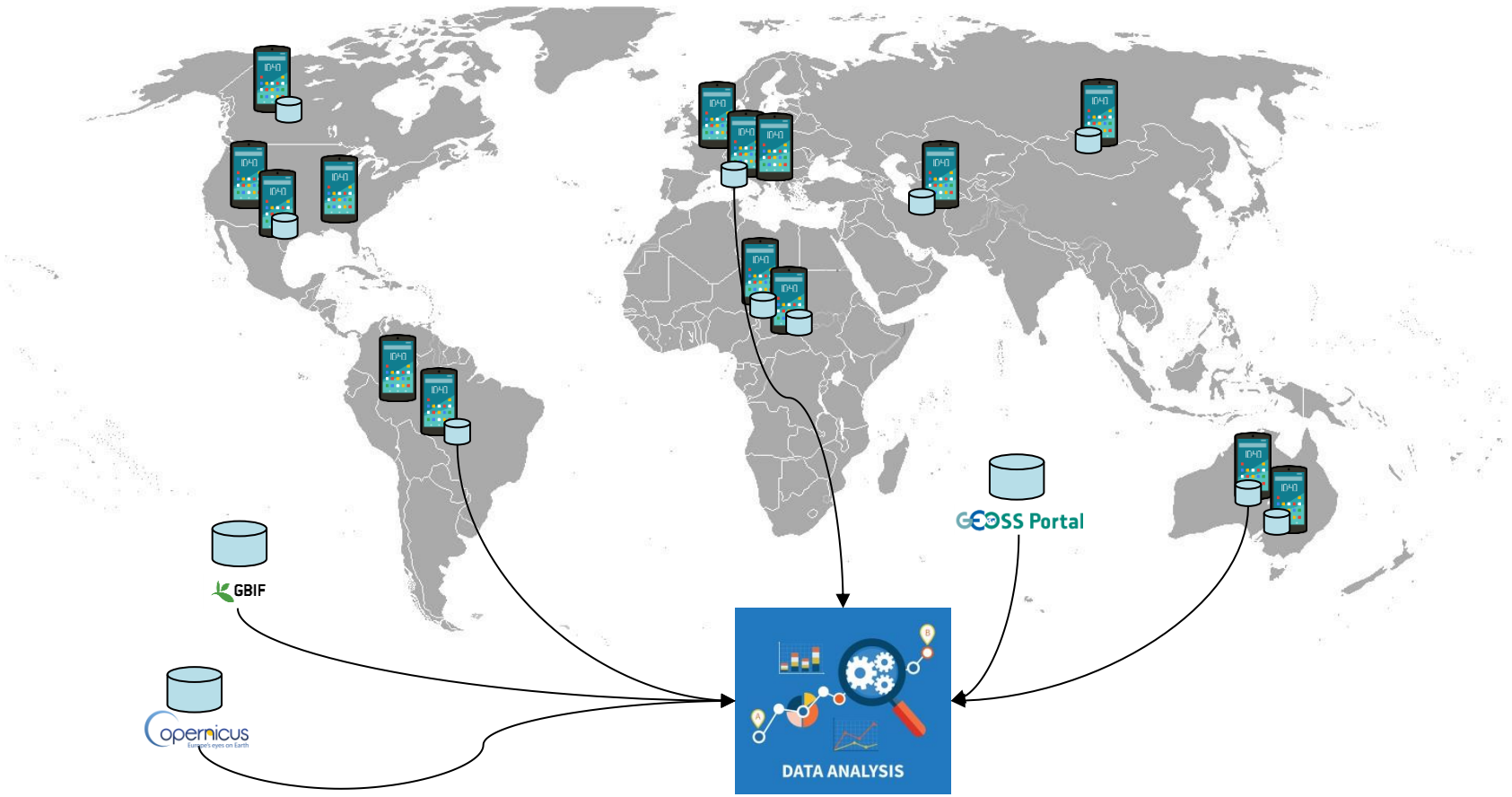
Bringing Semantics to Citizen Science Data – A Semantic Extension of Open Data Kit 1

Markus D. Steinberg[1], Sirko Schindler[2], Friederike Klan[3]
[1]*Friedrich-Schiller-Universität, Jena*
[2]*Data Management Technologies Group, DLR Institute of Data Science, Jena*
[3]*Citizen Science Group, DLR Institute of Data Science, Jena*

EGU General Assembly 2019
Vienna, 2019-04-12

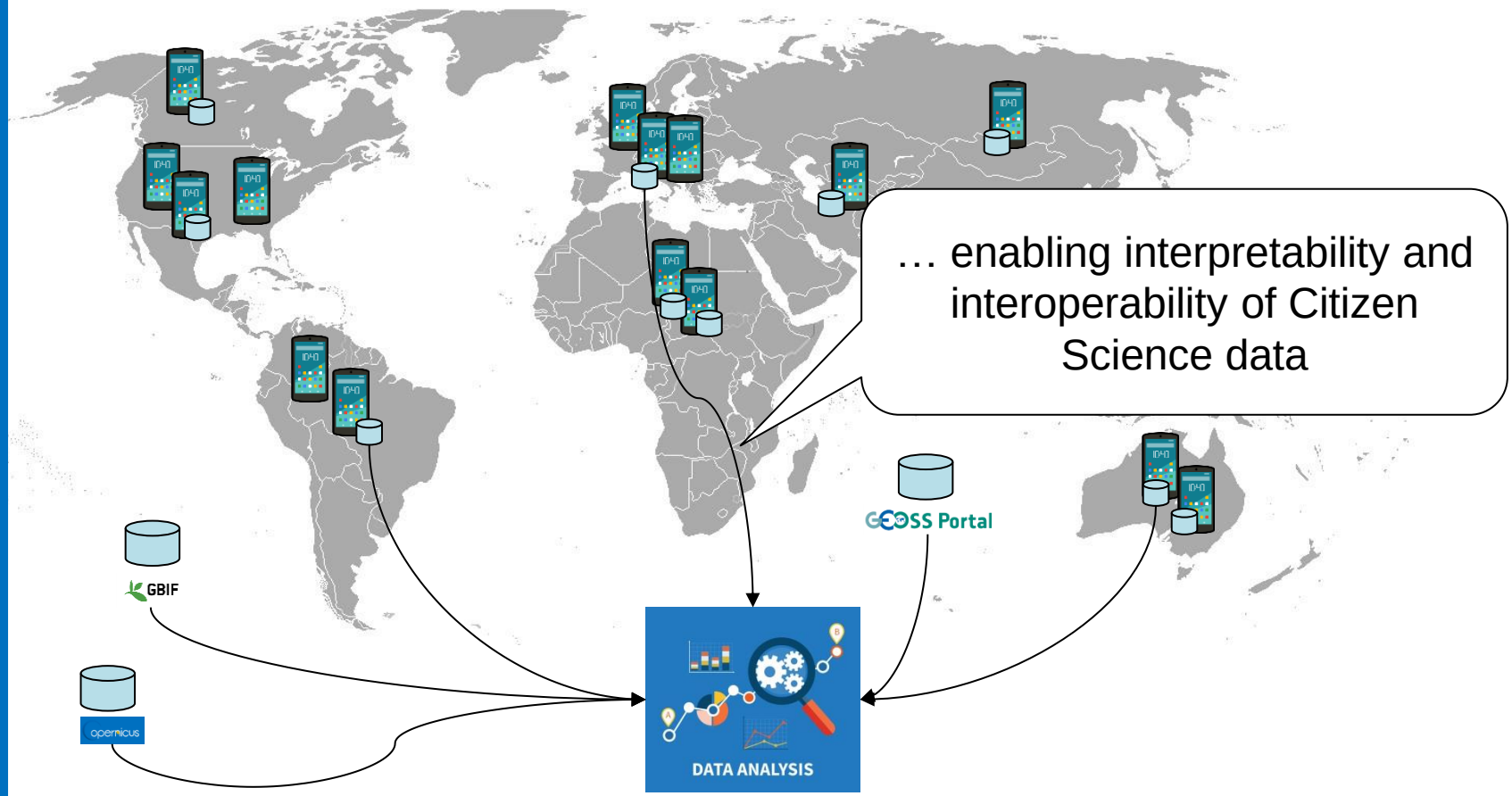
EGU2019-19123

Unlocking the Value of Citizen Contributions by …



Unlocking the Value of Citizen Contributions by …

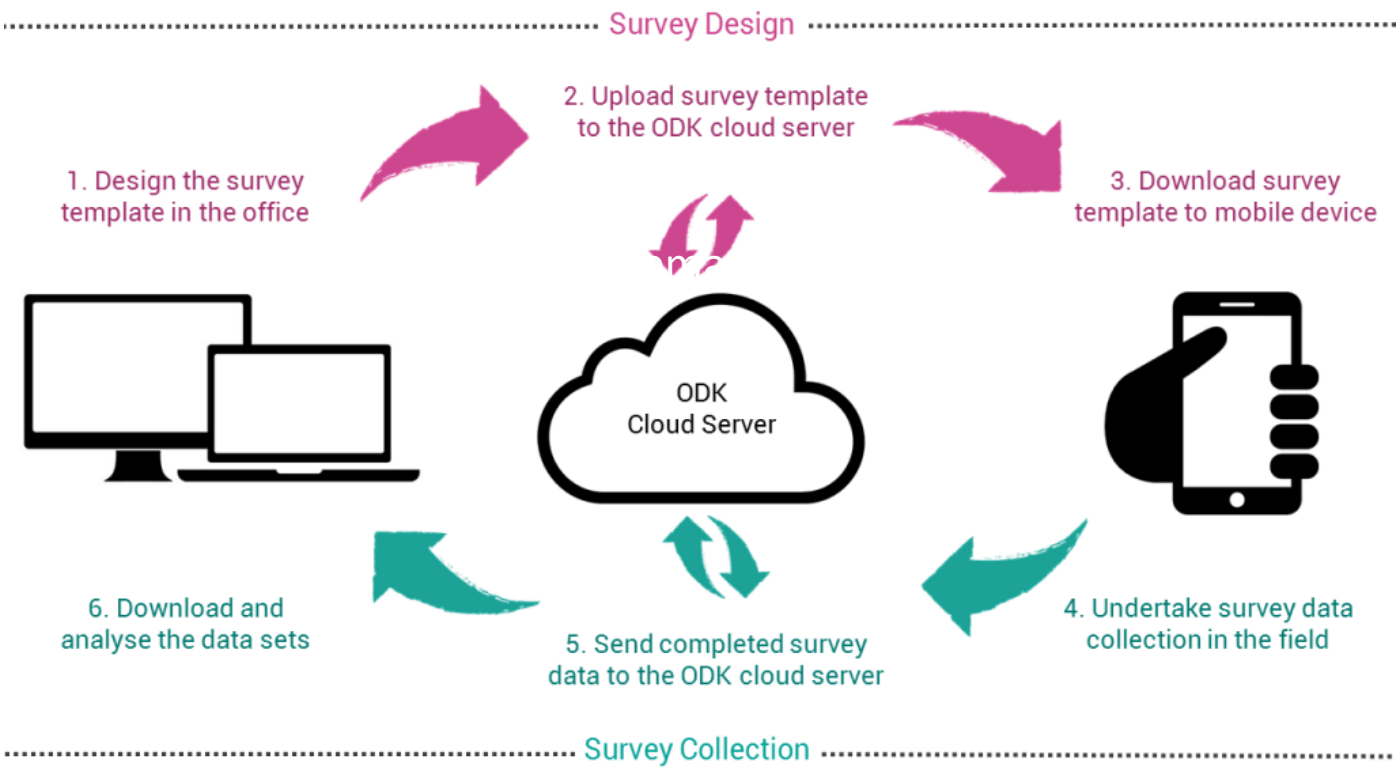
Reusable Software & Best Practices



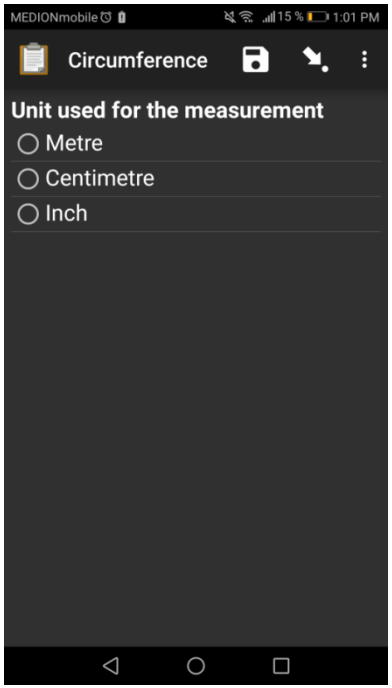
DLR

FRIEDRICH-SCHILLER-UNIVERSITÄT JENA

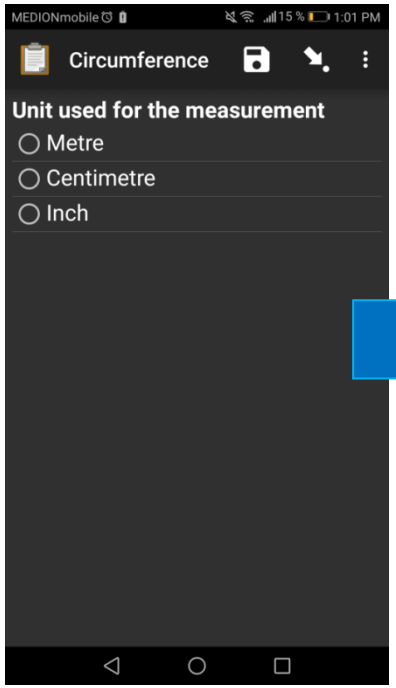
Mobile Data Collection (with ODK)

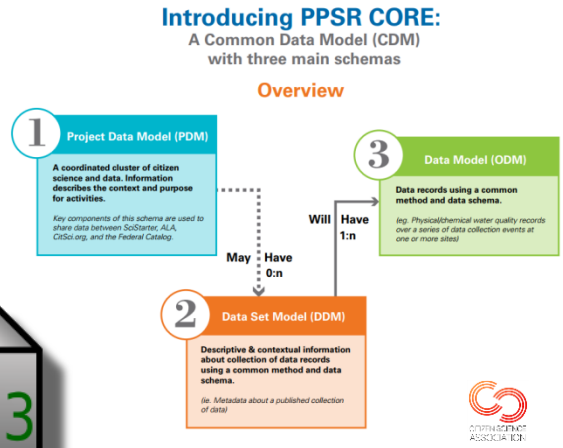


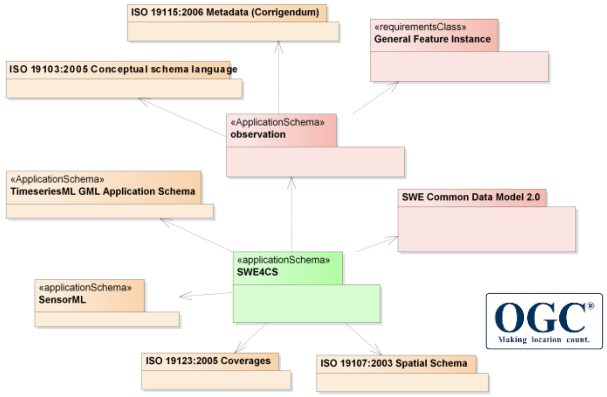
Enabling CS Data Interoperability & Reuse



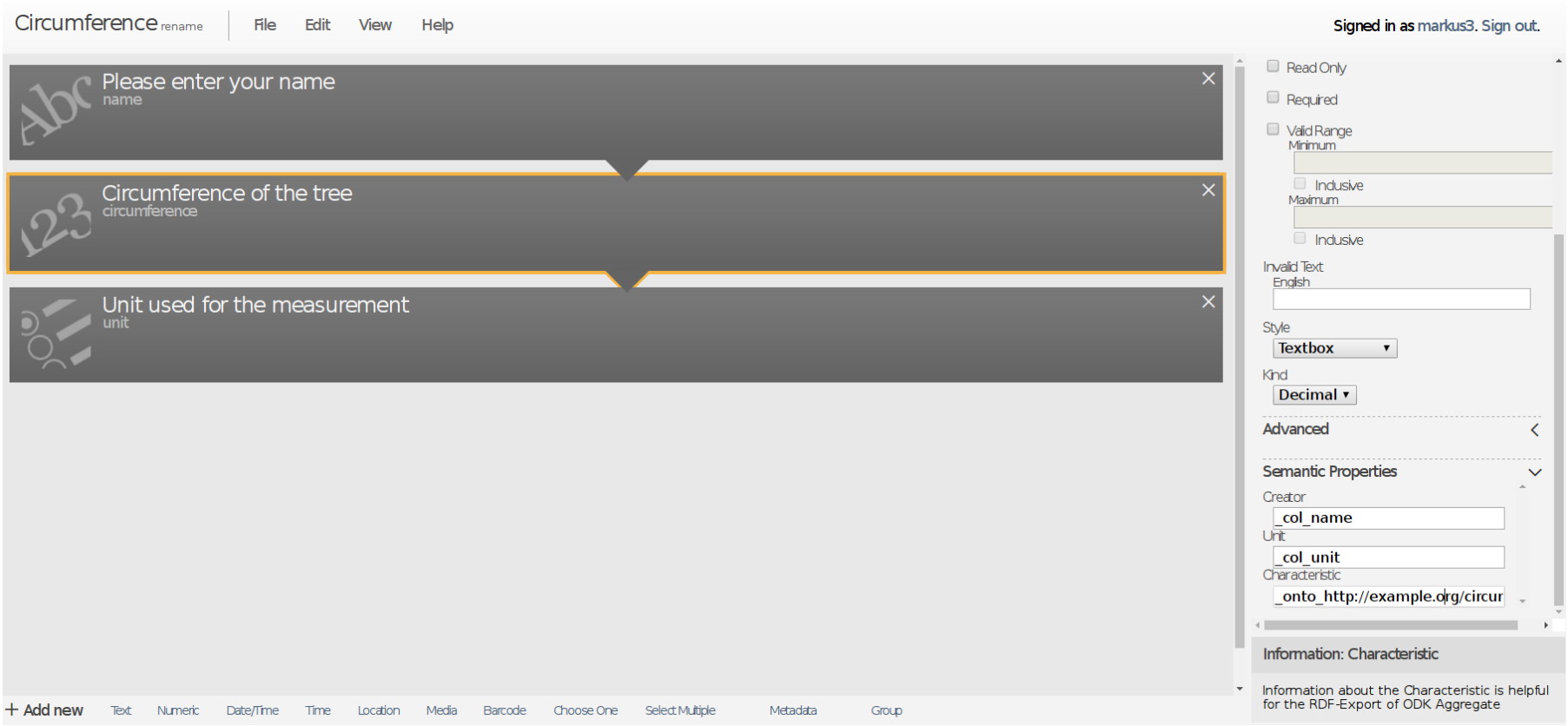
Enabling CS Data Interoperability & Reuse



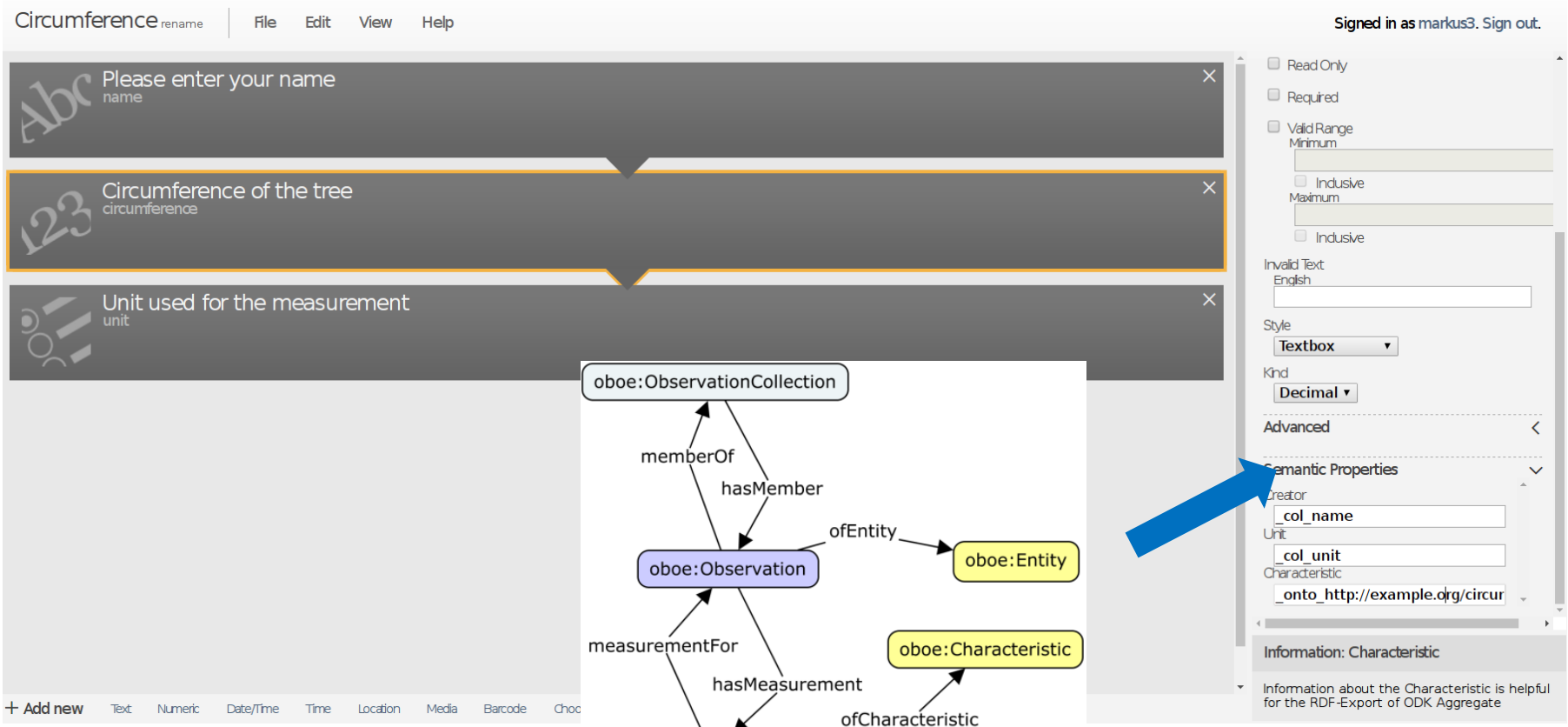


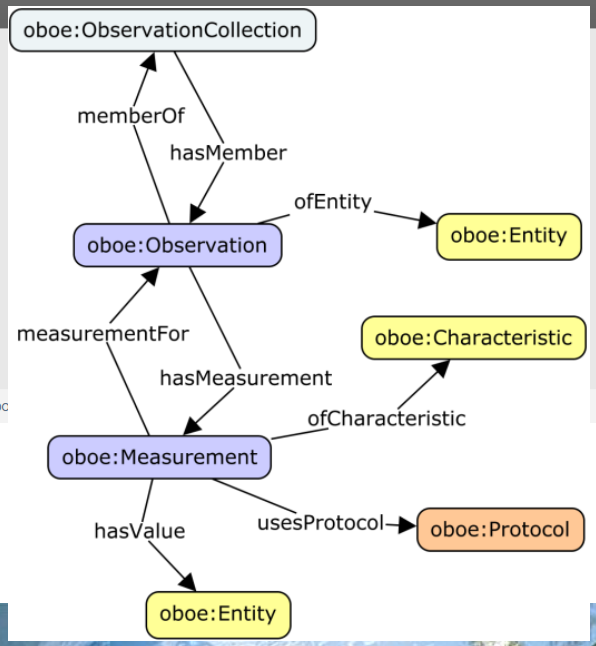


(1) Semantic Annotation at Design Time

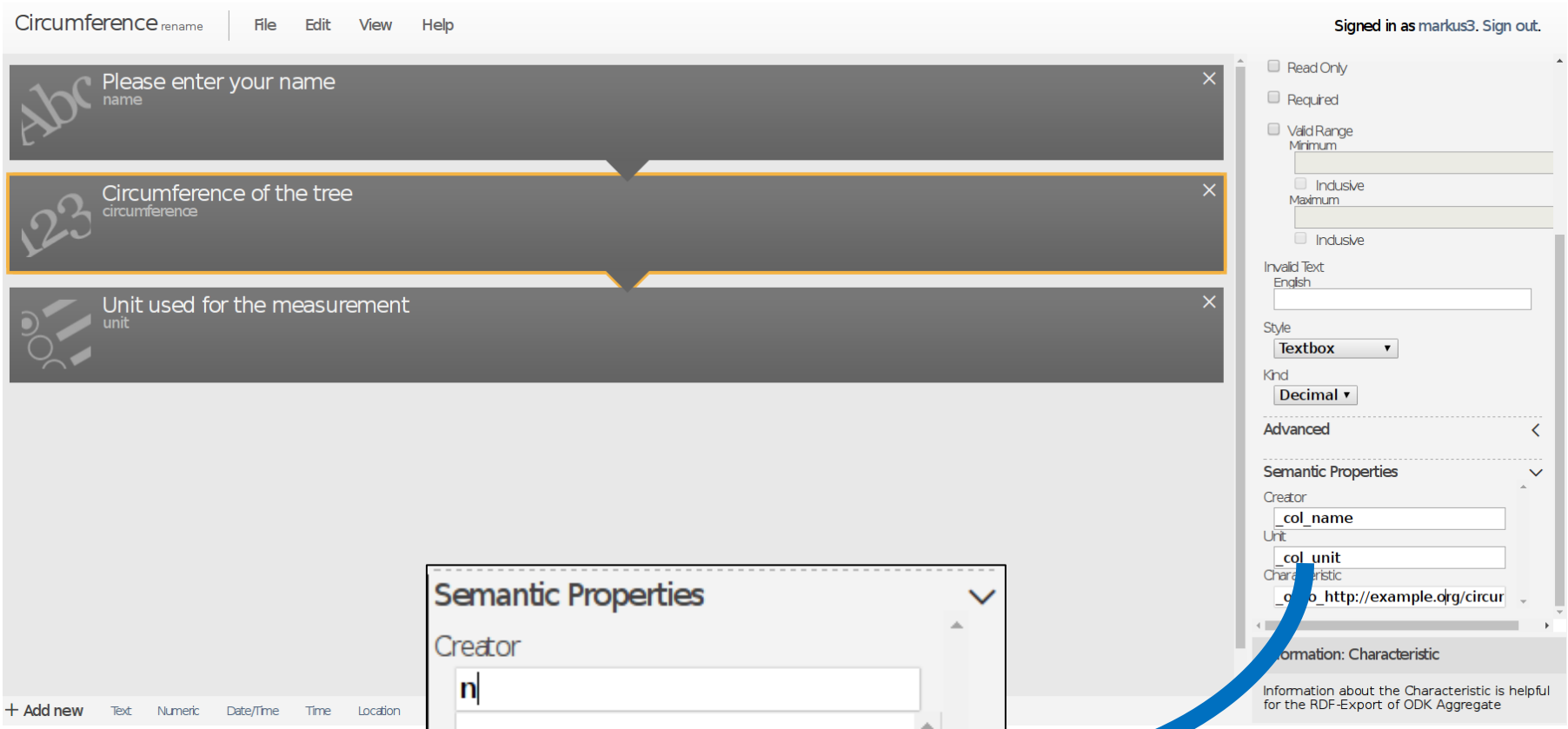


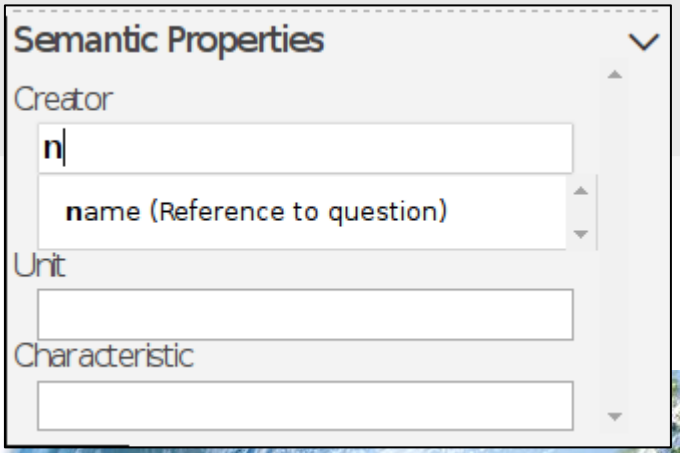
(1) Semantic Annotation at Design Time



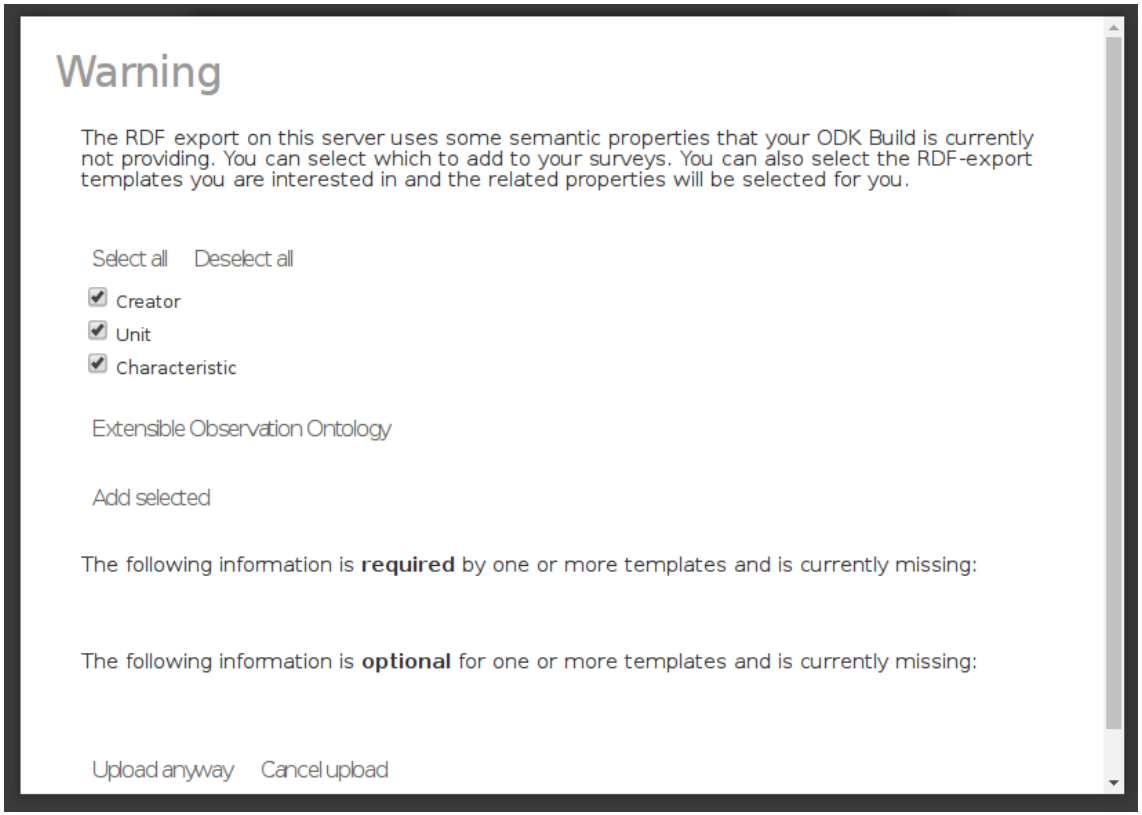


(1) Semantic Annotation at Design Time

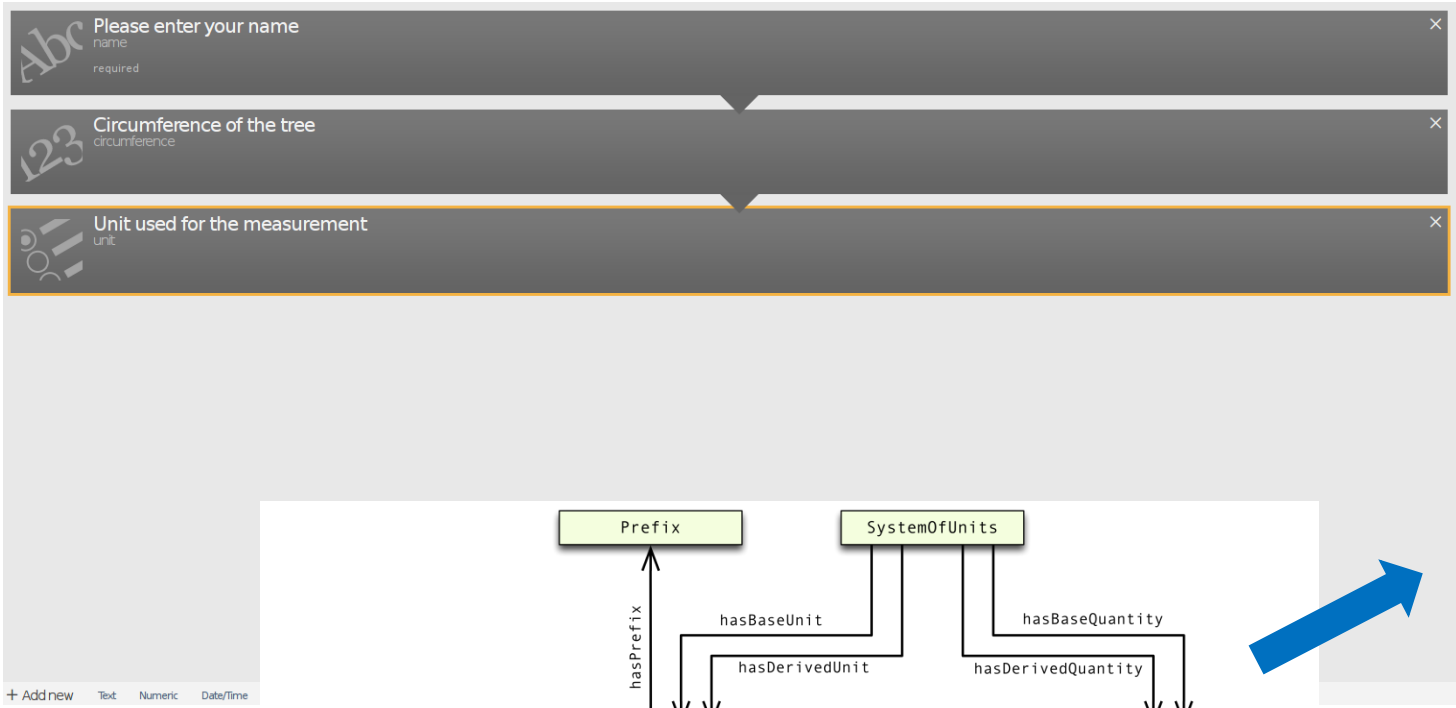


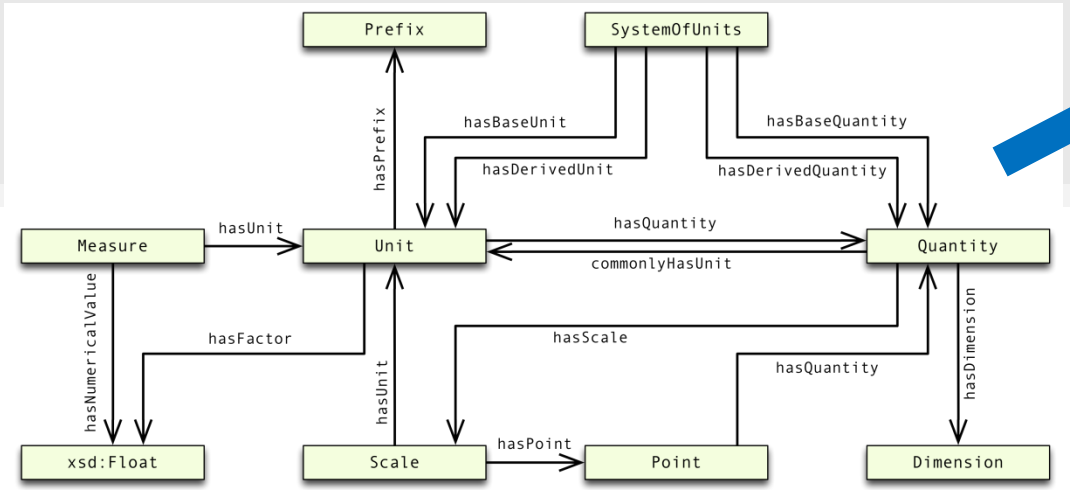


(1) Semantic Annotation at Design Time

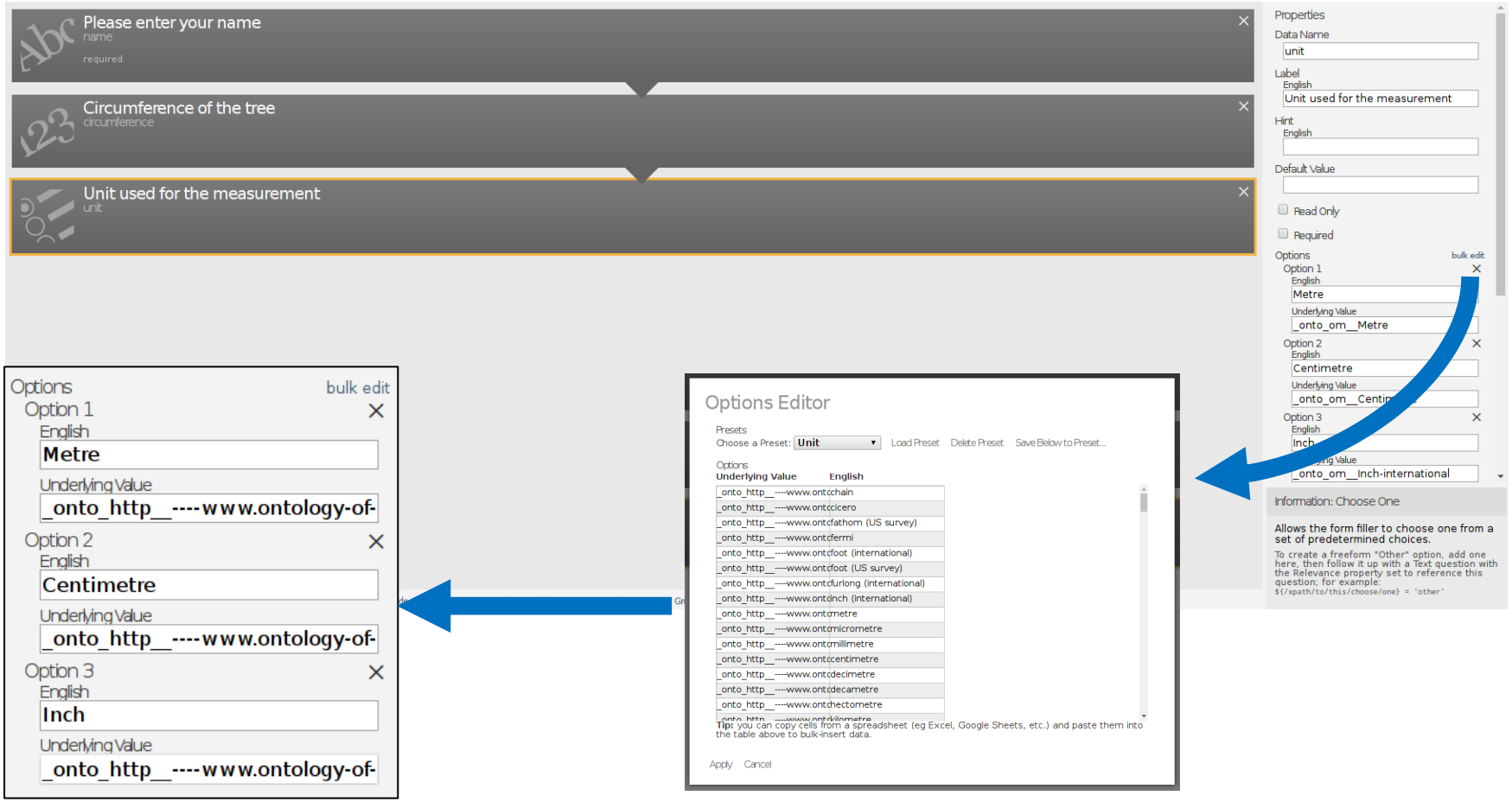


(1) Semantic Annotation at Design Time

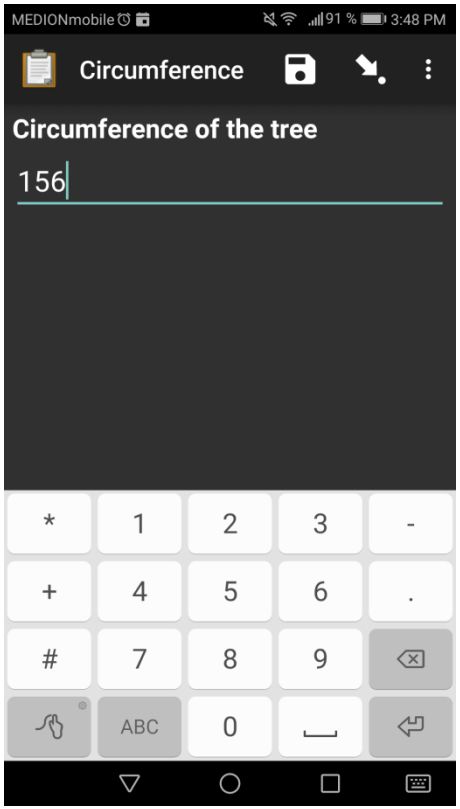


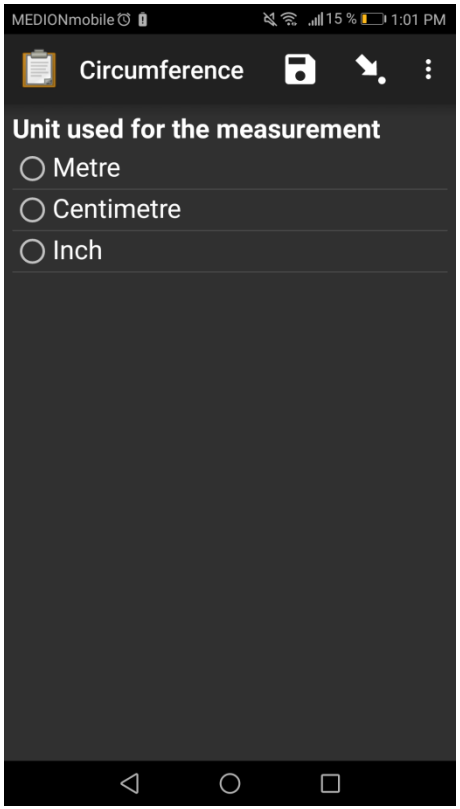


(1) Semantic Annotation at Design Time

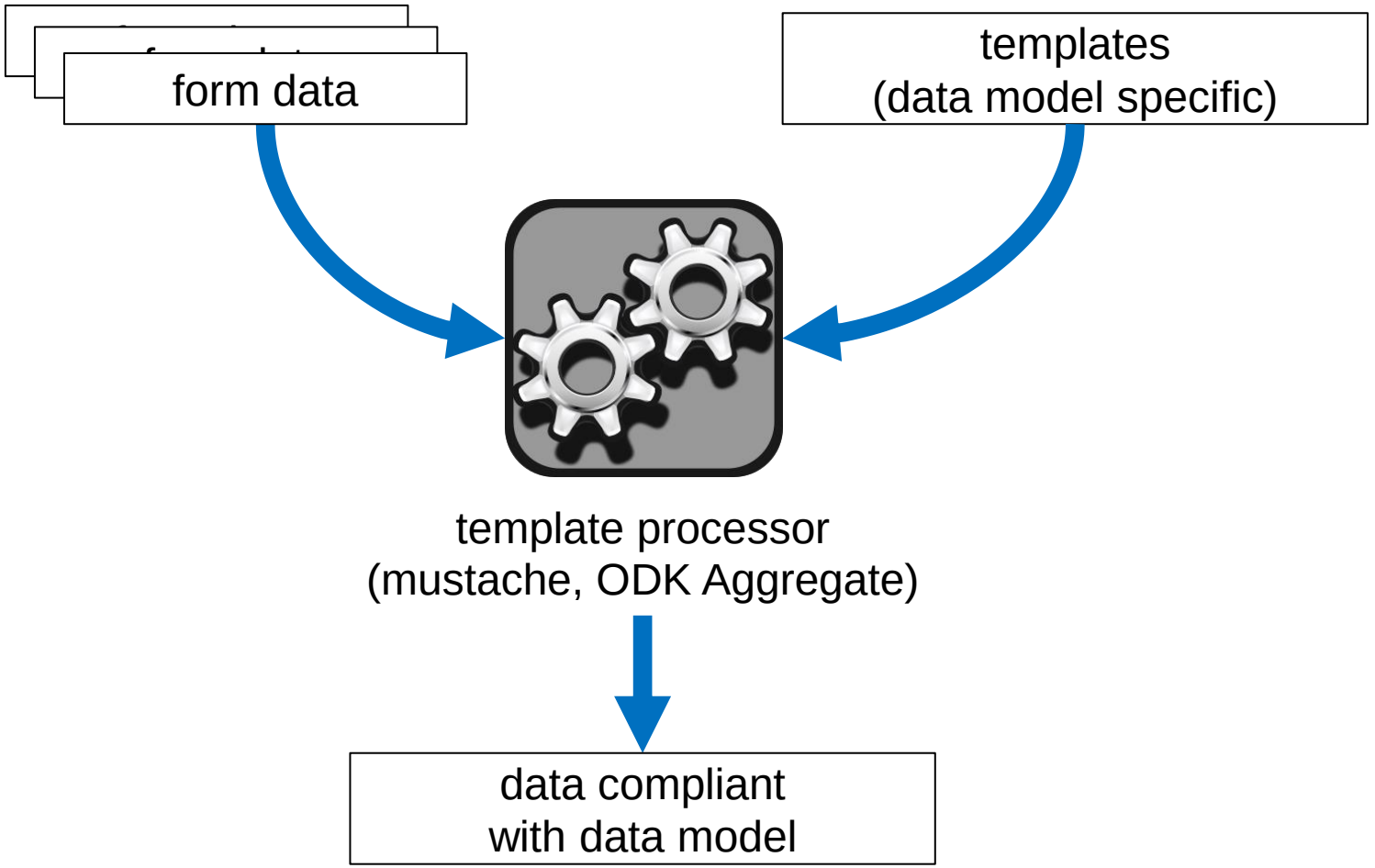


(2) Transparent during Data Collection



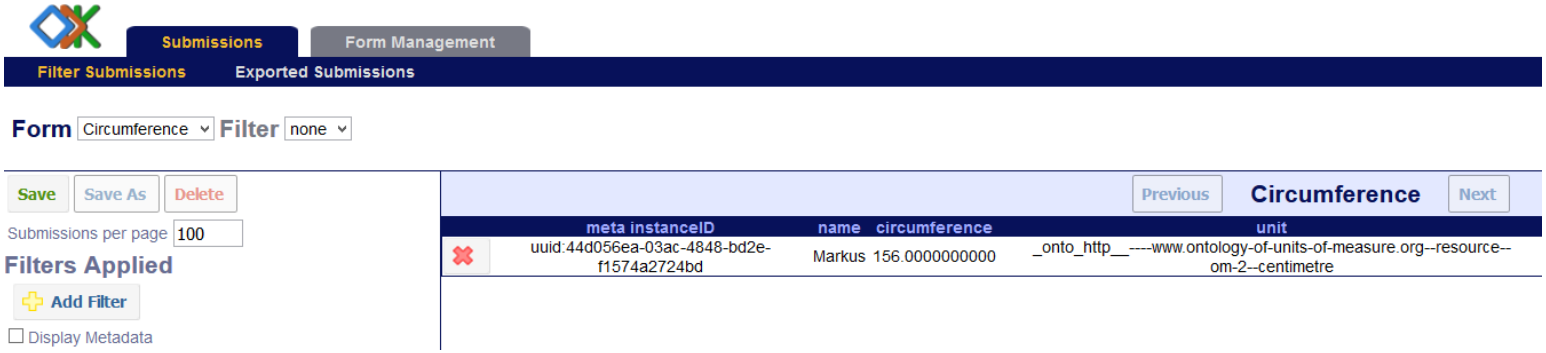


(3) Flexible Data Export via Templates



mustache - https://mustache.github.io/
Open Data Kit (ODK) - https://opendatakit.org/

(3) Flexible Data Export via Templates




```
{{!Construct an oboe-core:Measurement for the cell}}
{{cellEntityIdentifier}} rdf:type oboe-core:Measurement .

{{!Relate the Measurement to the Observation}}
{{columnModel.columnEntityIdentifier}} oboe-core:hasMeasurement {{cellEntityIdentifier}} .

{{!Construct an oboe-core:MeasuredValue for the cell}}
:measuredValue_{{rowModel.rowId}}_{{columnModel.columnHeader}}
        rdf:type oboe-core:MeasuredValue .

{{!Relate the Measurement to the MeasuredValue}}
{{cellEntityIdentifier}}
        oboe-core:hasValue :measuredValue_{{rowModel.rowId}}_{{columnModel.columnHeader}} .
```

```
{{^semantics.Unit.isLiteral}}
    {{cellEntityIdentifier}} om:hasUnit <{{{semantics.Unit.value}}}> .
{{/semantics.Unit.isLiteral}}
```

```
:measurement_1_circumference rdf:type oboe-core:Measurement .
:column_circumference oboe-core:hasMeasurement
:measurement_1_circumference .
:measuredValue_1_circumference rdf:type oboe-core:MeasuredValue .
:measurement_1_circumference oboe-core:hasValue
:measuredValue_1_circumference .

:measurement_1_circumference dcterms:creator
:creator_circumference_1 .
:creator_circumference_1 rdf:type foaf:Person .
:creator_circumference_1 foaf:name "Markus" .

:measurement_1_circumference
      oboe-core:ofCharacteristic <http://example.org/circumference>
.
:measurement_1_circumference
      om:hasUnit <http://www.ontology-of-units-of-
measure.org/resource/om-2/centimetre> .

:measuredValue_1_circumference oboe-core:hasCode
"156.0000000000"^^xsd:decimal .
```

Contact

Markus Daniel Steinberg markus.daniel.steinberg@uni-jena.de
Sirko Schindler sirko.schindler@dlr.de
Friederike Klan friederike.klan@dlr.de

Publications

M. D. Steinberg, S. Schindler, F. Klan: Software Solutions for Form-Based, Mobile Data Collection - A Comparative Evaluation,
18th Conference on Datenbanksysteme für Business, Technologie und Web (BTW 2019), Rostock, 2019, https://arxiv.org/abs/1901.11053

M.D. Steinberg, S. Schindler, F. Klan: Bringing Semantics to Citizen Data Collection - A Semantic Extension of Open Data Kit 1, Geophysical Research Abstracts, Vol. 21, EGU2019-19123, EGU General Assembly 2019

Software

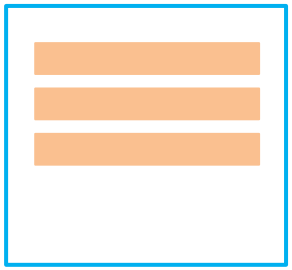
https://github.com/MaSteinberg/aggregate
https://github.com/MaSteinberg/build

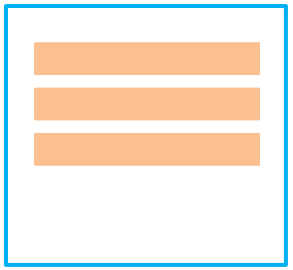
Image Sources

- p.2/3:
  - https://commons.wikimedia.org/wiki/File:BlankMap-World-v2.png
  - http://www.geoportal.org/documents/20181/24501/geoss-portal-beta.png/5b52dbab-a3c8-4962-8708-010ccff65c94?t=1480584717161
  - https://tr1.cbsistatic.com/hub/i/r/2015/09/30/f66fef72-5d57-48cc-bd8c-7ec0941e73ef/resize/770x/ac0363c3e414d51bafb1dd9b1c8a604a/dataanalysisistockrobuart.jpg
  - https://www.eea.europa.eu/about-us/who/copernicus-1/copernicus-logo/image
  - https://images.ctfassets.net/uo17ejk9rkwj/4rmEF9F4ZGiCwQk2WSMcce/b47146eacaf7b0dfc166656678c7ffe6/GBIF-2015.png
- p.4/15:
  - https://opendatakit.org/
- p.6:
  - https://openclipart.org
  - A. Bowser, P. Brenton, R. Stevenson, G. Newman, S. Schade, L. Bastin, A. Parker, J. Oliver: Citizen Science Association Data & Metadata Working Group: Report from CSA 2017 and Future Outlook, Wilson Center 2017
  - Ingo Simonis, Rob Atkinson: OGC Discussion Paper Standardized Information Models to Optimize Exchange, Reusability and Comparability of Citizen Science Data, OGC 16-129, http://www.opengis.net/doc/DP/16-129, 2016
  - http://www.opengeospatial.org/pub/www/files/OGC_Logo_2D_Blue_x_0_0.png
- p.8:
  - https://www.w3.org/TR/vocab-ssn, Fig.22
- p.11:
  - https://github.com/HajoRijgersberg/OM/blob/master/images/OM2.0-UML-diagram.png
- p.14:
  - https://www.clipartsfree.net

(3) Flexible Data Export via Templates

time





namespaces.ttl

toplevel.ttl

column.ttl . . . column.ttl

row.ttl

cell.ttl

booleanCell.ttl | dateCell.ttl | decimalCell.ttl

geoshapeCell.ttl | . . . | multipleChoiceCell.ttl

select1Cell.ttl | string.ttl | timeCell.ttl

. . .

row.ttl

.
.
.

(3) Flexible Data Export via Templates

namespaces.ttl.mustache

```
@BASE <{{base}}> .
@PREFIX : <{{base}}> .
{{#namespaces}}
    @PREFIX {{prefix}}: <{{uri}}> .
{{/namespaces}}
@PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@PREFIX owl: <http://www.w3.org/2002/07/owl#> .
@PREFIX oboe-core: <http://ecoinformatics.org/oboe/oboe.1.2/oboe-core.owl#> .
...
```

toplevel.ttl.mustache

```
{{toplevelEntityIdentifier}} rdf:type oboe-core:ObservationCollection .
```

(3) Flexible Data Export via Templates

column.ttl.mustache

```
{{columnEntityIdentifier}} rdf:type oboe-core:Observation .
{{topLevelModel.toplevelEntityIdentifier}} oboe-core:hasMember {{columnEntityIdentifier}} .
```

cell.ttl.mustache

```
{{cellEntityIdentifier}}
      rdf:type             oboe-core:Measurement ;
      oboe-core:hasValue   :measuredValue_{{rowModel.rowId}}_{{columnModel.columnHeader}} .
      om:hasUnit           <{{{semantics.Unit.value}}}> .
```

decimalCell.ttl.mustache

```
:measuredValue_{{rowModel.rowId}}_{{columnModel.columnHeader}}
      oboe-core:hasCode "{{{cellValue}}}"^^xsd:decimal .
```

Exported Data

```
:measurement_1_circumference rdf:type oboe-core:Measurement .
:column_circumference oboe-core:hasMeasurement :measurement_1_circumference .
:measuredValue_1_circumference rdf:type oboe-core:MeasuredValue .
:measurement_1_circumference oboe-core:hasValue :measuredValue_1_circumference .

:measurement_1_circumference dcterms:creator :creator_circumference_1 .
:creator_circumference_1 rdf:type foaf:Person .
:creator_circumference_1 foaf:name "Markus" .

:measurement_1_circumference oboe-core:ofCharacteristic <http://example.org/circumference> .
:measurement_1_circumference
         om:hasUnit <http://www.ontology-of-units-of-measure.org/resource/om-2/centimetre> .

:measuredValue_1_circumference oboe-core:hasCode "156.0000000000"^^xsd:decimal .
```

Template Configuration

The following snippets show a parts of a template configuration:

```yaml
displayName: Extensible Observation Ontology
templateProperties:
  optionalProperties:
    - Creator
  requiredProperties:
    - Characteristic
    - Unit
```

```yaml
availableProperties:
  Creator:
    Endpoint:
    Query:
  Unit:
    Endpoint: http://192.168.0.8:7200/repositories/om
    Query: |-
      PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
      PREFIX : <http://ecoinformatics.org/oboe/oboe.1.2/oboe.owl#>
      PREFIX om: <http://www.ontology-of-units-of-measure.org/resource/om-2/>
      PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

      SELECT DISTINCT ?uri ?displayName
      WHERE {
          ?uri rdf:type om:Unit .
          OPTIONAL{
              ?uri rdfs:label ?displayName .
              FILTER (lang(?displayName) = 'en')
          }
      }
```