters(year, description, classDescription, bwLocation, bathingWaterId)

- classifications(classificationYear, parameter) and parameter is (certainty, certaintyBand, classificationParameter, contributingLocationCode, lessThanGood, reportingParameter, reportingParameterCode, status, statusDescription)
- protectedAreas(obsValue, timePeriod, freq, dataSet, indic_md, geo, obsStatus, structure, seeAlso, comment)
- environmentProtectionRegion(env_dom, dataSet, unit, nace_r2, obsValue, freq, env_exp, type, geo, timePeriod, obsStatus, structure, seeAlso, comment)
- freshWater(obsValue, timePeriod, freq, dataSet, src, geo, type, obsStatus, structure, seeAlso, comment)
- transportNetwork(obsValue, timePeriod, freq, dataset, unit, geo, obsStatus, structure, seeAlso, comment)

During the following experiments plural nouns were converted to single.

6.1 Test 1: EnvironmentalProtectionExpenditureRegions

6.1.1 Test 1.1

The query that we tested is the following:
environmentalProtectionExpenditureRegion(env_dom, dataSet, unit, nace_r2, obsValue, freq, env_exp, type, geo, timePeriod, obsStatus, structure, seeAlso, comment)

The narrow-down process returned the following matching candidates:

- protectedArea
- bathingWaterLoacation
- environmentProtectionRegion

However SPSM did not return any results.
Timings:

- real 0m43.178s
- user 0m41.968s
- sys 0m3.003s

Difference to previous evaluation Tests Because of the depluralisation process we have three new results: protectedArea, bathingWaterLoacation(s) and environmentProtectionRegion(s) datasets.

6.1.2 Test 1.2

We change the query predicate from environmentalProtectionExpenditureRegion to environmentalProtectionRegion. The query that we tested is the following:

environmentalProtectionRegions(env_dom, dataSet, unit, nace_r2, obsValue, freq, env_exp, type, geo, timePeriod, obsStatus, structure, seeAlso, comment)
The narrow down process returned the following matching candidates:

- protectedArea
- bathingWaterLoacation
- environmentProtectionRegion

SPSM returned the same dataset as the query.

Timings:

- real 1m9.010s
- user 0m59.695s
- sys 0m4.249s

6.2 Test 2: FreshWaters

6.2.1 Test 2.1

The query that we tested is the following:

freshWater(obsValue, timePeriod, freq, dataSet, src, geo, type, obsStatus, structure, seeAlso, comment)
The narrow-down process retuned the following matching candidates:

- waterBodyPressure
- waterBodyMeasure
- surfaceWaterBody
- freshWater
- bathingWater
- bathingWaterLocation

SPSM returned only the same dataset as the query.

Timings:

- real 1m19.009s
- user 1m14.748s
- sys 0m7.868s

Difference to previous evaluation Tests Because of the depluralisation process we have one matching candidate: bathingWater(s) dataset.

6.3 Test 3: Uk_BathingWaters

6.3.1 Test 3.1

The query that we tested is the following:

uk_BathingWaters(sampleClassification, prefLabel, long, lat, northing, easting, latestComplianceAssessment, type, country, district, envelope, latestBathingWaterProfile, sedimentTypesPresent, uriSet, regionalOrganization, yearDesignated, latestSampleAssessment, eubWidNotation, waterQualityImpactedByHeavyRain, zoneOfInfluence, samplingPoint, complianceClassification, primaryTopic, extendedMetadataVersion, definition, label)

The narrow-down process returned the following candidates:

- bathingWaterLocation
- bathingWater
- surfaceWaterBody
- waterBodyMeasure
- waterBodyPressure
- freshWater

SPMS returned the following matches:

Match1.

uk_BathingWater ↦ waterBodyMeasure
uk_BathingWater(label) ↦ waterBodyMeasure(waterBodyId)
uk_BathingWater(primaryTopic) ↦ waterBodyMeasure(secondaryMeasure)

Match2.

uk_BathingWater ↦ waterBodyPressure
uk_BathingWater(label) ↦ waterBodyPressure(waterBodyId)
uk_BathingWater(type) ↦ waterBodyPressure(pressureType)
uk_BathingWater(prefLabel) ↦ waterBodyPressure(activity)
uk_BathingWater(sampleClassification) ↦ waterBodyPressure(affectsGroundwater)

Match3.

uk_BathingWater ↦ bathingWaterLocation
uk_BathingWater(lat) ↦ bathingWaterLocation(lat)
uk_BathingWater(long) ↦ bathingWaterLocation(long)

Match4.

uk_BathingWater ↦ bathingWater
uk_BathingWater(label) ↦ bathingWater(description)
uk_BathingWater(waterQualityImpactedByHeavyRain) ↦ bathingWater(bathingWaterId).

Timings:

- real 2m5.248s
- user 1m41.483s
- sys 0m15.254s

Difference to previous evaluation Tests Because of the depluralisation process we have one new matching candidate, the bathingWaterLocation(s) dataset, and the corresponding matches.

7 Test 4: SPA_ProtectedAreas

7.1 Test 4.1

The query that we send is the following:
spa_protectedArea(area, perimeter, paCode, siteName, siteHa, status, eurCode)
The narrow-down process retuned the following results:

- environmentProtectionRegion
- protectedArea

SPSM returned the following results:
Match 1:
spa_protectedAreas $\mapsto$ protectedAreas
spa_protectedAreas(status) $\mapsto$ protectedAreas(obsStatus)
spa_protectedAreas(area) $\mapsto$ protectedAreas(structure)
Timings:

- real 0m50.696s
- user 0m38.000s
- sys 0m2.152s

A SEPA datasets

A.1 European Statistics, Fresh Water Resources dataset: Test 1.3

Original Query Schema: water(resource,measure,geo,timePeriod).

Suggested Query Schema: waterBodyPressures(identifiedDate,waterBodyId,assessmentCategory,source).

Schema Mappings:
- water $\mapsto$ waterBodyPressures
- resource $\mapsto$ source
- measure $\mapsto$ assessmentCategory
- geo $\mapsto$ waterBodyId
- timePeriod $\mapsto$ identifiedDate

Some of the answers we got are presented in 1.

Table 1: Answers for query waterBodyPressures(identifiedDate,waterBodyId,assessmentCategory,source)

identifiedDate	waterBodyId	assessmentCategory	source
"2008-04-01"	sepaidw:23088	"Morphology and Fish Continuity"	"River"
"2008-04-01"	sepaidw:20497	"Morphology and Fish Continuity"	"River"
"2008-04-01"	sepaidw:20788	"Morphology and Fish Continuity"	"Lake"
"2008-04-01"	sepaidw:23332	"Morphology and Fish Continuity"	"River"
"2008-04-01"	sepaidw:20587	"Morphology and Fish Continuity"	"River"
"2008-04-01"	sepaidw:20176	"Morphology and Fish Continuity"	"Lake"
"2008-04-01"	sepaidw:23092	"Morphology and Fish Continuity"	"River"
"2008-04-01"	sepaidw:20508	"Morphology and Fish Continuity"	"Lake"
"2008-04-01"	sepaidw:23334	"Morphology and Fish Continuity"	"River"
"2008-04-01"	sepaidw:20261	"Morphology and Fish Continuity"	"River"
"2008-04-01"	sepaidw:20304	"Morphology and Fish Continuity"	"Lake"
"2008-04-01"	sepaidw:20199	"Morphology and Fish Continuity"	"River"
"2008-04-01"	sepaidw:20770	"Morphology and Fish Continuity"	"River"
"2008-04-01"	sepaidw:23345	"Morphology and Fish Continuity"	"River"
"2008-04-01"	sepaidw:20289	"Morphology and Fish Continuity"	"Lake"
"2008-04-01"	sepaidw:23158	"Morphology and Fish Continuity"	"River"
"2008-04-01"	sepaidw:20778	"Morphology and Fish Continuity"	"Lake"
"2008-04-01"	sepaidw:20098	"Morphology and Fish Continuity"	"River"
"2008-04-01"	sepaidw:20303	"Morphology and Fish Continuity"	"Lake"
"2008-04-01"	sepaidw:23086	"Morphology and Fish Continuity"	"River"
"2008-04-01"	sepaidw:23092	"Morphology and Fish Continuity"	"River"
"2008-04-01"	sepaidw:23641	"Morphology and Fish Continuity"	"River"
"2008-04-01"	sepaidw:20173	"Morphology and Fish Continuity"	"River"
"2008-04-01"	sepaidw:20175	"Morphology and Fish Continuity"	"Lake"
"2008-04-01"	sepaidw:20257	"Morphology and Fish Continuity"	"Lake"
"2008-04-01"	sepaidw:20281	"Morphology and Fish Continuity"	"Lake"
"2008-04-01"	sepaidw:23138	"Morphology and Fish Continuity"	"River"
"2008-04-01"	sepaidw:23339	"Morphology and Fish Continuity"	"River"
"2008-04-01"	sepaidw:23350	"Morphology and Fish Continuity"	"River"
"2008-04-01"	sepaidw:20227	"Morphology and Fish Continuity"	"River"
"2008-04-01"	sepaidw:23080	"Morphology and Fish Continuity"	"River"

A.2 European Statistics, Environmental Protection Expenditure dataset: Test 2.2

Original Query Schema: location(dimension,unit,environmentDomain,environmentExposure,geo).

Suggested Query Schema: bathingWaterLocation(localAuthority)

Schema Mappings:

- unit $\mapsto$ localAuthority
- dimension $\mapsto$ no match
- environmentDomain $\mapsto$ no match
- environmentExposure $\mapsto$ no match

Table 2: Answers for query bathingWaterLocation(localAuthority)

localAuthority
"Fife"
"Highland"
"Argyll Bute"
"East Lothian"
"City of Edinburgh"
"Fife"
"Moray"
"Aberdeenshire"
"Aberdeenshire"
"Scottish Borders"
"Fife"
"Argyll Bute"
"South Ayrshire"
"Highland"
"North Ayrshire"
"Dumfries Galloway"
"City of Edinburgh"
"Moray"
"Aberdeen City"
"East Lothian"
"Highland"
"East Lothian"
"Fife"
"Dumfries Galloway"
"Dumfries Galloway"
"Inverclyde"
"East Lothian"
"Fife"
"North Ayrshire"
"Fife"
"South Ayrshire"

- geo ↦ no match

Some of the answers we got are presented in 2

A.3 European Statistics, Environmental Protection Expenditure dataset: Test 2.3

Original Query Schema: water(associatedGroundwaterId,dimension,unit,environmentDomain,environmentExposure,geo)

Suggested Query Schema: waterBodyMeasures(waterBodyId,secondaryMeasure)

Schema Mappings:
- unit ↦ secondaryMeasure
- geo ↦ waterBodyId

- associatedGroundwaterId ↦ no match
- dimension ↦ no match
- environmentDomain ↦ no match
- environmentExposure ↦ no match

Some of the answers we got are presented in 7

Original Query Schema: water(associatedGroundwaterId,dimension,unit,environmentDomain,environmentExposure,geo)

Suggested Query Schema: waterBodyPressures(waterBodyId,activity,affectsGroundwater,dataSource,industrySector)

Schema Mappings:
- water ↦ waterBodyPressures
- dimension ↦ industrySector
- unit ↦ dataSource
- environmentDomain ↦ affectsGroundwater
- environmentExposure ↦ activity
- geo ↦ waterBodyId

The above query did not return any answers. This is because there does not exist within the dataset any given object to satisfy all given columns at the same time.

A.4 Scottish Protection Agency: Protected Areas: Test 3.3

Original Query Schema: water(area,perimeter,paCode,siteName,siteHa,status,eurCode).

Suggested Query Schema: waterBodyMeasures(locationCode,waterBodyId)

Schema Mappings:
- water ↦ waterBodyMeasures
- area ↦ waterBodyId
- eurCode ↦ locationCode

Some of the answers we got are presented in 4

A.5 Scottish Protection Agency: Protected Areas: Test 3.4

Original Query Schema: Original Query : location(area,perimeter,paCode,siteName,siteHa,status,eurCode).

Suggested Query Schema: bathingWaterLocation(locationCode).

Schema Mappings:
- location ↦ bathingWaterLocation
- eurCode ↦ locationCode
- area ↦ no match
- perimeter ↦ no match
- paCode ↦ no match
- siteName ↦ no match
- siteHa ↦ no match
- status ↦ no match

Some of the answers we got are presented in 5.

A.6 The UK Environmental Agency: Bathing Waters dataset: Test 4.1

Original Query Schema: uk_BathingWaters(sampleClassification,prefLabel,long,lat,northing,easting,latestComplianceAssessment,ty

Suggested Query Schema: bathingWaters(description)

Schema Mappings:
- label ↦ description
- sampleClassification ↦ no match
- prefLabel ↦ no match
- long ↦ no match
- lat ↦ no match
- northing ↦ no match
- easting ↦ no match
- latestComplianceAssessment ↦ no match
- type ↦ no match
- country ↦ no match
- district ↦ no match
- envelope ↦ no match
- latestBathingWaterProfile ↦ no match
- sedimentTypesPresent ↦ no match
- uriSet ↦ no match
- regionalOrganization ↦ no match
- yearDesignated ↦ no match
- latestSampleAssessment ↦ no match
- eubWidNotation ↦ no match
- waterQualityImpactedByHeavyRain ↦ no match
- zoneOfInfluence ↦ no match
- samplingPoint ↦ no match
- complianceClassification ↦ no match
- primaryTopic ↦ no match
- extendedMetadataVersion ↦ no match
- definition ↦ no match

Some of the answers we got are presented in 6.

Original Query Schema: uk_BathingWaters(sampleClassification,prefLabel,long,lat,northing,easting,latestComplianceAssessment,ty

Suggested Query Schema: waterBodyPressures(waterBodyId,pressureType,activity,affectsGroundwater).

Schema Mappings:
- sampleClassification ↦ affectsGroundwater
- prefLabel ↦ activity
- type ↦ pressureType

- label ↦ waterBodyId
- long ↦ no match
- lat ↦ no match
- northing ↦ no match
- easting ↦ no match
- latestComplianceAssessment ↦ no match
- country ↦ no match
- district ↦ no match
- envelope ↦ no match
- latestBathingWaterProfile ↦ no match
- sedimentTypesPresent ↦ no match
- uriSet ↦ no match
- regionalOrganization ↦ no match
- yearDesignated ↦ no match
- latestSampleAssessment ↦ no match
- eubWidNotation ↦ no match
- waterQualityImpactedByHeavyRain ↦ no match
- zoneOfInfluence ↦ no match
- samplingPoint ↦ no match
- complianceClassification ↦ no match
- primaryTopic ↦ no match
- extendedMetadataVersion ↦ no match
- definition ↦ no match

No returned Answers.

Original Query Schema: uk_BathingWaters(sampleClassification,prefLabel,long,lat,northing,easting,latestComplianceAssessment,ty

Suggested Query Schema: waterBodyMeasures(waterBodyId,secondaryMeasure).

Schema Mappings:
- primaryTopic ↦ secondaryMeasure
- label ↦ waterBodyId
- sampleClassification ↦ no match
- prefLabel ↦ no match
- long ↦ no match
- lat ↦ no match
- northing ↦ no match
- easting ↦ no match
- latestComplianceAssessment ↦ no match
- type ↦ no match
- country ↦ no match

- district ↦ no match
- envelope ↦ no match
- latestBathingWaterProfile ↦ no match
- sedimentTypesPresent ↦ no match
- uriSet ↦ no match
- regionalOrganization ↦ no match
- yearDesignated ↦ no match
- latestSampleAssessment ↦ no match
- eubWidNotation ↦ no match
- waterQualityImpactedByHeavyRain ↦ no match
- zoneOfInfluence ↦ no match
- samplingPoint ↦ no match
- complianceClassification ↦ no match
- extendedMetadataVersion ↦ no match
- definition ↦ no match

Some of the answers we got are presented in 7.

B European Statistics, ProtectedAreas dataset

B.1 Scottish Protection Agency: Test 1.1

Original Query Schema: spa_protectedAreas(area,perimeter,paCode,siteName,siteHa,status,eurCode).

Suggested Query Schema: protectedAreas(obsStatus)

Schema Mappings:
- spa_protectedAreas ↦ waterBodyPressures status ↦ obsStatus
- area ↦ structure
- perimeter ↦ no match
- paCode ↦ no match
- siteName ↦ no match
- siteHa ↦ no match
- eurCode ↦ no match

Some of the answers we got are presented in 8.

B.2 European Statistics, Environmental Protection Expenditure dataset: Test 2.2

Original Query Schema: protectedRegions(dimension,unit,environmentDomain,environmentExposure,geo)

Suggested Query Schema: protectedAreas(geo,timePeriod).

Schema Mappings:
- protectedRegions ↦ protectedAreas
- environmentDomain ↦ timePeriod

- geo ↦ geo
- dimension ↦ no match
- unit ↦ no match
- environmentExposure ↦ no match

Some of the answers we got are presented in 9.

B.3 UK Environmental Agency: Bathing Waters Dataset: Test 3.3

Original Query Schema: uk_protectedAreas(sampleClassification, prefLabel, long, lat, northing, easting, latestComplianceAssessment, type, country, district, envelope, latestBathingWaterProfile, sedimentTypesPresent, uriSet, regionalOrganization, yearDesignated, latestSampleAssessment, eubWidNotation, waterQualityImpactedByHeavyRain, zoneOfInfluence, samplingPoint, complianceClassification, primaryTopic, extendedMetadataVersion, definition, label)

Suggested Query Schema: protectedAreas(timePeriod)

Schema Mappings:
- uk_protectedAreas ↦ protectedAreas
- type ↦ no match
- yearDesignated ↦ timePeriod
- sampleClassification ↦ no match
- prefLabel ↦ no match
- long ↦ no match
- lat ↦ no match
- northing ↦ no match
- easting ↦ no match
- latestComplianceAssessment ↦ no match
- country ↦ no match
- district ↦ no match
- envelope ↦ no match
- latestBathingWaterProfile ↦ no match
- sedimentTypesPresent ↦ no match
- uriSet ↦ no match
- regionalOrganization ↦ no match
- latestSampleAssessment ↦ no match
- eubWidNotation ↦ no match
- waterQualityImpactedByHeavyRain ↦ no match
- zoneOfInfluence ↦ no match
- samplingPoint ↦ no match
- complianceClassification ↦ no match

- primaryTopic ↦ no match
- extendedMetadataVersion ↦ no match
- definition ↦ no match
- label ↦ no match

Some of the answers we got are presented in 10.

C European Statistics, Transport Network dataset

C.1 NAPTAN, Stop Availability dataset: Test 1.2

Original Query Schema: transport(atcoCode, startDate, endDate, availabilityStatus, note,noteLang, transferStopAtcoCode, creationDateTime, modificationDateTime, revisionNumber, modification)

Suggested Query Schema: transportNetwork(comment)

Schema Mappings:
- transport ↦ transportNetwork
- note ↦ comment
- atcoCode ↦ no match
- startDate ↦ no match
- endDate ↦ no match
- availabilityStatus ↦ no match
- noteLang ↦ no match
- transferStopAtcoCode ↦ no match
- creationDateTime ↦ no match
- modificationDateTime ↦ no match
- revisionNumber ↦ no match
- modification ↦ no match

Some of the answers we got are presented in 11.

Table 3: Answers for query waterBodyMeasures(waterBodyId,secondaryMeasure)

waterBodyId	secondaryMeasure
sepaidw:150342	"Control Abstraction"
sepaidw:100269	"Removal of barriers or provision of mechanisms to enable fish migration"
sepaidw:10757	"Non-urban land management measures"
sepaidw:23330	"Improvement to condition of riparian zone and/or wetland habitats"
sepaidw:23190	"Improvement to condition of riparian zone and/or wetland habitats"
sepaidw:5802	"Improvement to condition of riparian zone and/or wetland habitats"
sepaidw:10211	"Improvement to condition of riparian zone and/or wetland habitats"
sepaidw:10042	"Change timing or frequency of discharge"
sepaidw:10214	"Improvement to condition of riparian zone and/or wetland habitats"
sepaidw:150104	"Control Abstraction"
sepaidw:3002	"Reduce at source"
sepaidw:23482	"Control pattern/timing of abstraction (Hands off flow/utilisation of storage (new/existing)'
sepaidw:20278	"Control Abstraction"
sepaidw:10000	"Reduce at source"
sepaidw:6503	"Control Abstraction"
sepaidw:6525	"Improve Modified Habitat"
sepaidw:5904	"Increase treatment"
sepaidw:23322	"Removal of barriers or provision of mechanisms to enable fish migration"
sepaidw:100291	"Improve Modified Habitat"
sepaidw:6563	"Relocate all or part of discharge"
sepaidw:10202	"Control pattern/timing of abstraction (Hands off flow/utilisation of storage (new/existing)'
sepaidw:20307	"Improve Modified Habitat"
sepaidw:10063	"Removal of barriers or provision of mechanisms to enable fish migration"
sepaidw:4206	"Reduce Diffuse Source Inputs"
sepaidw:6525	"Control Abstraction"
sepaidw:20342	"Control pattern/timing of abstraction (Hands off flow/utilisation of storage (new/existing)'
sepaidw:20609	"Removal of barriers or provision of mechanisms to enable fish migration"
sepaidw:4000	"Increase treatment"
sepaidw:10022	"Change timing or frequency of discharge"
sepaidw:10673	"Provide appropriate baseline flow regime downstream of impoundment"
sepaidw:10251	"Improvement to condition of riparian zone and/or wetland habitats"
sepaidw:23182	"Removal of barriers or provision of mechanisms to enable fish migration"
sepaidw:10047	"Reduce at source"
sepaidw:10003	"Change timing or frequency of discharge"
sepaidw:10029	"Removal of barriers or provision of mechanisms to enable fish migration"
sepaidw:10458	"Removal of barriers or provision of mechanisms to enable fish migration"
sepaidw:20245	"Improvement to condition of channel/bed and/or banks/shoreline"
sepaidw:23223	"Reduce Diffuse Source Inputs"
sepaidw:23270	"Reduce at source"
sepaidw:3000	"Reduce at source"
sepaidw:6882	"Control Abstraction"
sepaidw:4207	"Reduce Point Source Inputs"
sepaidw:20387	"Reduce Diffuse Source Inputs"

Table 4: Answers for query waterBodyMeasures(locationCode,waterBodyId)

locationCode	waterBodyId
sepaidloc:333346	sepaidw:150112
sepaidloc:329473	sepaidw:5901
sepaidloc:811	sepaidw:3105
sepaidloc:331030	sepaidw:150257
sepaidloc:331476	sepaidw:150340
sepaidloc:112953	sepaidw:10002
sepaidloc:337942	sepaidw:4728
sepaidloc:315756	sepaidw:10341
sepaidloc:125190	sepaidw:10666
sepaidloc:337783	sepaidw:100311
sepaidloc:334244	sepaidw:10790
sepaidloc:311241	sepaidw:10397
sepaidloc:339861	sepaidw:200242
sepaidloc:311099	sepaidw:10374
sepaidloc:315120	sepaidw:20111
sepaidloc:330454	sepaidw:150261
sepaidloc:341476	sepaidw:150194
sepaidloc:331899	sepaidw:23034
sepaidloc:373044	sepaidw:23571
sepaidloc:311347	sepaidw:10747
sepaidloc:328579	sepaidw:6109
sepaidloc:334233	sepaidw:23138
sepaidloc:311857	sepaidw:10829
sepaidloc:315411	sepaidw:20220
sepaidloc:332865	sepaidw:20342
sepaidloc:200781	sepaidw:20798
sepaidloc:315456	sepaidw:23392
sepaidloc:7	sepaidw:3107
sepaidloc:332275	sepaidw:10600
sepaidloc:16487	sepaidw:150243
sepaidloc:1175	sepaidw:4001
sepaidloc:344655	sepaidw:100299
sepaidloc:324428	sepaidw:23020
sepaidloc:125173	sepaidw:200011
sepaidloc:77777	sepaidw:10069
sepaidloc:127765	sepaidw:10040
sepaidloc:128568	sepaidw:10918

Table 5: Answers for query bathingWaterLocation(locationCode)

locationCode
sepaidloc:366962
sepaidloc:4561
sepaidloc:124793
sepaidloc:114582
sepaidloc:206179
sepaidloc:11789
sepaidloc:9318
sepaidloc:9334
sepaidloc:4552
sepaidloc:233617
sepaidloc:124770
sepaidloc:206178
sepaidloc:9335
sepaidloc:235334
sepaidloc:124482
sepaidloc:235333
sepaidloc:9320
sepaidloc:9316
sepaidloc:206238
sepaidloc:124817
sepaidloc:4593

Table 6: Sample answers for query bathingWaters(description)

description
"Largs (Pencil Beach)"
"Cullen Bay"
"Achmelvich"
"Lunan Bay"
"Mossyard"
"Nairn (East)"
"Dores"
"Irvine"
"Kinghorn (Harbour Beach)"
"Prestwick"
"Thurso"
"Fraserburgh (Philorth)"
"Dunbar (East)"
"Rockcliffe"
"Dhoon Bay"
"Inverboyndie"
"Cruden Bay"
"Leven"
"Brighouse Bay"
"Monifieth"
"Crail (Roome Bay)"
"Findhorn"
"Broad Sands"
"Girvan"
"St. Andrews (West Sands)"
"Yellowcraig"
"Portobello (Central)"
"North Berwick (West)"
"Dunbar (Belhaven)"
"Burntisland"
"Southerness"
"Portobello (West)"
"Thorntonloch"
"Saltcoats/Ardrossan "
"Elie (Ruby Bay)"
"Culzean"
"Peterhead (Lido)"
"Seacliff"
"Kinghorn (Pettycur)"
"Rosemarkie"
"Ganavan "
"Arbroath (West Links)"
"Rosehearty"
"Kingsbarns"
"St. Andrews (East Sands)"
"Aberdour (Silver Sands)"
"Seamill"
"Sandyhills"
"Montrose"

Table 7: Sample Results for query waterBodyMeasures(waterBodyId,secondaryMeasure)

waterBodyId	secondaryMeasure
sepaidw:150342	"Control Abstraction" sepaidw:100269
sepaidw:10757	"Non-urban land management measures"
sepaidw:23330	"Improvement to condition of riparian zone and/or wetland habitats"
sepaidw:23190	"Improvement to condition of riparian zone and/or wetland habitats"
sepaidw:5802	"Improvement to condition of riparian zone and/or wetland habitats"
sepaidw:10211	"Improvement to condition of riparian zone and/or wetland habitats"
sepaidw:10042	"Change timing or frequency of discharge"
sepaidw:10214	"Improvement to condition of riparian zone and/or wetland habitats"
sepaidw:150104	"Control Abstraction"
sepaidw:3002	"Reduce at source"
sepaidw:23482	"Control pattern/timing of abstraction (Hands off flow/utilisation of storage (new/existing)'
sepaidw:20278	"Control Abstraction"
sepaidw:10000	"Reduce at source"
sepaidw:6503	"Control Abstraction"
sepaidw:6525	"Improve Modified Habitat"
sepaidw:5904	"Increase treatment"
sepaidw:23322	"Removal of barriers or provision of mechanisms to enable fish migration"
sepaidw:100291	"Improve Modified Habitat"
sepaidw:6563	"Relocate all or part of discharge"
sepaidw:10202	"Control pattern/timing of abstraction (Hands off flow/utilisation of storage (new/existing)'
sepaidw:20307	"Improve Modified Habitat"
sepaidw:10063	"Removal of barriers or provision of mechanisms to enable fish migration"
sepaidw:4206	"Reduce Diffuse Source Inputs"
sepaidw:6525	"Control Abstraction"
sepaidw:20342	"Control pattern/timing of abstraction (Hands off flow/utilisation of storage (new/existing)'
sepaidw:20609	"Removal of barriers or provision of mechanisms to enable fish migration"
sepaidw:4000	"Increase treatment"
sepaidw:10022	"Change timing or frequency of discharge"
sepaidw:10673	"Provide appropriate baseline flow regime downstream of impoundment"
sepaidw:10251	"Improvement to condition of riparian zone and/or wetland habitats"
sepaidw:23182	"Removal of barriers or provision of mechanisms to enable fish migration"
sepaidw:10047	"Reduce at source"
sepaidw:10003	"Change timing or frequency of discharge"
sepaidw:10029	"Removal of barriers or provision of mechanisms to enable fish migration"
sepaidw:10458	"Removal of barriers or provision of mechanisms to enable fish migration"
sepaidw:20245	"Improvement to condition of channel/bed and/or banks/shoreline"
sepaidw:23223	"Reduce Diffuse Source Inputs"
sepaidw:23270	"Reduce at source"
sepaidw:3000	"Reduce at source"
sepaidw:6882	"Control Abstraction"
sepaidw:4207	"Reduce Point Source Inputs"
sepaidw:20387	"Reduce Diffuse Source Inputs"
sepaidw:6311	"Change timing or frequency of discharge"
sepaidw:100232	"Appropriate management of rate and range of artificial drawdown"

Table 8: Sample Results for query protectedAreas(obsStatus)

7

obsStatus
”¡http://eurostat.linked-statistics.org/dic/obs_status#s¿”
” ¡http://eurostat.linked-statistics.org/dic/obs_status#s¿”
”¡http://eurostat.linked-statistics.org/dic/obs_status#i¿”
” ¡http://eurostat.linked-statistics.org/dic/obs_status#s¿”
” ¡http://eurostat.linked-statistics.org/dic/obs_status#s¿”
” ¡http://eurostat.linked-statistics.org/dic/obs_status#s¿”

Table 9: Sample Results for query protectedAreas(geo,timePeriod)

geo	timePeriod
¡http://eurostat.linked-statistics.org/dic/geo#JO¿	”2006-01-01”
¡http://eurostat.linked-statistics.org/dic/geo#SY¿	”2003-01-01”
¡http://eurostat.linked-statistics.org/dic/geo#PS¿	”2002-01-01”
¡http://eurostat.linked-statistics.org/dic/geo#TN¿	”2008-01-01”
¡http://eurostat.linked-statistics.org/dic/geo#MA¿	”2001-01-01”
¡http://eurostat.linked-statistics.org/dic/geo#MA¿	”2009-01-01”
¡http://eurostat.linked-statistics.org/dic/geo#MA¿	”2000-01-01”
¡http://eurostat.linked-statistics.org/dic/geo#JO¿	”2003-01-01”
¡http://eurostat.linked-statistics.org/dic/geo#MA¿	”2011-01-01”
¡http://eurostat.linked-statistics.org/dic/geo#LB¿	”2000-01-01”
¡http://eurostat.linked-statistics.org/dic/geo#JO¿	”2004-01-01”
¡http://eurostat.linked-statistics.org/dic/geo#EG¿	”2004-01-01”
¡http://eurostat.linked-statistics.org/dic/geo#SY¿	”2003-01-01”
¡http://eurostat.linked-statistics.org/dic/geo#EG¿	”2010-01-01”
¡http://eurostat.linked-statistics.org/dic/geo#PS¿	”2004-01-01”
¡http://eurostat.linked-statistics.org/dic/geo#PS¿	”2000-01-01”
¡http://eurostat.linked-statistics.org/dic/geo#LB¿	”2005-01-01”
¡http://eurostat.linked-statistics.org/dic/geo#MA¿	”2006-01-01”
¡http://eurostat.linked-statistics.org/dic/geo#EG¿	”2004-01-01”
¡http://eurostat.linked-statistics.org/dic/geo#LB¿	”2005-01-01”
¡http://eurostat.linked-statistics.org/dic/geo#MA¿	”2005-01-01”
¡http://eurostat.linked-statistics.org/dic/geo#PS¿	”2002-01-01”
¡http://eurostat.linked-statistics.org/dic/geo#TN¿	”2008-01-01”
¡http://eurostat.linked-statistics.org/dic/geo#MA¿	”2003-01-01”
¡http://eurostat.linked-statistics.org/dic/geo#MA¿	”2005-01-01”8sd:dat
¡http://eurostat.linked-statistics.org/dic/geo#PS¿	”2005-01-01”
¡http://eurostat.linked-statistics.org/dic/geo#EG¿	”2007-01-01”

Table 10: Sample Results for query protectedAreas(timePeriod)

timePeriod
"2006-01-01"
"2003-01-01"
"2002-01-01"
"2008-01-01"
"2001-01-01"
"2009-01-01"
"2000-01-01"
"2003-01-01"
"2011-01-01"
"2000-01-01"
"2004-01-01"
"2004-01-01"
"2003-01-01"
"2010-01-01"
"2004-01-01"
"2000-01-01"
"2005-01-01"
"2006-01-01"
"2004-01-01"
"2005-01-01"
"2005-01-01"
"2002-01-01"
"2008-01-01"
"2003-01-01"
"2005-01-01"8sd:dat
"2005-01-01"
"2007-01-01"

Table 11: Sample Results for query transportNetwork(comment)

comment
Reused Eurostat Linked Data Wrapper (http://estatwrap.ontologycentral.com/) to rdfize Eurostat datasets (h