The xEBR Ontology
Transforming the XBRL Europe Business Registers Taxonomy into an OWL Ontology: How & What For

Hans-Ulrich Krieger & Thierry Declerck
German Research Center For Artificial Intelligence (DFKI)
Overview

- Background, Goals & Benefits
- xEBR Ontology
- Showcases
Background: The Monnet & TrendMiner Projects

- European Integrated Projects
  - Monnet (www.monnet-project.eu)
  - TrendMiner (www.trendminer-project.eu)
- information extraction from
  - stock exchange http pages
    - Deutsche Börse, NYSE Euronext, Bolsa de Madrid
  - machine-readable financial documents
    - financial reports from different XBRL jurisdictions
- record extracted data in a Semantic Repository
  - transform data into a meaningful form w.r.t. ontology
  - use well-defined standards (e.g., XSD, RDF, OWL)
  - carry out inferences and queries
    - to make implicit knowledge explicit
    - to unveil inconsistencies
- xEBR ontology mediates between different XBRL jurisdictions
xEBR—What is This?

- international standard for financial reporting: XBRL
  - XBRL = eXtensible Business Reporting Language
  - free and open standard
  - makes use of usual XML technologies
  - describes important financial information for one full year

- structure of XBRL document is the same for each jurisdiction
- **but**
  - namespace and tag names differ
  - some of the tags in one jurisdiction do not exist in another one

- circumvent these issues (in part): xEBR
  - xEBR = XBRL Europe Business Registers
  - XBRL working group on xEBR
  - core taxonomy (≈ interlingua) for all XBRL jurisdictions
  - information that is shared by all jurisdictions
  - canonical names
  - further info: other key financial ratios, company identification
Background: Schema—The Monnet Financial Ontology

- green: industry sector classification
- orange: financial reporting
- brown: stock exchange master data
- red: interface axioms
- blue: financial instruments
Background: Instance Data—Company Snapshots

- company information is harvested from stock exchange pages
- caption-value pairs are translated into instance data compliant with the DAX & EN ontology schemata
- this includes
  - the proper use of class and property names
  - the introduction of fresh URIs for individuals of interest
  - the syntactical transformation of values
  - the combination of values from different places
- binary facts + temporal duration: quintuples

<table>
<thead>
<tr>
<th>company identification</th>
<th>property type</th>
<th>property value</th>
</tr>
</thead>
<tbody>
<tr>
<td>DE000A1EWWW0_1324055698213</td>
<td>dax:endOfYear</td>
<td>&quot;2011-12-31&quot;^^xsd:gMonthDay</td>
</tr>
<tr>
<td>DE000A1EWWW0_1324055698213</td>
<td>dax:totalCapitalStock</td>
<td>&quot;209216186EUR&quot;^^xsd:monetary</td>
</tr>
<tr>
<td>DE000A1EWWW0_1324055698213</td>
<td>dateTime</td>
<td>&quot;2011-12-21T16:24:20&quot;</td>
</tr>
<tr>
<td>DE000A1EWWW0_1324055698213</td>
<td>dateTime</td>
<td>&quot;2011-12-21T16:24:20&quot;</td>
</tr>
</tbody>
</table>
Background: Instance Data—xEBR Reports

- financial facts are harvested from XBRL reports
- xEBR serves as a mediator between XBRL jurisdictions
- XML-to-N-Triples translation

```xml
<pfs:Assets contextRef="CurrentInstant" decimals="INF" unitRef="U-EUR">128413</pfs:Assets>
```

```xml
rep42 rdf:type xebr:Report
derf:hasKeyBalanceSheetFiguresReport kbsfr4711
```

```
XEBR2XBRL & IF OWL axioms
xebr:AssetsPresentation ≡ xbrl_be:AssetsTitle
xebr:AssetsPresentation ≡ xbrl_es:ActivoPresentacion
xebr:hasAssetsTotal ≡ xbrl_be:hasAssets
xebr:hasAssetsTotal ≡ xbrl_es:hasTotalActivo
en:revenue ≡ xebr:hasNetTurnover
en:operatingProfit ≡ xebr:hasOperatingProfitLossTotal
```
Using xEBR: Benefits

- the representation of the xEBR taxonomy as an OWL ontology enables the use of semantic technologies
- short-time varying stock exchange data can be merged with long(er)-time static data found in XBRL reports
- data redundancy permits consistency checking and the generation of financial summaries
- the use of xEBR (and industry sector classification) makes it possible to compare companies across XBRL jurisdictions
- multilingual labels (e.g., English, French, German) attached to xEBR concepts ease the work of human analysts
<xsd:element name="KeyBalanceSheetFiguresReport" id="xebr_KeyBalanceSheetFiguresReport"
type="xbrli:stringItemType" substitutionGroup="xbrli:item" abstract="true" nillable="true"
xbri:periodType="instant"/>
<xsd:element name="AssetsPresentation" id="bach_AssetsPresentation" type="xbrli:stringItemType"
substitutionGroup="xbrli:item" abstract="true" nillable="true" xbrli:periodType="instant"/>
<xsd:element name="SubscribedCapitalUnpaid" id="bach_SubscribedCapitalUnpaid"
type="xbrli:monetaryItemType" substitutionGroup="xbrli:item" nillable="true"
xbri:balance="debit" xbrli:periodType="instant"/>
<xsd:element name="FixedAssetsPresentation" id="xebr_FixedAssetsPresentation"
type="xbrli:stringItemType" substitutionGroup="xbrli:item" abstract="true" nillable="true"
xbri:periodType="instant"/>
<xsd:element name="IntangibleFixedAssetsPresentation" id="bach_IntangibleFixedAssetsPresentation"
xbri:periodType="instant"/>
<xsd:element name="FormationPreliminaryExpenses" id="bach_FormationPreliminaryExpenses"
type="xbrli:monetaryItemType" substitutionGroup="xbrli:item" nillable="true"
xbri:balance="debit" xbrli:periodType="instant"/>
<xsd:element name="OtherIntangibleFixedAssets" id="bach_OtherIntangibleFixedAssets"
type="xbrli:monetaryItemType" substitutionGroup="xbrli:item" nillable="true"
xbri:balance="debit" xbrli:periodType="instant"/>
<xsd:element name="IntangibleFixedAssetsTotal" id="bach_IntangibleFixedAssetsTotal"
type="xbrli:monetaryItemType" substitutionGroup="xbrli:item" nillable="true"
xbri:balance="debit" xbrli:periodType="instant"/>
<xsd:element name="TangibleFixedAssetsPresentation" id="bach_TangibleFixedAssetsPresentation"
xbri:periodType="instant"/>
.........
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Assets [Presentation]</td>
<td>Subscribed capital unpaid [Presentation]</td>
<td>abstract</td>
</tr>
<tr>
<td>Subscribed Capital Unpaid [Presentation]</td>
<td>Subscribed capital unpaid [Total]</td>
<td>monetary</td>
</tr>
<tr>
<td>Intangible Fixed Assets [Presentation]</td>
<td>Intangible fixed assets [Presentation]</td>
<td>abstract</td>
</tr>
<tr>
<td>Formation Preliminary Expenses</td>
<td>Formation (preliminary) expenses</td>
<td>monetary</td>
</tr>
<tr>
<td>Other Intangible Fixed Assets</td>
<td>Other intangible fixed assets</td>
<td>monetary</td>
</tr>
<tr>
<td>Intangible Fixed Assets [Total]</td>
<td>Intangible fixed assets [Total]</td>
<td>monetary</td>
</tr>
<tr>
<td>Tangible Fixed Assets [Presentation]</td>
<td>Tangible fixed assets [Presentation]</td>
<td>abstract</td>
</tr>
<tr>
<td>Property Plant And Equipment [Presentation]</td>
<td>Property, plant, and equipment [Presentation]</td>
<td>abstract</td>
</tr>
<tr>
<td>Land And Buildings</td>
<td>Land and buildings</td>
<td>monetary</td>
</tr>
<tr>
<td>Plant And Machinery</td>
<td>Plant and machinery</td>
<td>monetary</td>
</tr>
<tr>
<td>Furniture Fixtures And Equipment</td>
<td>Furniture, fixtures, and equipment</td>
<td>monetary</td>
</tr>
<tr>
<td>Other Property Plant And Equipment</td>
<td>Other property, plant, and equipment</td>
<td>monetary</td>
</tr>
<tr>
<td>Other Fixtures</td>
<td>Other fixtures</td>
<td>monetary</td>
</tr>
<tr>
<td>Property Plant And Equipment [Total]</td>
<td>Property, plant, and equipment [Total]</td>
<td>monetary</td>
</tr>
<tr>
<td>Payments On Account And Assets In Construction</td>
<td>Payments on account and assets in construction</td>
<td>monetary</td>
</tr>
<tr>
<td>Other Tangible Fixed Assets</td>
<td>Other tangible fixed assets</td>
<td>monetary</td>
</tr>
<tr>
<td>Tangible Fixed Assets [Total]</td>
<td>Tangible fixed assets [Total]</td>
<td>monetary</td>
</tr>
<tr>
<td>Affiliated Enterprises</td>
<td>Affiliated enterprises</td>
<td>monetary</td>
</tr>
<tr>
<td>Other Financial Assets</td>
<td>Other financial assets</td>
<td>monetary</td>
</tr>
<tr>
<td>Financial Fixed Assets [Total]</td>
<td>Financial fixed assets [Total]</td>
<td>monetary</td>
</tr>
<tr>
<td>Fixed Assets [Total]</td>
<td>Fixed assets [Total]</td>
<td>monetary</td>
</tr>
</tbody>
</table>
xEBR: From a Taxonomy to an Ontology

- specification: make use of W3C standards RDF & OWL
  - declarative formalism
  - precise semantics
  - computational properties: decidable, terminating
- taxonomy is a tree, expressing meronymy (not hyponymy !)
- every xEBR concept in the specification comes with
  - tag (xEBR concept): OWL class and/or property
  - (multilingual) label: annotation property rdfs:label
  - xEBR type
    - role: 4 fundamental parts of an xEBR report; example: *Key Balance Sheet Figures [Report]*
    - abstract: a presentation has parts; example: *Assets [Presentation]*
    - tuple: like abstract, but parts might be instantiated multiple times; example: *Company Official [List]*
    - monetary, string, date, etc.: leaves of the taxonomy
**xEBR: From a Taxonomy to an Ontology—Idea**

- **whole** (indent $i$) $\supset$ **part** (indent $i + 1$): xEBR types
  - role $\supset$ abstract $|$ tuple
  - abstract $|$ tuple $\supset$ abstract $|$ tuple
  - abstract $|$ tuple $\supset$ atomic type
  - **no** other combinations possible, i.e., atomic type $\not\supset$ abstract $|$ tuple

- transformation to OWL
  - role, abstract, tuple: OWL classes
  - xEBR atomic types: XSD types
  - parthood ($\supset$): dedicated OWL properties

- `<whole> has<part> <part>`
  - ... $\supset$ **abstract**: functional *object* property
  - ... $\supset$ **tuple**: (relational) *object* property
  - ... $\supset$ **atomic type**: functional *datatype* property
xEBR: From a Taxonomy to an Ontology—Example

- ontology schema is auto-generated by a Java program
- N-Triples file is transformed into OWL-XML using raptor
- result can be viewed and modified using ontology editors
- example: $\text{Assets [Presentation]} \supseteq \text{Fixed Assets [Presentation]}$

\begin{verbatim}
xeb:AssetsPresentation rdf:type owl:Class
xeb:AssetsPresentation rdfs:label "Assets [Presentation]"@en
xeb:AssetsPresentation rdfs:subClassOf xeb:Abstract
xeb:hasFixedAssetsPresentation rdf:type owl:FunctionalProperty
xeb:hasFixedAssetsPresentation rdf:type owl:ObjectProperty
xeb:hasFixedAssetsPresentation rdfs:domain xeb:AssetsPresentation
xeb:hasFixedAssetsPresentation rdfs:range xeb:FixedAssetsPresentation

xeb:FixedAssetsPresentation rdf:type owl:Class
xeb:FixedAssetsPresentation rdfs:label "Fixed Assets [Presentation]"@en
xeb:FixedAssetsPresentation rdfs:subClassOf xeb:Abstract
\end{verbatim}
xEBR: Version 7 OWL Ontology—Protégé View
xEBR: Version 7 OWL Ontology—Protégé View, Cont.
xEBR: From a Taxonomy to an Ontology—Alternative

- meronymy can **not** be expressed *intensionally* for all reports
  - OWL’s TBox axiom constructors: $\equiv$ and $\sqsubseteq$
  - geared towards hyponymy
  - no constructor for parthood $\ni$
  - our solution: individual properties for part-whole pairs
  - **but:** introduces a great number of (possibly empty) container objects over and over again during ontology population

- solution: meta-model $\ni$ as a TBox axiom constructor
  - $\ni \subseteq$ owl:Class $\times$ owl:Class
  - no longer OWL-DL, but still RDF, accessible through SPARQL
  - requires property `belongsTo` to link concrete data (e.g., financial numbers) to a specific xEBR report
  - relational properties (tuple !) can **not** be modeled properly
xEBR: Problems With Hyponymy

- Problem: change of xEBR type; true example: role → tuple → abstract → tuple → string ??
- Hyponymy would lead to the “inheritance” of properties
- hasTangibleFixedAssets defined on FixedAssets is valid for TangibleFixedAssets and PropertyPlantAndEquipment, given PropPlantEquip ⊑ TangibleFixedAssets ⊑ FixedAssets
- Solution: local cardinality (= 0) restriction defined on classes
- More serious problem: entailment (universal instantiation) + querying (SPARQL): more than 1 value
- SELECT ?t WHERE {?p rdf:type xebr:FixedAssets . ?p xebr:hasTotal ?t}
Showcases

showcases make use of *HFC*, an in-memory semantic repository and extended forward chainer, going beyond RDF & OWL

- connecting snapshots and reports
- multilingual properties and classes
- evolution of financial data
- finding competitors across borders
- financial summaries: filling missing data
- consistency checking
snapshots (quintuples) and reports (triples) are not related when they are uploaded; this rule connects them through if:hasReport rules and queries below depend on the application of this rule

```sparql
?comp rdf:type dax:Company ?s1 ?e1
?comp dax:isin ?isin ?s1 ?e1
?rep rdf:type xebr:Report
?rep xebr:hasCompanyIdentificationReport ?idrep
?idrep xebr:hasCompanyIdNumberList ?idlist
?idlist xebr:hasCompanyIdValue ?isin
?rep xebr:starts ?s2
?rep xebr:ends ?e2
->
?comp if:hasReport ?rep ?s2 ?e2
```
interface axioms and multilingual labels can ease the work of a human analyst in a specialized GUI (not shown here, only query)

xebr:hasAssetsTotal ≡ xbrl_be:hasAssets
xebr:hasAssetsTotal ≡ xbrl_es:hasTotalActivo
xebr:hasKeyBalanceSheetFiguresReport ≡ xbrl_es:hasBalanceSituacionPresentacion
xebr:KeyBalanceSheetFiguresReport ≡ xbrl_es:BalanceSituacionPresentacion
xebr:AssetsPresentation ≡ xbrl_es:ActivoPresentacion

xbrl_be:hasAssets rdfs:label "Summe der Aktiva"de
xbrl_be:hasAssets rdfs:label "Total assets"en
xbrl_be:hasAssets rdfs:label "Total de l’actif"fr
xbrl_be:hasAssets rdfs:label "Totaal van de activa"nl
xbrl_es:hasTotalActivo rdfs:label "Total activo"es
xebr:hasAssetsTotal rdfs:label "Assets [Total]"en
Spanish analyst looking for FDC’s total assets for 2007 (start of business year can be obtained elsewhere, viz., stock exchange)

SELECT ?total_assets
WHERE ?rep rdf:type xebr:Report

  ?rep xebr:hasCompanyIdentificationReport ?idrep
  ?idrep xebr:hasCompanyIdNumberList ?idlist
  ?idlist xebr:hasCompanyIdValue "BE0468567012"
  ?rep xebr:starts "2007-04-01"^^xsd:date

?rep xbrl_es:hasBalanceSituacionPresentacion ?kbsf
?kbsf xbrl_es:hasActivoPresentacion ?ap
?ap xbrl_es:hasTotalActivo ?total_assets
query tangible & intangible fixed assets for *adidas* over the years (not listed on German DAX pages)

```
SELECT ?start2 ?end2 ?tass ?iass
WHERE ?comp rdf:type dax:Company ?start ?end
  ?comp dax:isin "DE000A1EWWW0" ?start ?end
  ?rep xebr:hasKeyBalanceSheetFiguresReport ?kbsf
  ?kbsf xebr:hasAssetsPresentation ?ap
  ?ap xebr:hasFixedAssetsPresentation ?fap
  ?fap xebr:hasTangibleFixedAssetsPresentation ?tfap
  ?tfap xebr:hasTangibleFixedAssetsTotal ?tass
  ?fap xebr:hasIntangibleFixedAssetsPresentation ?ifap
  ?ifap xebr:hasIntangibleFixedAssetsTotal ?iass
```
Showcase: Finding Competitors Across Borders

we are looking for competitors of Deutsche Bank and their financial numbers (xEBR reports), both on DAX & Euronext

IF interface axiom (DAX ↔ EN): financial institutions

icb:ICB8300 ≡ dax:Banks ≡ nace:nace_64.1

SELECT DISTINCT ?rival ?rep ?s3 ?e3
WHERE ?db dax:name "Deutsche Bank" ?s ?e
 ?rival rdf:type ?type ?s2 ?e2
 ?rival if:hasReport ?rep ?s3 ?e3
FILTER ?db != ?rival
Showcase: Financial Summaries—Filling Missing Data

Information that should be listed on DAX/EN is missing, but is available in the xEBR report for the same company (vice versa)

Example: EADS’ *net profit* for 2007-2012 on Euronext

```
?rep rdf:type xebr:Report
?rep xebr:hasCompanyIdentificationReport ?idrep
?idrep xebr:hasCompanyIdNumberList ?idlist
?idlist xebr:hasCompanyIdValue "NL0000235190"
?rep xebr:starts ?start
?rep xebr:ends ?end
?rep xebr:hasKeyBalanceSheetFiguresReport ?kbsf
?kbsf xebr:hasEquityAndLiabilitiesPresentation ?elp
?elp xebr:hasEquityPresentation ?ep
?ep xebr:hasProfitLossForThePeriodTotal ?prof
->
```
consistency checking through rules implements conditions that should never happen (like LTL safety properties):

- leave a *memento* in the repository when things go wrong
- by marking URIs as being inconsistent (assign bottom type)

```prolog
?p rdf:type owl:FunctionalProperty
?p rdf:type owl:DatatypeProperty
?x ?p ?y ?s1 ?e1
?x ?p ?z ?s2 ?e2
->
?x rdf:type owl:Nothing ?s ?e
@test
?y != ?z
IntersectionNotEmpty ?s1 ?e1 ?s2 ?e2
@action
?s = Max2 ?s1 ?s2
?e = Min2 ?e1 ?e2
```
Thank you!

Questions?


The xEBR Ontology: Transforming the XBRL Europe Business Register Taxonomy into an OWL Ontology. XBRL 2013.