

Universität
Rostock



Traditio et Innovatio

XML to Graph Mapping Tool

Bachelorarbeit Präsentation

Safwat Zakkor

217205264

Betreuer:

Dr. Holger Meyer

Holger Meyer (PhD), U of Rostock, CSEE, Database Research Group

Rostock, 29.03.2022

Inhalt

DATUM UND UHRZEIT:
29. MÄRZ.2022, 15:00 UHR

Via Zoom

1

Einleitung

- Was ist das Ziel dieser Arbeit

2

Übersicht

- Allgemeine Übersicht
- Motivation
- Modulstruktur

3

Filter

- Die verschiedenen Varianten von XML-Dokumente
- Beispiele XML
- XML-Schema
- XPath in XML-Dokumente Funktionsweise

4

Graph Generator

- Generator
- Schritt für Schritt wie ein Graph erstellen wird
- Match (Find)
- CreateNode
- CreateEdge
- Node replacement cases
 - Problemdarstellung
 - Lösungsvarianten
 - Beispiele

5

Exporter

- Die Formate
- Exporter als
 - Node.csv
 - Edge.csv
- Andere Vorschläge
 - Gexf
 - Graphml



6

Ergebnisse

- Ausgabe des generierten Graphen

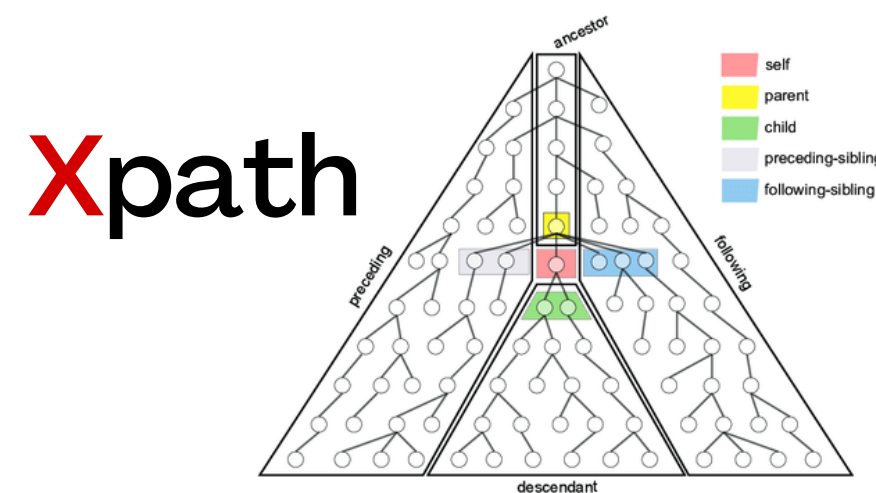
7

Zusammenfassung und Ausblick

- Was wurde geschafft? 
- Welche Teile wurden implementiert?
- Teile die nicht gemacht wurden? 
- Erweiterungen
- Ausgabeformat

Das Ziel

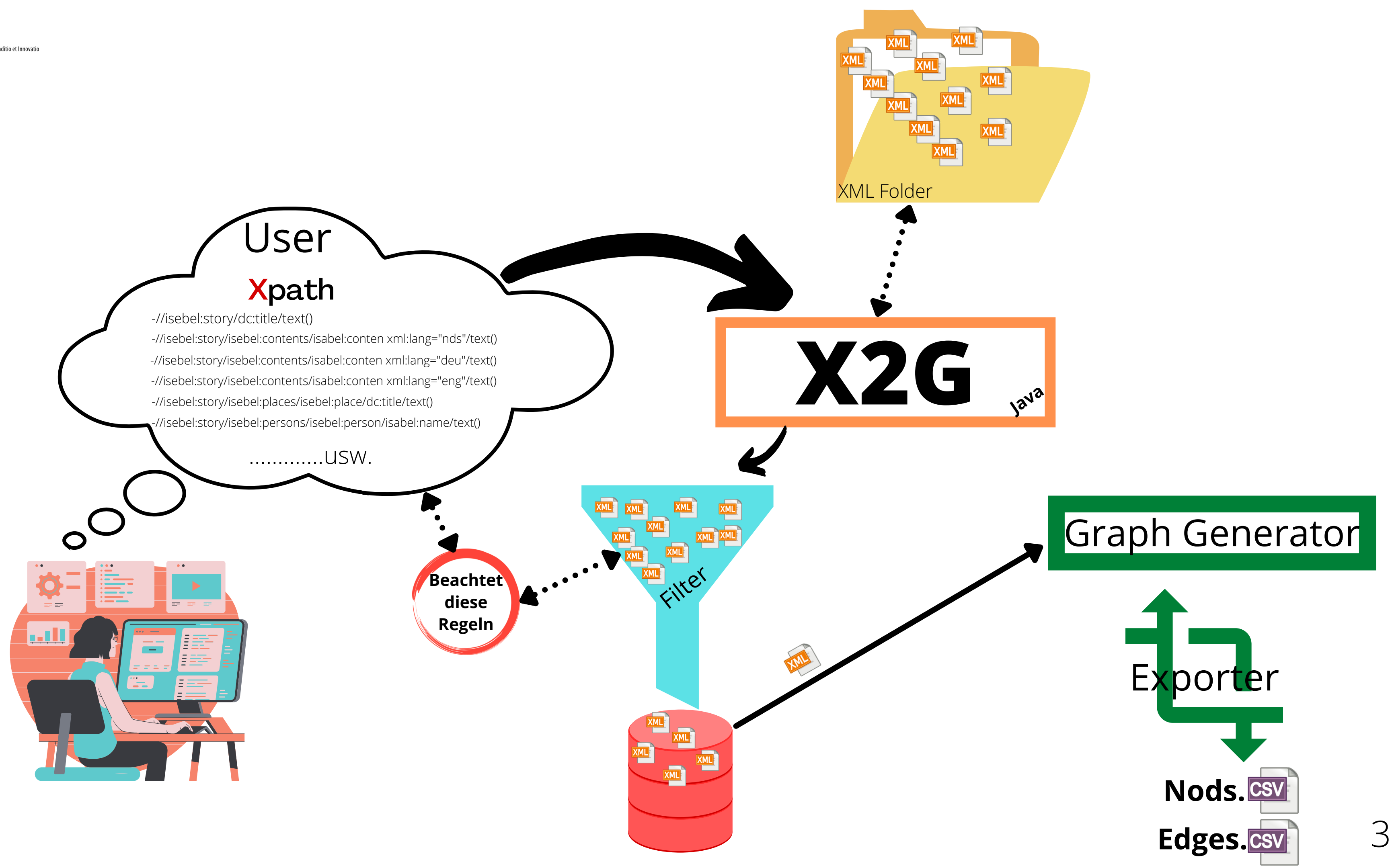
- Entwicklung und Konzeption einer Anwendung zur Ermöglichung einer Analyse und Visualisierung für XML-Dokumenten in Property Graph model
- Prototypische Implementierung und qualitätsbasierte Evaluierung anhand von ISEBEL-Beispielszenarien
- Die Transformation sollte von benutzerdefinierten Regeln geleitet werden
- Ausgabe des generierten Graphen in verschiedenen Formaten



Gephi (Beispiel)

X2G beantwortet die folgenden Fragen

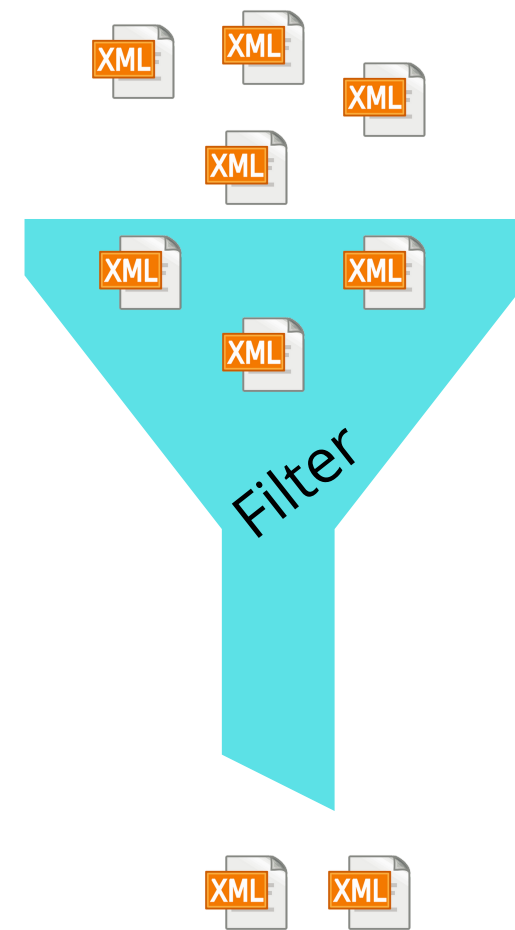
- Dokumentstruktur wird nicht 1-zu-1 in Graphstruktur überführt
- Welche XML-Dokumente werden bearbeitet
- Die Property-Graphstruktur: Knoten und Kanten mit Eigenschaften
- Bestimmte Regeln für die Erstellung von Knoten und Kanten
- Das Format der Ausgabe der generierten Graphen




Wie funktioniert der Filter?

und warum brauchen wir einen Filter?

Filter verwendet XPath-Ausdrücke, um jede XML-Datei zu überprüfen, ob sie die Kriterien (Regeln) des Benutzers erfüllt.



Beispiele --> nächste Folie 

```

<isebel:story xmlns:isebel="http://www.isebel.eu/ns/isebel" xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:html="http://www.w3.org/1999/xhtml" xml:id="xmd-s001-000-000-030" xml:lang="deu">
  <dc:identifier>de.wossidia.xmd-s001-000-000-030</dc:identifier>
  <isebel:purl>https://apps.wossidia.de/webapp/run/node/XMD-S001-000-000-030</isebel:purl>
  <dc:title>[30] Der gefundene Schatz macht standesgemäß</dc:title>
  <dc:title xml:lang="nds"/>
  <dc:title xml:lang="deu"/>
  <dc:title xml:lang="eng"/>
  <isebel:contents>
    <isebel:content xml:lang="nds">Kleekamp willen poor Daglöhner Holt stietzen. 'Dat' is 'noch tiedigen' Morgen. 'Se kamen an 'ne' Kuul mit Steen 'vörbi'. Een mööt ut de Hosen. 'He' will 'solang'! Dorvon het de Herr en Stuuv bugen laten in't Herrenhuus. Erzähler: H. Müller, Hohen Viechein, Kreis Wismar; Aufzeichner: Wossidlo, 21. 5. 1932.</isebel:content>
    <isebel:content xml:lang="deu">In Kleekamp wollen ein paar Tagelöhner Holz stehlen. Es ist noch sehr zeitig am Morgen. Sie kommen an einer Steingrube vorbei. Einer von ihnen muß mal in die altmodisch gewesen. Er hat es aber genommen und umgewechselt. Davon hat ihm der Herr im Herrenhaus eine Stube bauen lassen.</isebel:content>
    <isebel:content xml:lang="eng">A few daytaller wanted to steal some wood. It was very early in the morning. They passed a stone pit. One of them had to go in the bushes. Whilst he wanted to From that profit, the squire had a room in the manor built for him.</isebel:content>
  </isebel:contents>
  <isebel:places>
    <isebel:place id="10001174">
      <dc:title>Hohen Viechein</dc:title>
      <isebel:point>
        <datacite:pointLatitude xmlns:datacite="http://datacite.org/schema/kernel-4">53.78404</datacite:pointLatitude>
        <datacite:pointLongitude xmlns:datacite="http://datacite.org/schema/kernel-4">11.51169</datacite:pointLongitude>
      </isebel:point>
      <isebel:role>narration</isebel:role>
    </isebel:place>
  </isebel:places>
  <isebel:events>
    <isebel:event id="990013673">
      <isebel:date>1932-05-21</isebel:date>
      <isebel:role>narration</isebel:role>
    </isebel:event>
  </isebel:events>
  <isebel:keywords>
    <isebel:keyword id="XMD-M030-000-000-640" type="controlled vocabulary" provenance="WossiDiA-Index">Schätze</isebel:keyword>
    <isebel:keyword id="XMD-M030-000-004-718" type="controlled vocabulary" provenance="WossiDiA-Index">finden</isebel:keyword>
    <isebel:keyword id="XMD-M030-000-008-000" type="controlled vocabulary" provenance="WossiDiA-Index">nehmen (untersch. Bed.)</isebel:keyword>
    <isebel:keyword xml:lang="deu" id="ZAW-B802" type="controlled vocabulary" provenance="RW-Index">Schätze</isebel:keyword>
    <isebel:keyword xml:lang="eng" id="ZAW-B802" type="controlled vocabulary" provenance="RW-Index">treasure</isebel:keyword>
    <isebel:keyword xml:lang="deu" id="ZAW-B802-017" type="controlled vocabulary" provenance="RW-Index">Den gefundenen Schatz nimmt ein anderer</isebel:keyword>
    <isebel:keyword xml:lang="eng" id="ZAW-B802-017" type="controlled vocabulary" provenance="RW-Index">The found treasure takes another</isebel:keyword>
  </isebel:keywords>
  <isebel:imageResources>
    <isebel:imageResource id="ZAW-B802-017-009-001">
      <isebel:purl>https://apps.wossidia.de/digipool/3babc7eb/master</isebel:purl>
    </isebel:imageResource>
    <isebel:imageResource id="ZAW-B802-017-009-002">
      <isebel:purl>https://apps.wossidia.de/digipool/3babc7ec/master</isebel:purl>
    </isebel:imageResource>
  </isebel:imageResources>
</isebel:story>

```

```
<isebel:story xmlns:isebel="http://www.isebel.eu/ns/isebel" xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:html="http://www.w3.org/1999/xhtml" xml:id="xmd-s001-000-000-005" xml:lang="deu">
  <dc:identifier>de.wossidia.xmd-s001-000-000-005</dc:identifier>
  <isebel:purl>https://apps.wossidia.de/webapp/run/node/XMD-S001-000-000-005</isebel:purl>
  <dc:title>[5] Edelmann wird totgeschlagen, nachdem er Mätresse misshandelt hat</dc:title>
  <dc:title xml:lang="nds"/>
  <dc:title xml:lang="deu"/>
  <isebel:contents>
    <isebel:content xml:lang="nds">De Eddelmann in Stellshagen het en Diern so mißhannelt, de het toletzt ’n Stoeker nahmen un is up em ingahn. Un de Lüüd hebben em von’n Pierd krägen un dootsl
    de hebben sik flüchtt, de annern sünd faatt. – Gerichts schlag heet dat hüüt noch. As fief all richtt sünd, un de sösst ran sall, kümmt Ordre von Swerin, se süllen bloß enen richten. Dat he
    <isebel:content xml:lang="deu">Der Edelmann in Stellshagen hat ein Mädchen so mißhandelt, die hat zuletzt den Stock genommen und ist auf ihn eingedrungen. Und seine Leute haben ihn vom Pfer
    Wassertümpel, da sind viele hindurchgeschwommen, die haben sich geflüchtet. Die anderen sind gefaßt worden. – Gerichtsschlag heißt es ja noch heute dort. – Als fünf schon hingerichtet sind
    </isebel:content>
  </isebel:contents>
  <isebel:places>
    <isebel:place id="100001174">
      <dc:title>Hohen Viecheln</dc:title>
      <isebel:point>
        <datacite:pointLatitude xmlns:datacite="http://datacite.org/schema/kernel-4">53.78404</datacite:pointLatitude>
        <datacite:pointLongitude xmlns:datacite="http://datacite.org/schema/kernel-4">11.51169</datacite:pointLongitude>
      </isebel:point>
      <isebel:role>narration</isebel:role>
    </isebel:place>
    <isebel:place id="100002701">
      <dc:title>Stellshagen</dc:title>
      <isebel:point>
        <datacite:pointLatitude xmlns:datacite="http://datacite.org/schema/kernel-4">53.94487</datacite:pointLatitude>
        <datacite:pointLongitude xmlns:datacite="http://datacite.org/schema/kernel-4">11.13973</datacite:pointLongitude>
      </isebel:point>
      <isebel:role>action</isebel:role>
    </isebel:place>
  </isebel:places>
  <isebel:persons>
    <isebel:person id="110002742">
      <isebel:name>Hinrichs</isebel:name>
      <isebel:gender>male</isebel:gender>
      <isebel:role>narrator</isebel:role>
      <isebel:profession>Pantoffelmacher</isebel:profession>
    </isebel:person>
  </isebel:persons>
  <isebel:keywords>
    <isebel:keyword id="XMD-M030-000-001-053" type="controlled vocabulary" provenance="WossiDiA-Index">Edelleute</isebel:keyword>
    <isebel:keyword id="XMD-M030-000-005-606" type="controlled vocabulary" provenance="WossiDiA-Index">Haberlandt</isebel:keyword>
    <isebel:keyword id="XMD-M030-000-030-625" type="controlled vocabulary" provenance="WossiDiA-Index">Frevelsage</isebel:keyword>
    <isebel:keyword id="XMD-M030-000-000-023" type="controlled vocabulary" provenance="WossiDiA-Index">Wasser</isebel:keyword>
    <isebel:keyword id="XMD-M030-000-000-595" type="controlled vocabulary" provenance="WossiDiA-Index">Wasserlauf</isebel:keyword>
    <isebel:keyword id="XMD-M030-000-000-782" type="controlled vocabulary" provenance="WossiDiA-Index">Edelmann</isebel:keyword>
    <isebel:keyword id="XMD-M030-000-001-198" type="controlled vocabulary" provenance="WossiDiA-Index">Geld</isebel:keyword>
    <isebel:keyword id="XMD-M030-000-001-526" type="controlled vocabulary" provenance="WossiDiA-Index">totschlagen</isebel:keyword>
    <isebel:keyword id="XMD-M030-000-001-977" type="controlled vocabulary" provenance="WossiDiA-Index">Leute</isebel:keyword>
    <isebel:keyword xml:lang="nds" id="XMD-M030-000-002-413" type="controlled vocabulary" provenance="WossiDiA-Index">Dirn</isebel:keyword>
    <isebel:keyword xml:lang="nds" id="XMD-M030-000-018-226" type="controlled vocabulary" provenance="WossiDiA-Index">Scheid</isebel:keyword>
    <isebel:keyword id="XMD-M030-000-025-606" type="controlled vocabulary" provenance="WossiDiA-Index">Mätress</isebel:keyword>
    <isebel:keyword id="XMD-M030-000-038-232" type="controlled vocabulary" provenance="WossiDiA-Index">misshandeln</isebel:keyword>
    <isebel:keyword id="XMD-M030-000-040-933" type="controlled vocabulary" provenance="WossiDiA-Index">besprechen</isebel:keyword>
    <isebel:keyword xml:lang="deu" id="ZAW-B609" type="controlled vocabulary" provenance="RW-Index">Frevel</isebel:keyword>
    <isebel:keyword xml:lang="eng" id="ZAW-B609" type="controlled vocabulary" provenance="RW-Index">sacrilege</isebel:keyword>
    <isebel:keyword xml:lang="deu" id="ZAW-B609-001" type="controlled vocabulary" provenance="RW-Index">Haberlandt</isebel:keyword>
    <isebel:keyword xml:lang="eng" id="ZAW-B609-001" type="controlled vocabulary" provenance="RW-Index">Haberlandt</isebel:keyword>
  </isebel:keywords>
  <isebel:imageResources>
    <isebel:imageResource id="ZAW-B609-001-012-002">
      <isebel:purl>https://apps.wossidia.de/digipool/3bab5921/master</isebel:purl>
    </isebel:imageResource>
    <isebel:imageResource id="ZAW-B609-001-012-001">
      <isebel:purl>https://apps.wossidia.de/digipool/3bab5920/master</isebel:purl>
    </isebel:imageResource>
  </isebel:imageResources>
</isebel:story>
```

2.XML-Datei aus ISEBEL-Projekt: Stories (Sagen, Legenden, Märchen)


```

▼<isebel:story xmlns:isebel="http://www.isebel.eu/ns/isebel" xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:html="http://www.w3.org/1999/xhtml" xml:id="xmd-s001-000-000-026" xml:lang="deu">
  <dc:identifizier>de.wossidia.xmd-s001-000-000-026</dc:identifizier>
  <isebel:purl>https://apps.wossidia.de/webapp/run/node/XMD-S001-000-000-026</isebel:purl>
  <dc:title>[26] Den Schatz nimmt der Fürst</dc:title>
  <dc:title xml:lang="deu"/>
  <dc:title xml:lang="eng"/>
  ▼<isebel:contents>
    <isebel:content xml:lang="deu">In Brunow kommt ein Jäger spät abends nach Haus. Es ist sehr dunkel, da glimmt nach seiner Meinung ein alter Stamm dicht am Weg, wie er annimmt, von Schäferj
    Da findet er an der Stelle eine große Menge Geld. Dadurch, daß er davon erzählt, soll's ihm später unter den Händen verschwunden sein oder der Fürst soll es zur Hälfte eingezogen haben. Er
    <isebel:content xml:lang="eng">Late in the evening, a hunter arrived at his home in Brunow. It was very dark. There glew, according to him, an old stem close to the way. He assumes that yo
    damage. Then, he found an awful lot of money. Consequently that he told about it, the money should disappeared or the ruler confiscated the half.</isebel:content>
  </isebel:contents>
</isebel:story>

```

3.Beispiel mit deutschen und englischen Text der Sage

```

▼<isebel:story xmlns:isebel="http://www.isebel.eu/ns/isebel" xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:html="http://www.w3.org/1999/xhtml" xml:id="xmd-s001-000-000-016" xml:lang="deu">
  <dc:identifizier>de.wossidia.xmd-s001-000-000-016</dc:identifizier>
  <isebel:purl>https://apps.wossidia.de/webapp/run/node/XMD-S001-000-000-016</isebel:purl>
  <dc:title>[16] Den gefundenen Schatz nimmt der Gutsherr</dc:title>
  <dc:title xml:lang="nds"/>
  <dc:title xml:lang="deu"/>
  <dc:title xml:lang="eng"/>
  ▼<isebel:contents>
    <isebel:content xml:lang="nds">In Walkendörp uppe Braak het en Jung Swien hött. 'Donn het en' Soeg 'n ganzen Schäpel Geld rutbraken. 'Donn seggt' de Herr, 'dat Geld gehüürt em un' het sik dat
    <isebel:content xml:lang="deu">In Walkendorf auf der Brache hat ein Junge Schweine gehütet. Da hat eine Sau einen ganzen Scheffel Geld herausgewühlt. Da sagt der Herr, das Geld gehöre ihm
    <isebel:content xml:lang="eng">A boy had herded pigs on the fallow in Walkendorf. Once, a sow has dug out of the earth a whole Scheffel* money. Then, the squire said the money would be his
  </isebel:contents>
</isebel:story>

```

4.Niederdeutsch/Hochdeutsch/Englisch

```
▼<isebel:story xmlns:isebel="http://www.isebel.eu/ns/isebel" xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:html="http://www.w3.org/1999/xhtml" xml:id="xmd-s001-000-000-218" xml:lang="deu">
  <dc:identifizier>de.wossidia.xmd-s001-000-000-218</dc:identifizier>
  <isebel:purl>https://apps.wossidia.de/webapp/run/node/XMD-S001-000-000-218</isebel:purl>
  <dc:title>[218] Bestrafte Habgier</dc:title>
  <isebel:contents/>
</isebel:story>
```

5.Beispiel

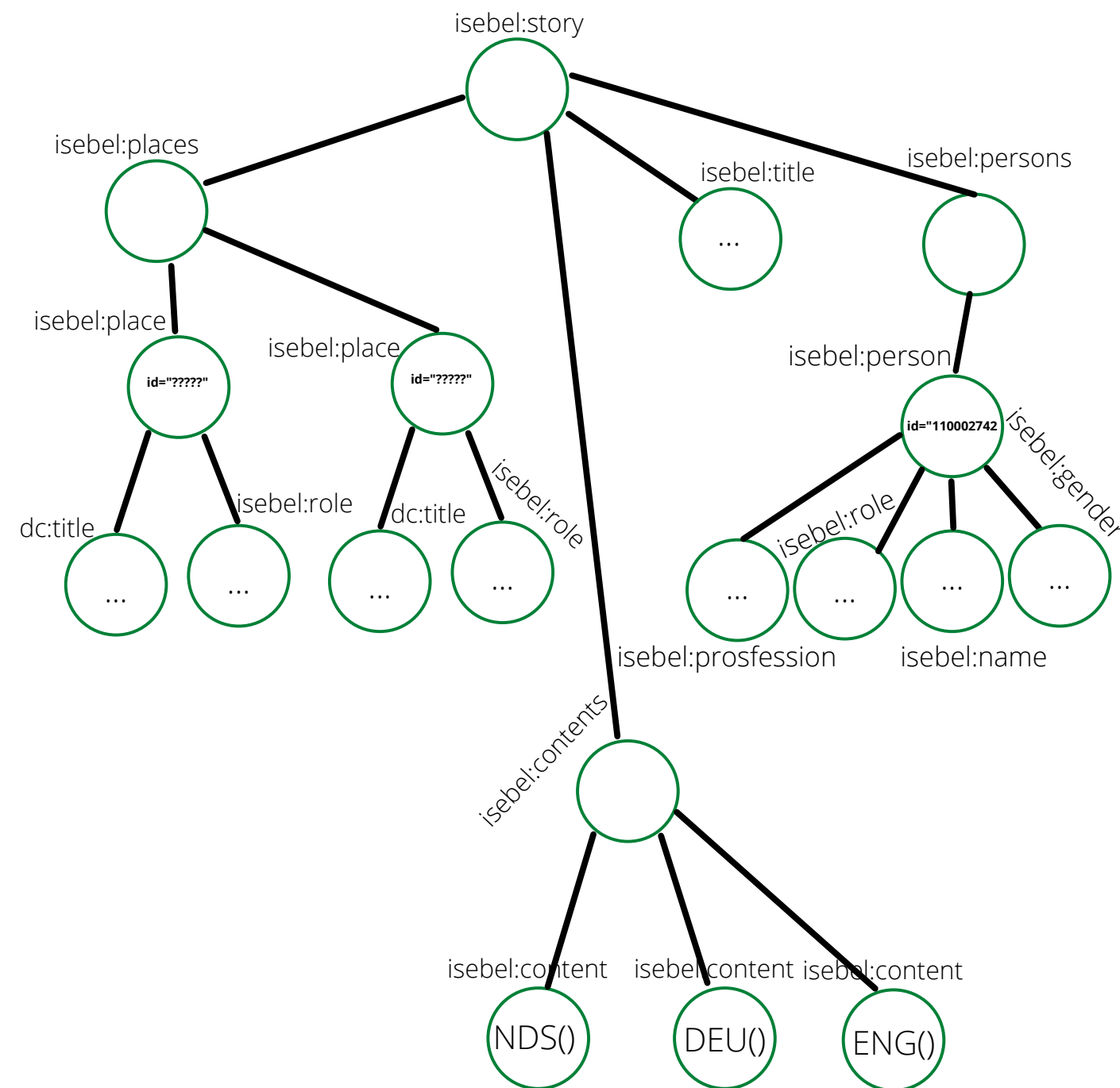
```
▼<isebel:story xmlns:isebel="http://www.isebel.eu/ns/isebel" xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:html="http://www.w3.org/1999/xhtml" xml:id="xmd-s001-000-000-081" xml:lang="deu">
  <dc:identifizier>de.wossidia.xmd-s001-000-000-081</dc:identifizier>
  <isebel:purl>https://apps.wossidia.de/webapp/run/node/XMD-S001-000-000-081</isebel:purl>
  <dc:title>[81] Mord</dc:title>
  <dc:title xml:lang="deu"/>
  <isebel:contents>
    <isebel:content xml:lang="deu">Wo der Weg Moltenow–Bernitt von einem Feldweg gekreuzt wird, dicht hinter Moltenow, steht noch jetzt eine alte Eiche. – Johannistag zwischen 9 und 10 sitzt d
der Vater wollte aber nichts davon wissen. Wenn nun der Vater am Sonntag nachmittag auf die Jagd ging, erhielt sie Besuch von ihrem Geliebten. Doch dieser mußte immer am Moltenower Schäfer
Schäfer der Tat. Der Gutsherr traf ihn, wie er gerade unter der Eiche saß und schlug ihn nieder. Aber er hatte keine Ruhe im Grabe, und am Johannistage, an dem dies geschehen war, kehrte er
B. zurück von der Arbeit. Sch. ist ängstlich, und als sie bei der Eiche vorbeikommen, nimmt er den Hut ab. „Worüm deist du dat?“ fragt der mutige F. „Hest du denn’ Griesen bi de Eik nich se
von der Sache überzeugen, faßt seinen Krückstock fest und geht zur Eiche und sieht einen Mann dort. Als Sch. sich umblickt, sieht er zu seinem Schreck, wie F. hinfällt. Schnell eilt Sch. hi
beleidigten Geistes. Erzähler: Untertertianer Scharf, Penzin, Kreis Bützow; Aufzeichner: Studienrat Barnewitz, Bützow, Kreis Bützow, 1920.</isebel:content>
  </isebel:contents>
</isebel:story>
```

Sage in Hochdeutsch mit direkter Rede in Niederdeutsch

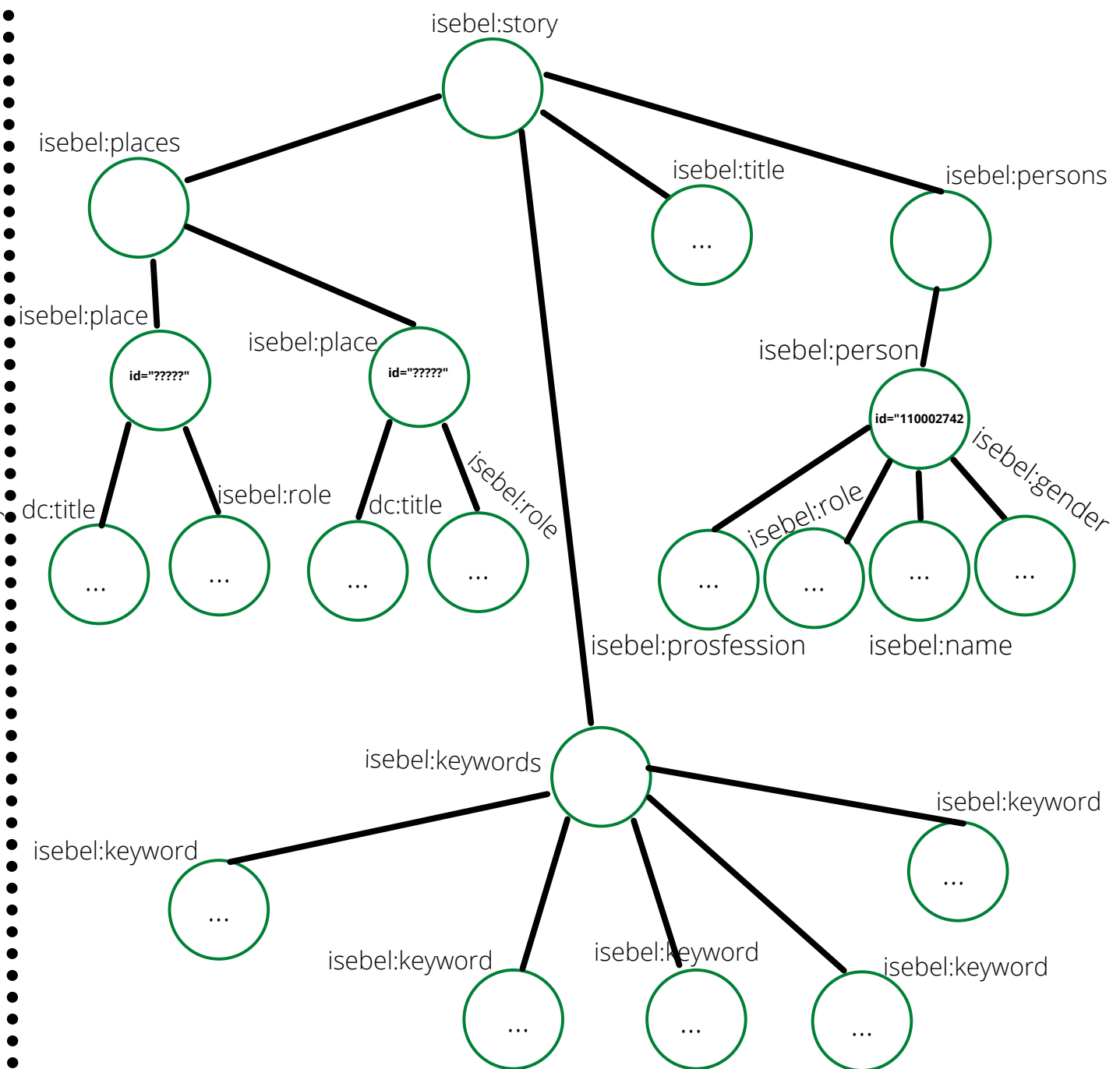
XML als Baumstruktur

- Persons names (Heinrich, Dettmann)
- Persons genders --> (z.B. male)
- Places names --> (Stellshagen, Rostck, ...)

XML File 1



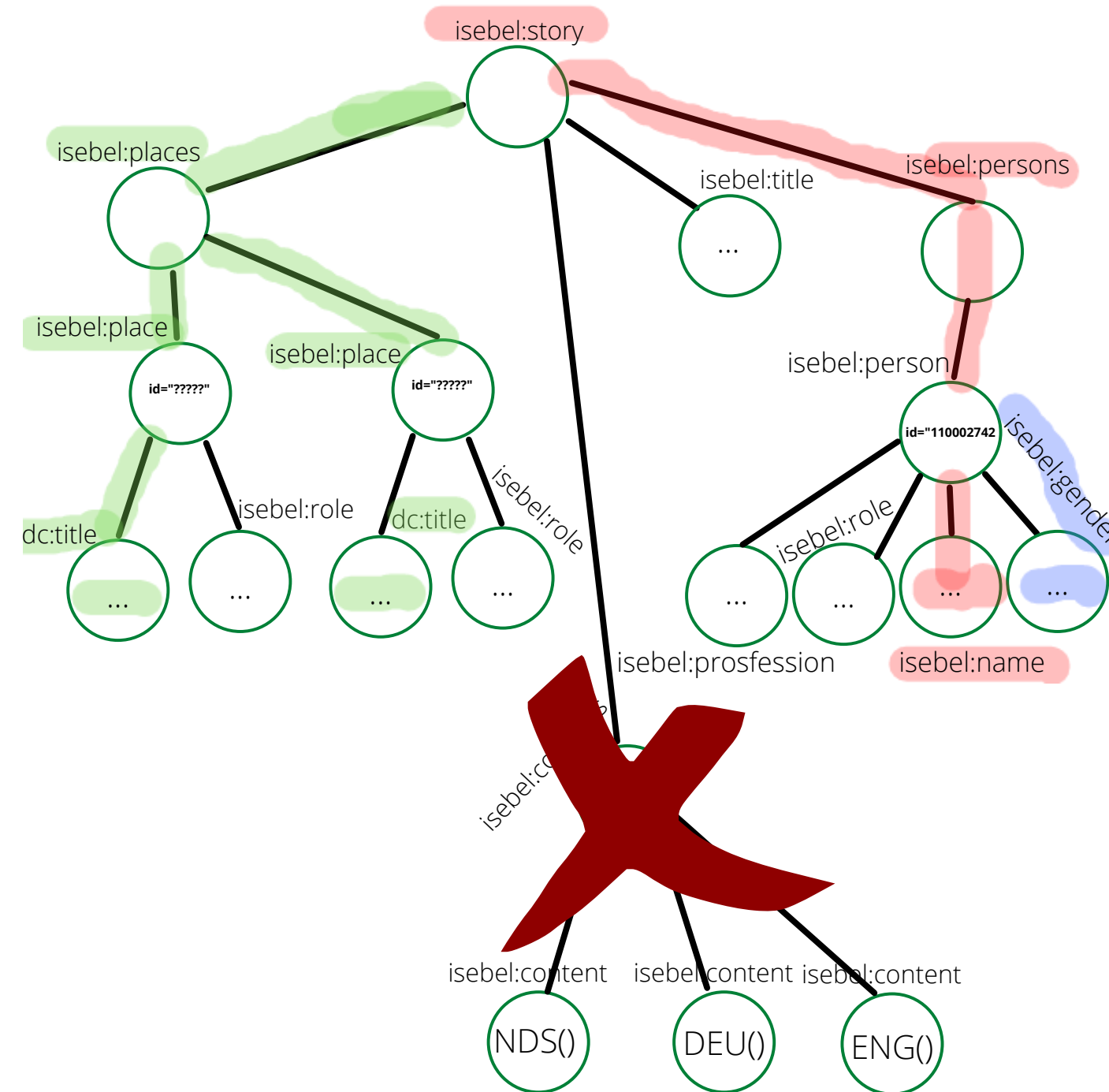
XML File 2



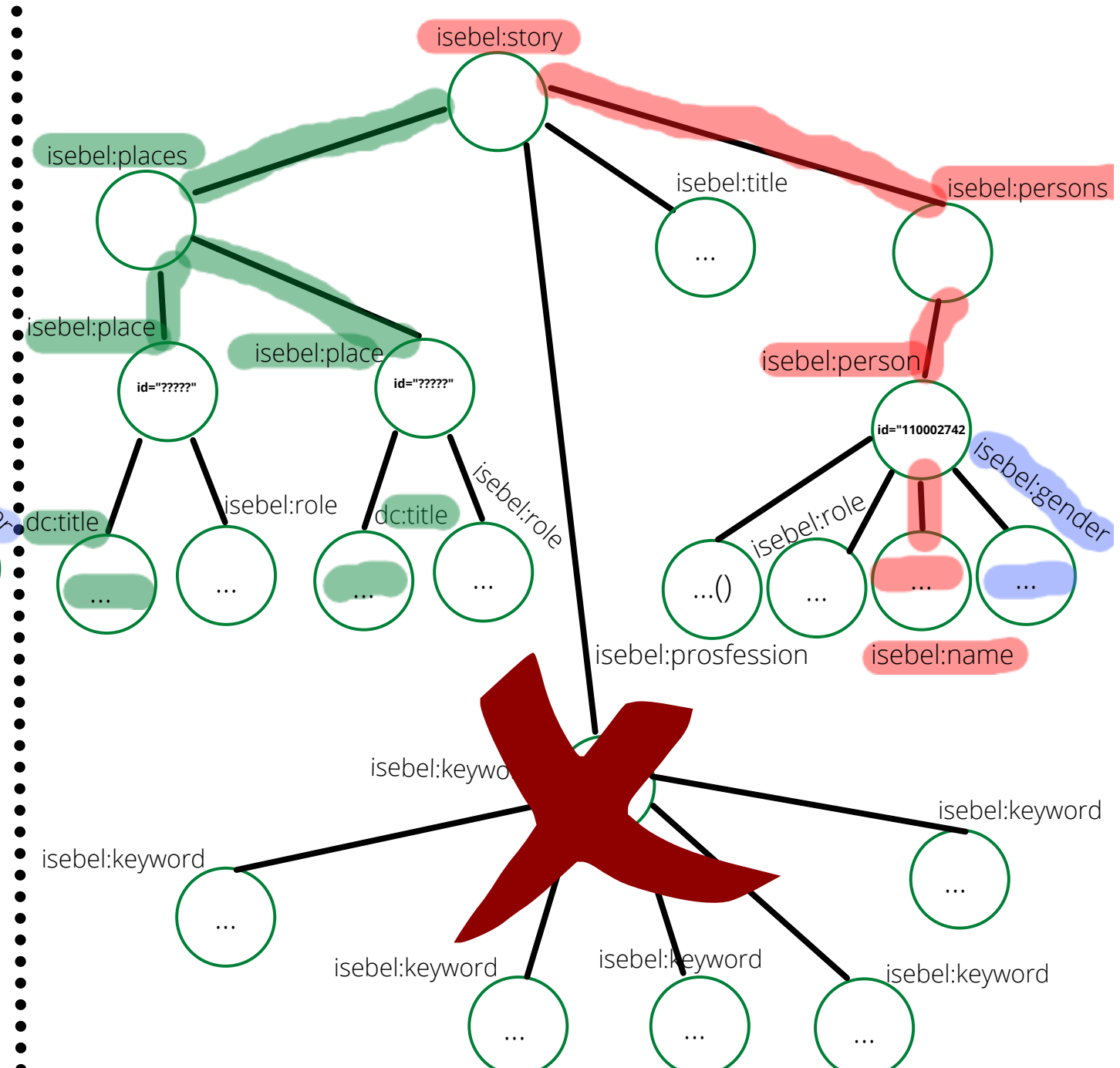
XML als Baumstruktur

`//isebel:story/isebel:persons/isebel:person/isebel:name/text()` Persons names
`//isebel:story/isebel:persons/isebel:person/isebel:gender/text()` Persons genders
`//isebel:story/isebel:places/isebel:place/dc:title/text()` Places names

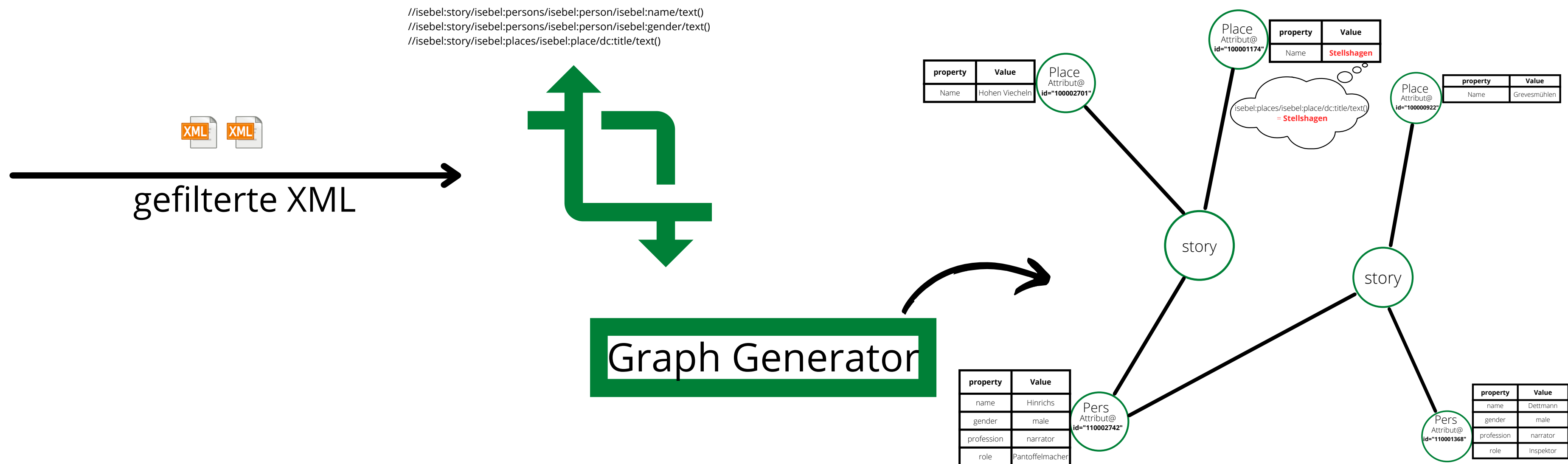
XML File 1



XML File 2



Graph Generator



Graph-Sicht -> nächste Folien

Wir zeigen schrittweise das Generieren des Graphens aus den extrahierten XML-Fragmenten am Beispiel!





create node "story"

create node "place"

Place

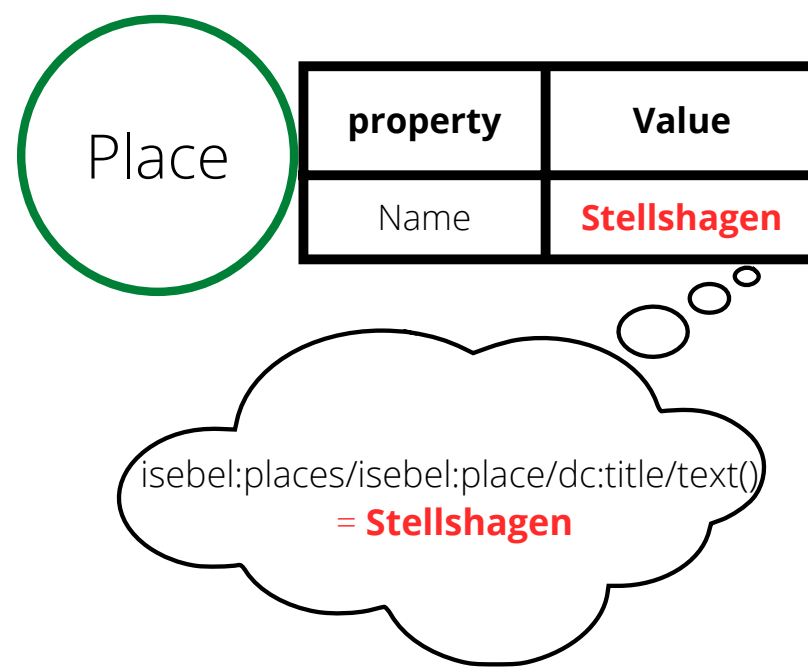
story

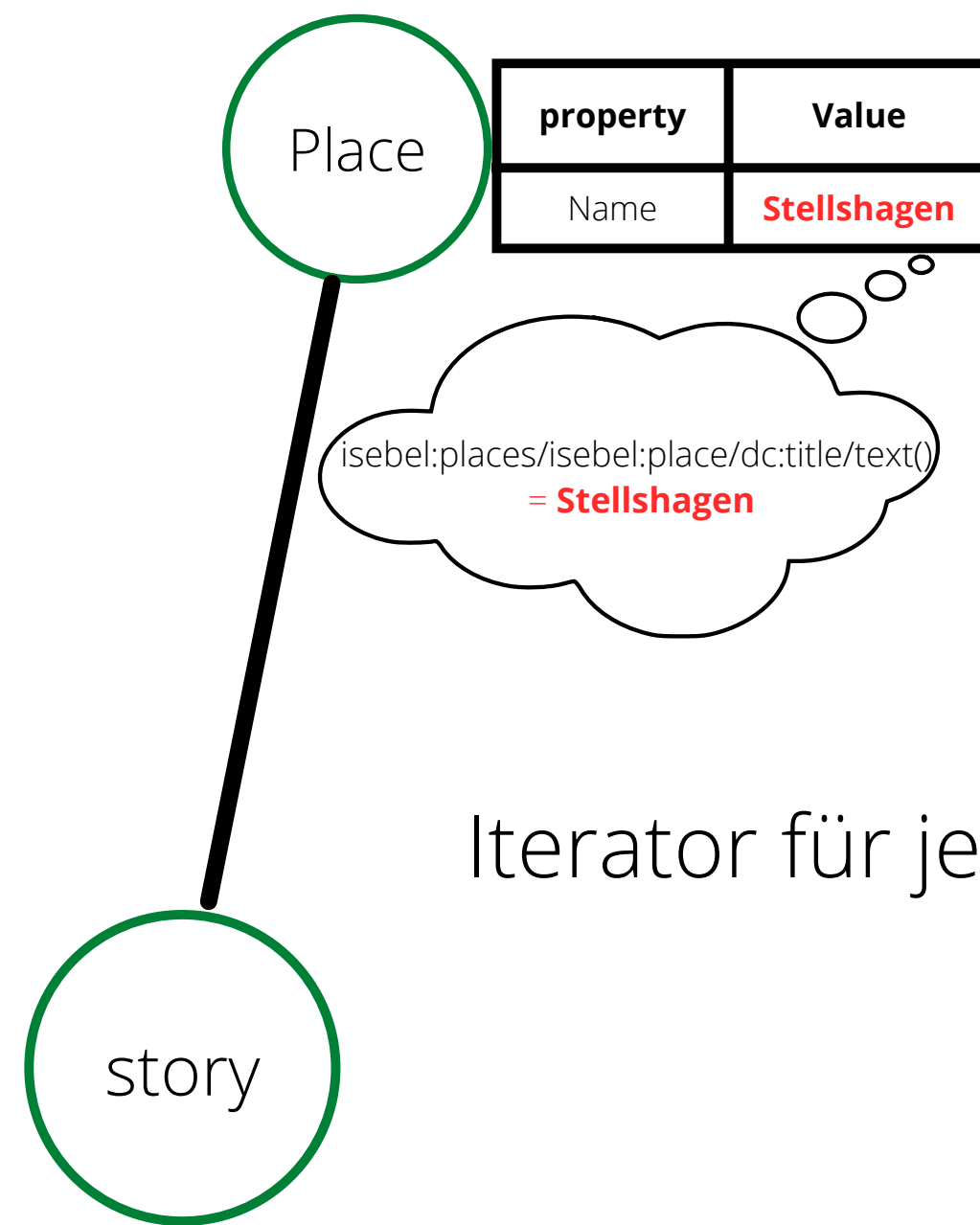
property name = "Stellshagen"

Place

property	Value
Name	Stellshagen

story



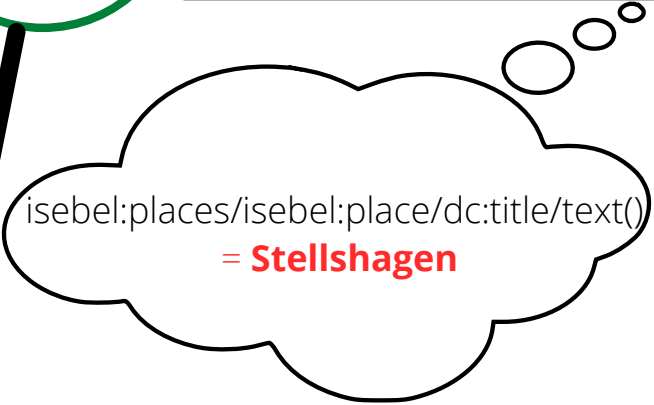


Iterator für jedes Place-Subelement im XML-Dokument!

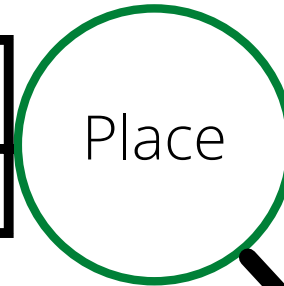
property	Value
Name	Hohen Viecheln



property	Value
Name	Stellshagen



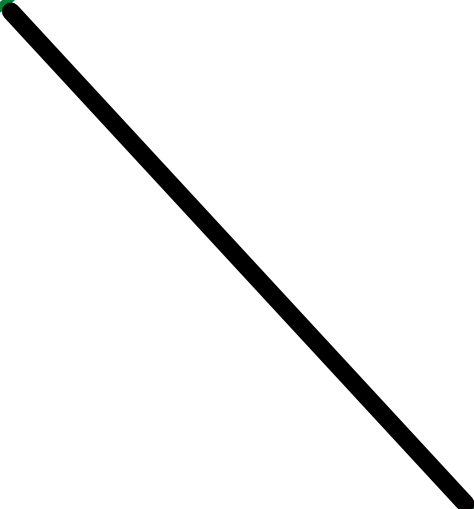
property	Value
Name	Hohen Viecheln



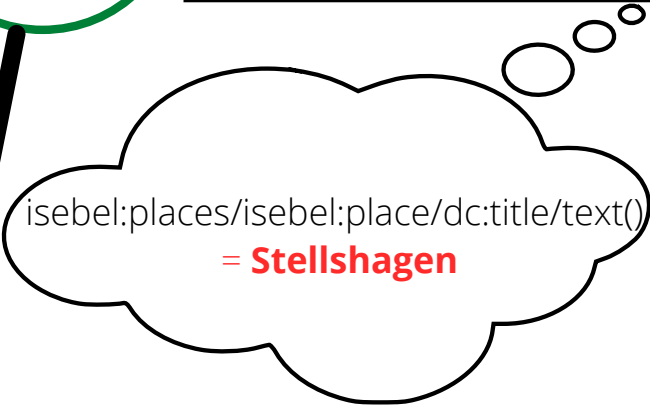
property	Value
Name	Stellshagen



property	Value
Name	Hohen Viecheln



property	Value
Name	Stellshagen



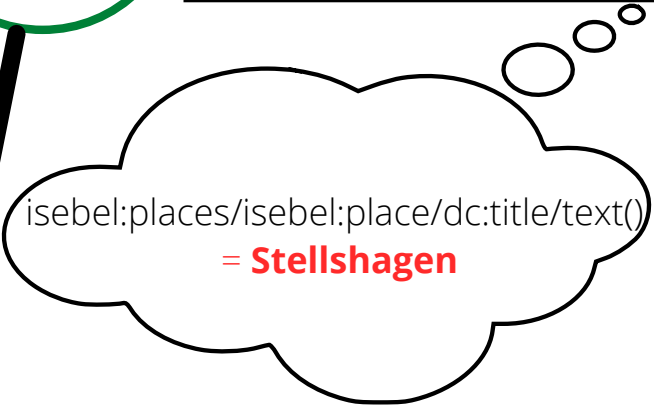
property	Value
name	Hinrichs
gender	male
profession	narrator
role	Pantoffelmacher



property	Value
Name	Hohen Viecheln

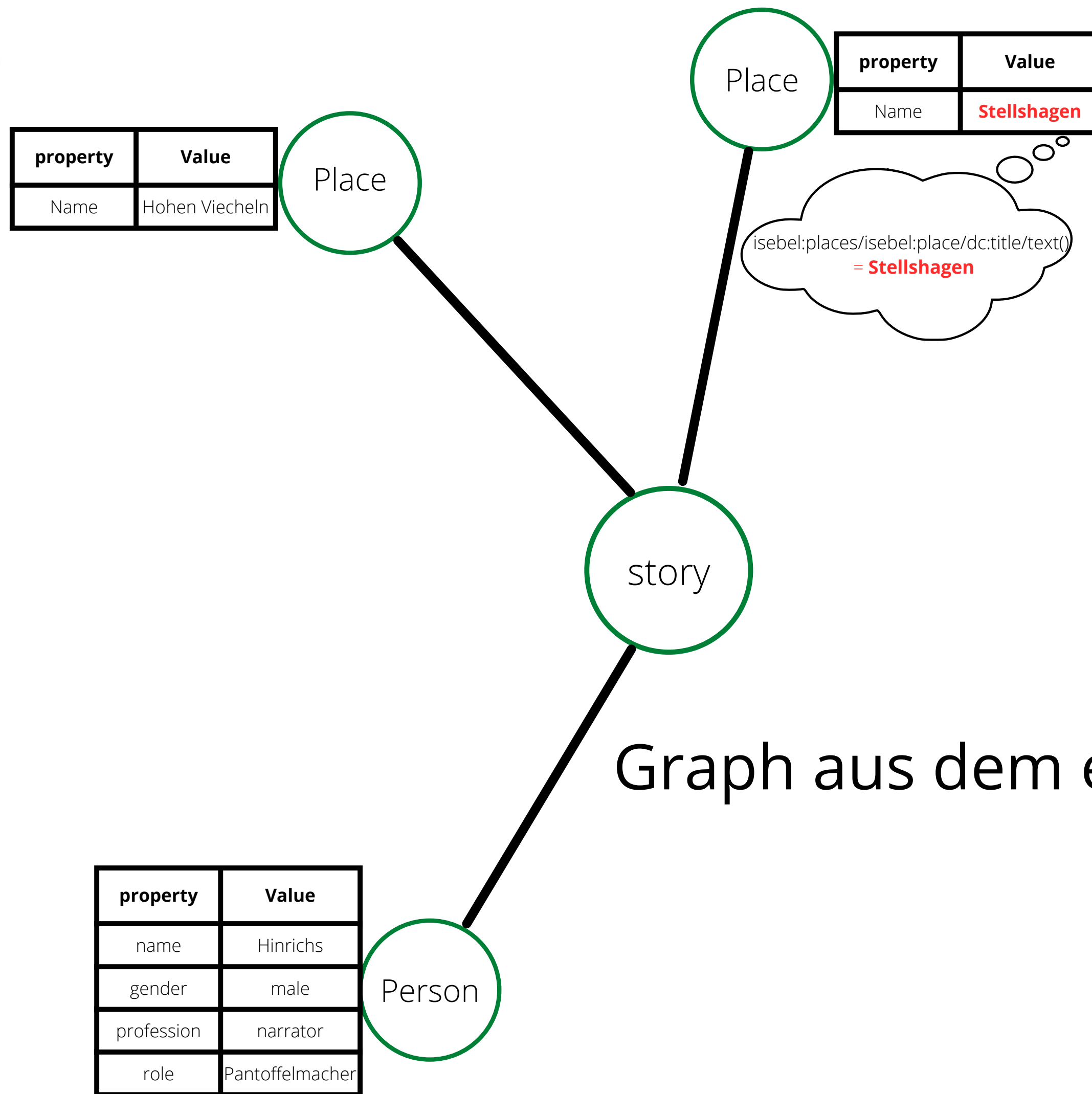


property	Value
Name	Stellshagen

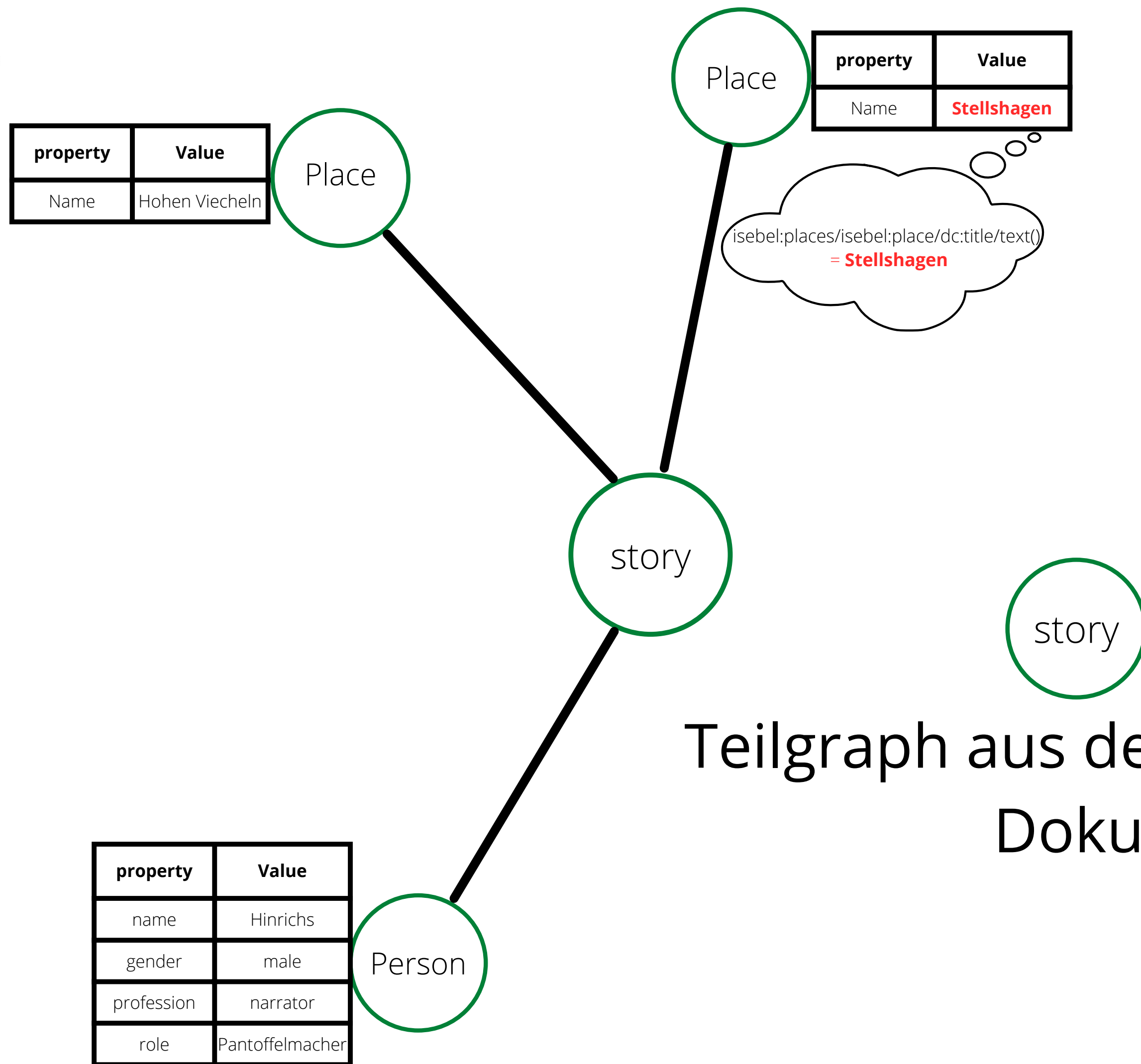


property	Value
name	Hinrichs
gender	male
profession	narrator
role	Pantoffelmacher

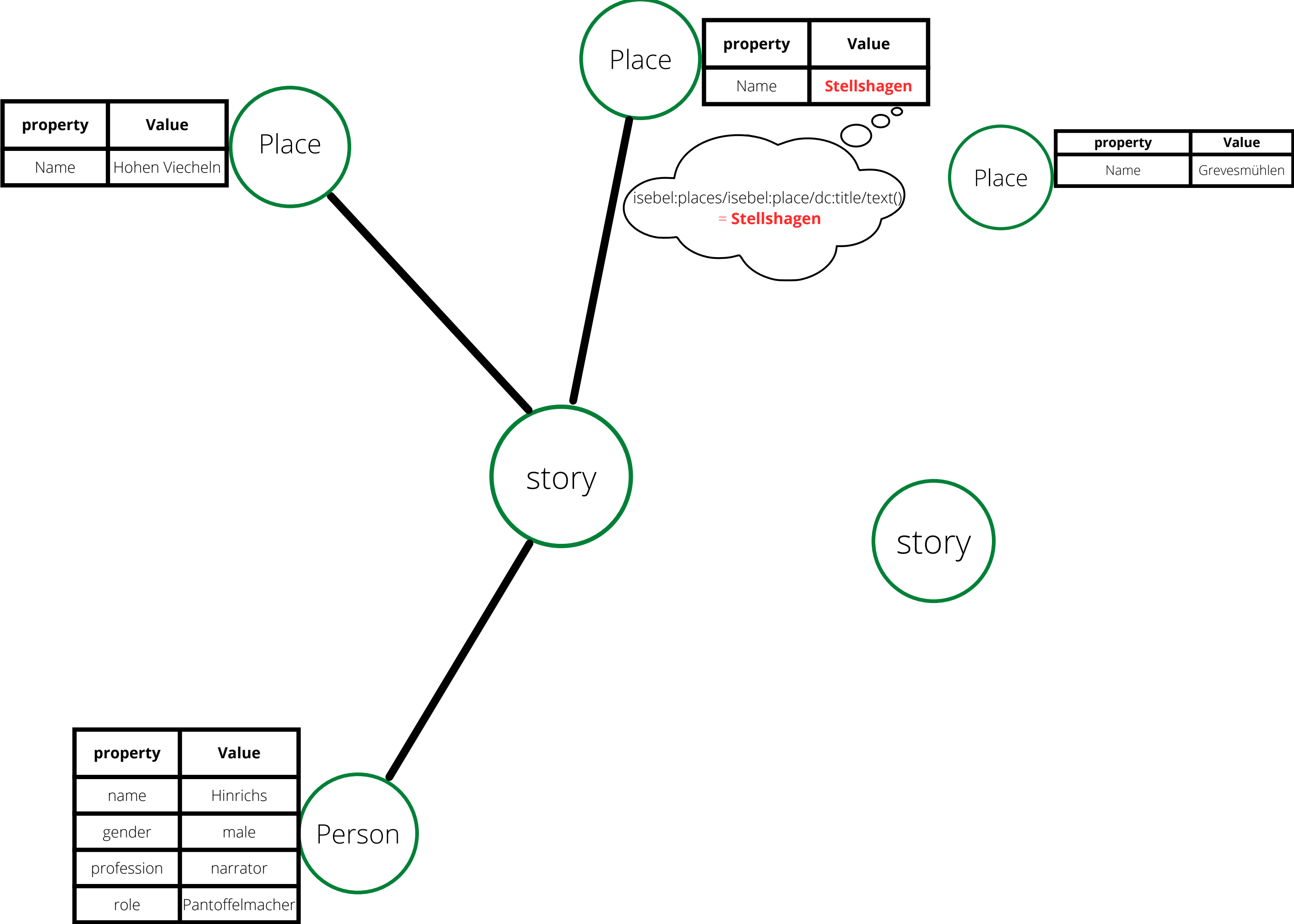


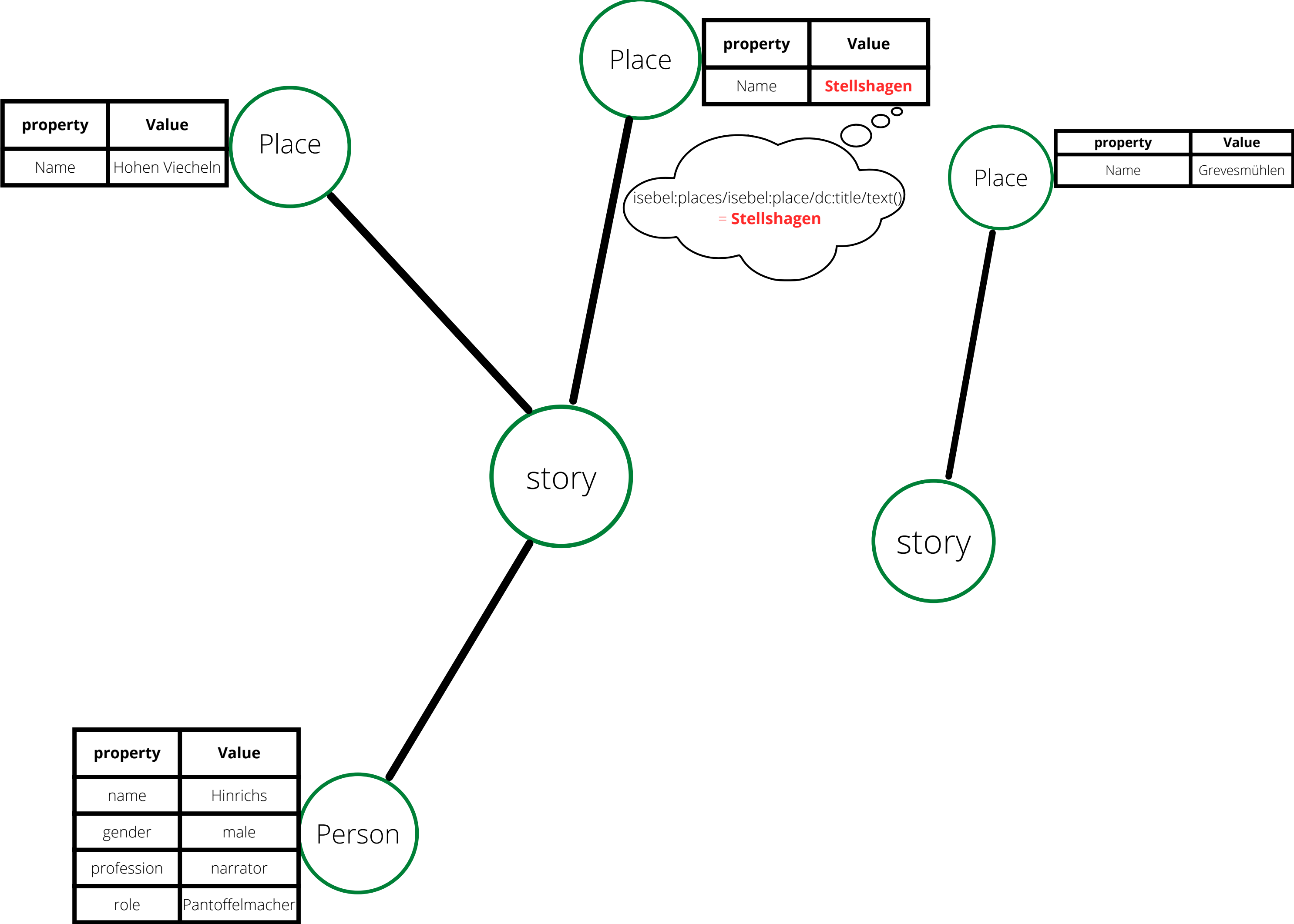


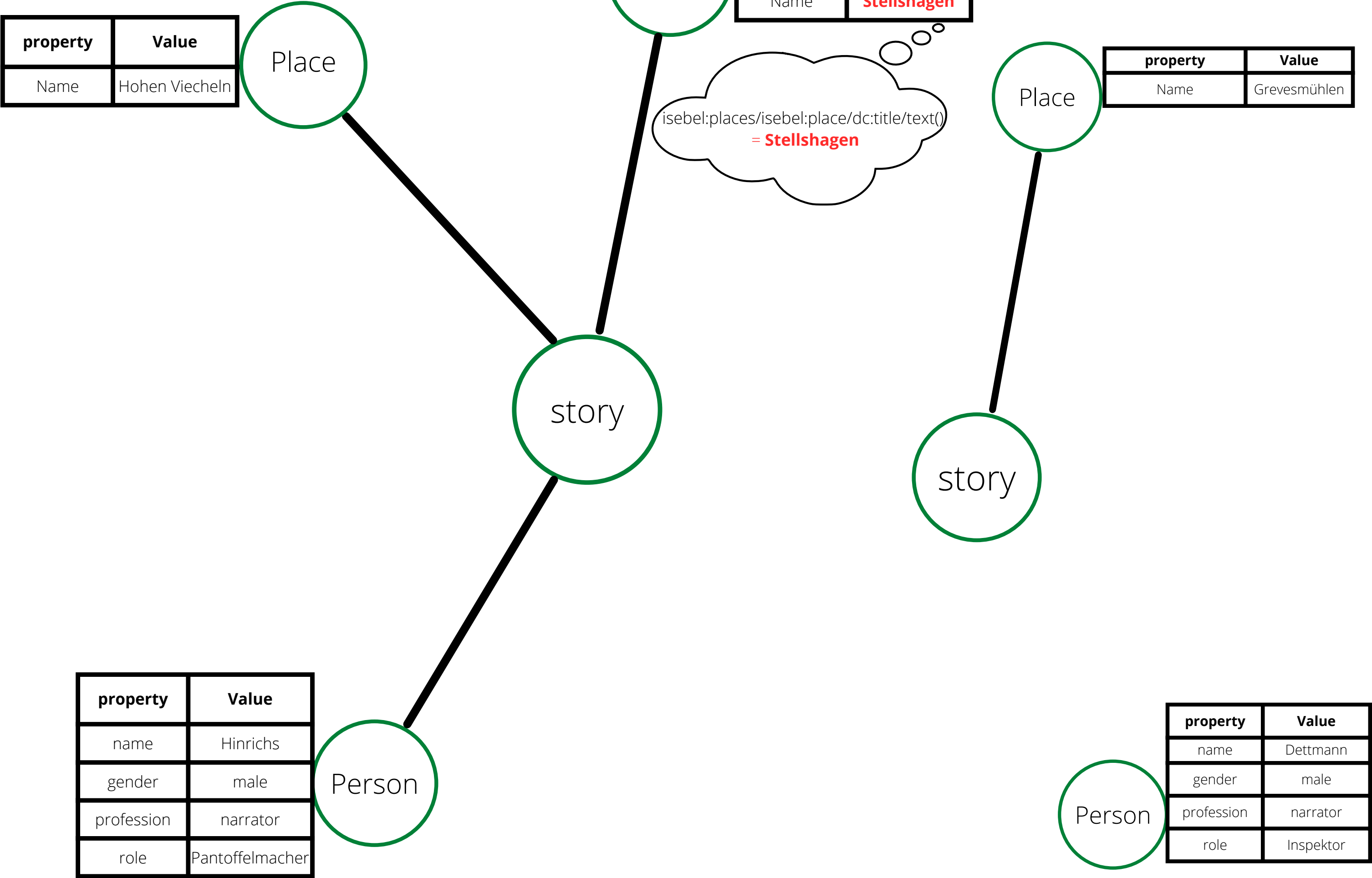
Graph aus dem ersten XML-Dokument



Teilgraph aus dem zweiten XML-Dokument



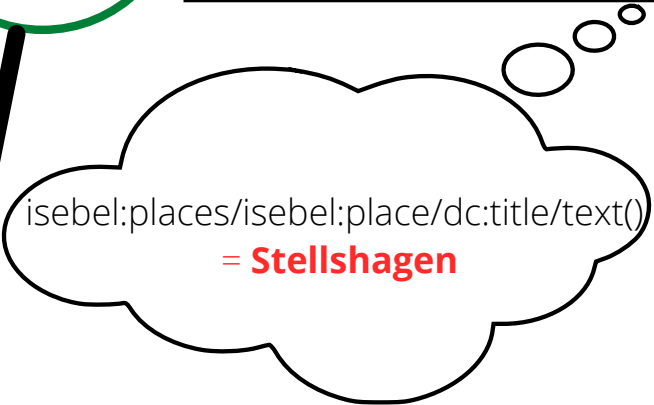




property	Value
Name	Hohen Viecheln



property	Value
Name	Stellshagen



property	Value
Name	Grevesmühlen



property	Value
name	Hinrichs
gender	male
profession	narrator
role	Pantoffelmacher



property	Value
name	Dettmann
gender	male
profession	narrator
role	Inspektor



property	Value
Name	Hohen Viecheln



property	Value
Name	Stellshagen



property	Value
Name	Grevesmühlen



property	Value
name	Hinrichs
gender	male
profession	narrator
role	Pantoffelmacher

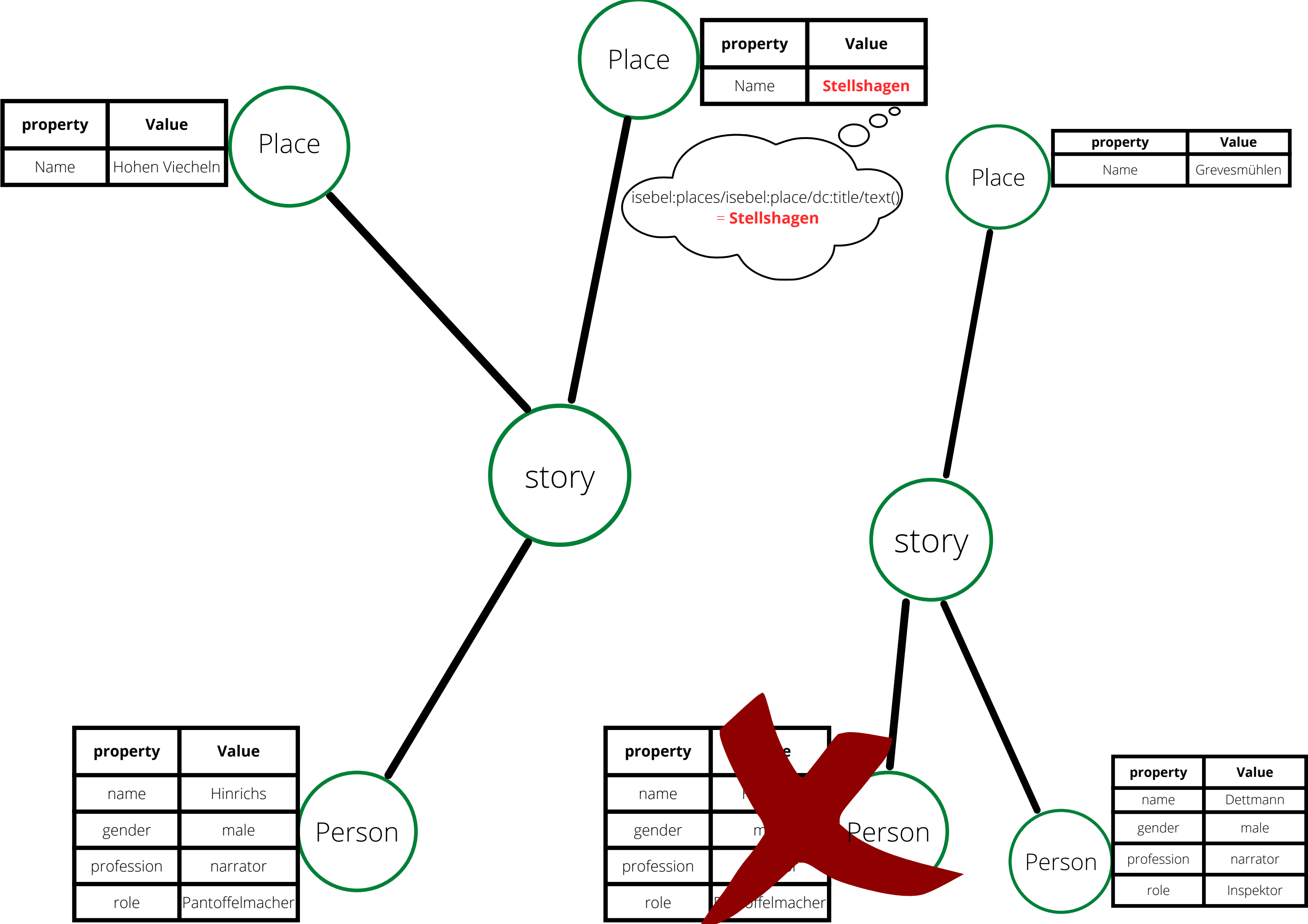


property	Value
name	Hinrichs
gender	male
profession	narrator
role	Pantoffelmacher



property	Value
name	Dettmann
gender	male
profession	narrator
role	Inspektor

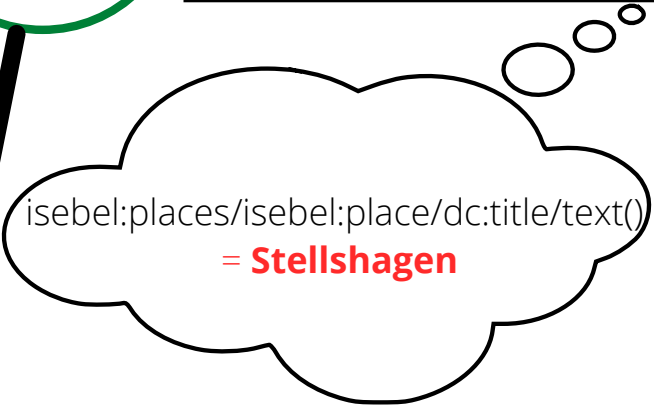




property	Value
Name	Hohen Viecheln



property	Value
Name	Stellshagen



property	Value
Name	Grevesmühlen



property	Value
name	Hinrichs
gender	male
profession	narrator
role	Pantoffelmacher



property	Value
name	Dettmann
gender	male
profession	narrator
role	Inspektor



Node replacement cases

Problemdarstellung

Problem beim Erzeugen von Knoten:

- Verschiedene Daten/Text-Inhalte in XML-Dokumenten
- Wann wird ein neuer Knoten bzw. Kanten erzeugt?
- Wann kann ein existierender Knoten verwendet werden?
- Wann sind Knoten identisch?
- Wann wird ähnlicher Knoten verlinkt (d.h. über Kante "ähnlich" verbunden)?
- Label gibt Zugehörigkeit zu einer Klasse an/Typangabe
- Properties können in einem Teil übereinstimmen
- Schlüsseleigenschaften/Properties gegeben?
- Wenn nicht: Inwieweit stimmen Properties überein?

Fallunterscheidung notwendig!

```
<isebel:persons>
  <isebel:person id="110001368">
    <isebel:name>Hinrichs</isebel:name>
    <isebel:gender>male</isebel:gender>
    <isebel:role>narrator</isebel:role>
    <isebel:profession>Example</isebel:profession>
  </isebel:person>
</isebel:persons>
```

```
<isebel:persons>
  <isebel:person id="110002742">
    <isebel:name>Hinrichs</isebel:name>
    <isebel:gender>male</isebel:gender>
    <isebel:role>narrator</isebel:role>
    <isebel:profession>Pantoffelmacher</isebel:profession>
  </isebel:person>
</isebel:persons>
```

```
<isebel:places>
  <isebel:place id="100002991">
    <dc:title>Warin</dc:title>
    <isebel:point>
      <datacite:pointLatitude xmlns:datacite="http://datacite.org/schema/kernel-4">53.80238</datacite:pointLatitude>
      <datacite:pointLongitude xmlns:datacite="http://datacite.org/schema/kernel-4">11.70502</datacite:pointLongitude>
    </isebel:point>
    <isebel:role>narration</isebel:role>
  </isebel:place>
```

```
<isebel:place id="100002991">
  <dc:title>Warin</dc:title>
  <isebel:point>
    <datacite:pointLatitude xmlns:datacite="http://datacite.org/schema/kernel-4">53.80238</datacite:pointLatitude>
    <datacite:pointLongitude xmlns:datacite="http://datacite.org/schema/kernel-4">11.70502</datacite:pointLongitude>
  </isebel:point>
  <isebel:role>action</isebel:role>
</isebel:place>
```

Nodes Replacement

Die verschiedenen Varianten

Zwei Hauptszenarien



Erzeuge keinen Knoten

- Verlinkt vorhandenen Knoten
- Ergänze evtl. weitere Properties

Erzeuge einen Knoten / Kanten

- Neuer Knoten wird erzeugt
- Neuer Knoten wird erzeugt + kante "Is Related" verbunden mit dem Vorherigen Kanten

Beispiele folgen



N1(G1)

property	Value
name	Hinrichs
gender	male
profession	narrator
role	Pantoffelmacher

N2(G2)

property	Value
name	Hinrichs
gender	male
profession	narrator
role	Pantoffelmacher

Es wird kein neuer Knoten erstellt

N1(G1)

property	Value
name	Hinrichs
gender	male
profession	narrator
role	Pantoffelmacher

N2(G2)

property	Value
name	Hinrichs
gender	male
profession	narrator

Es wird kein neuer Knoten erstellt

N = Knoten
G = Graph

N1(G1)

property	Value
name	Hinrichs
gender	male
profession	narrator

N2(G2)

property	Value
name	Hinrichs
role	Pantoffelmacher

Es wird kein neuer Knoten erstellt. Die Property "role=Pantoffelmacher" wird von N1(G1) zu N2(G2) übertragen

N1(G1)

property	Value
name	Hinrichs
gender	male
role	Pantoffelmacher

N2(G2)

property	Value
name	Hinrichs
profession	narrator

Es wird kein neuer Knoten erstellt

N1(G1)

property	Value
name	Hinrichs
gender	male
profession	narrator
role	Pantoffelmacher

N2(G2)

property	Value
gender	male
profession	narrator
role	Pantoffelmacher

Es wird kein neuer Knoten erstellt

N1(G1)

property	Value
name	Hinrichs
gender	male
profession	narrator
role	Pantoffelmacher

N2(G2)

property	Value
gender	male
profession	narrator

Es wird kein neuer Knoten erstellt

N = Konten

G = Graph

N1(G1)

property	Value
name	Hinrichs
gender	male
profession	narrator

N2(G2)

property	Value
name	Hinrichs

Es wird kein neuer Knoten erstellt

N1(G1)

property	Value
name	Hinrichs
gender	male

N2(G2)

property	Value
name	Hinrichs

Es wird kein neuer Knoten erstellt

N1(G1)

property	Value
name	Hinrichs
gender	male
profession	narrator
role	Pantoffelmacher

K2(G2)

property	Value
name	Dettmann
gender	male
profession	narrator
role	Inspektor

Ein neuer Knoten wird erstellt

N1(G1)

property	Value
name	Dettmann
gender	male
profession	narrator
role	Pantoffelmacher

N2(G2)

property	Value
name	Hinrichs
gender	male
profession	narrator
role	Pantoffelmacher

Ein neuer Knoten wird erstellt

N = Knoten

G = Graph

N1(G1)

property	Value
name	Hinrichs

N2(G2)

property	Value
name	Dettmann
gender	male
profession	narrator
role	Pantoffelmacher

Ein neuer Knoten wird erstellt

N1(G1)

property	Value
name	Hinrichs
gender	male

N2(G2)

property	Value
name	Hinrichs
gender	male
profession	narrator
role	Pantoffelmacher

Kein neuer Knoten wird erstellt

N1(G1)

property	Value
name	Hinrichs
gender	male
profession	narrator

N2(G2)

property	Value
name	Hinrichs
gender	male
profession	narrator
role	Inspektor

Kein neuer Knoten, Eigenschaften ergänzen beim alten Knoten!

N1(G1)

property	Value
name	Hinrichs
role	Pantoffelmacher

N2(G2)

property	Value
name	Hinrichs
gender	male
profession	narrator

Alten Knoten lassen und Eigenschaften vereinigen.

N = Knoten
G = Graph

N1(G1)

property	Value
name	Hinrichs
profession	narrator

N2(G2)

property	Value
name	Hinrichs
profession	narrator
role	Pantoffelmacher

Kein neuer Knoten, Eigenschaften ergänzen beim alten Knoten!

N1(G1)

property	Value
gender	male
profession	narrator
role	Pantoffelmacher

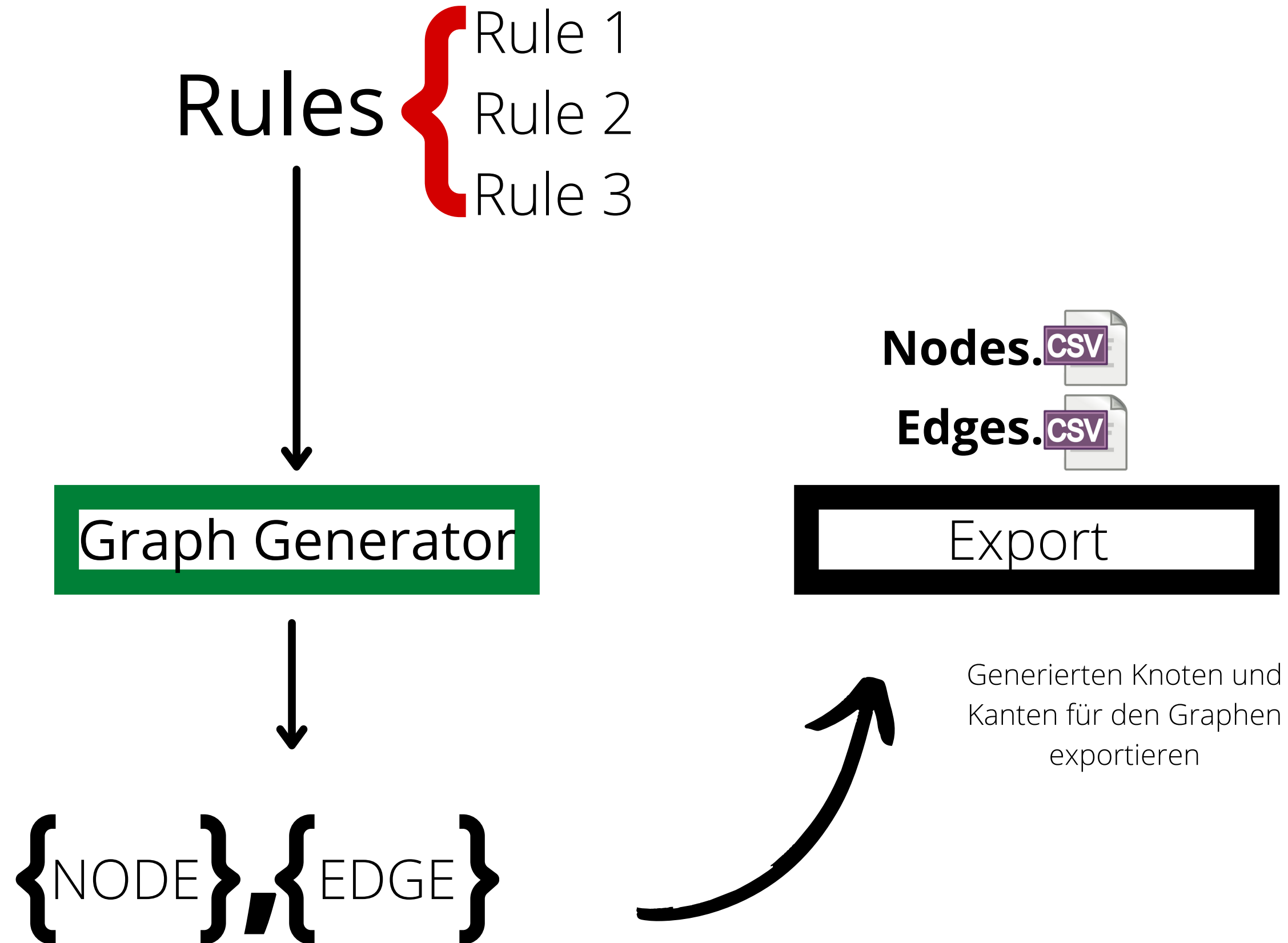
Is Related

N2(G2)

property	Value
name	Hinrichs
gender	male
profession	narrator
role	Pantoffelmacher

Neuer Knoten erzeugt, aber eine Kante "Is Related" wird da erzeugt

Exporter



Beispiel für die Ausgabe als .CSV

Node. 

```
"ID","Label","Type","Property"  
"1","xmd_s001_000_000_005.xml","story",""  
"2","Person","Person","name:Hinrichs;gender:male;role:narrator;profession:Pantoffelmacher"  
"3","Place","Place","name:Hohen Viecheln"  
"4","Place","Place","name:Stellshagen"  
"5","xmd_s001_000_000_006.xml","story",""  
"6","Place","Place","name:Grevesmühlen"
```

Edge. 

```
"ID","Source","Target","Label"  
"1","1","2","Person"  
"2","1","3","Place"  
"3","1","4","Place"  
"4","5","2","Person"  
"5","5","6","Place"
```

Beispiel-Szenarien für Graph-Generieren:

- Alle Storys von einer bestimmten Person XY
- Alle Storys, die an einem bestimmten Ort erzählt wurden (Stellshagen, Wismar...)
- Einfluss des Geschlechts des Erzählers auf die Hexen- oder Werwolf-Sagen
- Alle Storys, die in 3 Sprachen (DEU, NDS, ENG) verfügbar sind

Regelsprache Beispiel

```
// nested match example, the second match is evaluated within the context of the first
match xpath("//story") using $s {
  create node $sn label "story" {
    // properties
    content = $s.xpath("content/text()"),
    title = $s.xpath("title/text()"),
    unique (title)
  },
  match xpath("./person[@role='narrator']") using $p {
    create node $pn label "person" {
      // properties with optional type declaration
      name = $p.xpath("name/text()"),
      age = $p.xpath("age"),
      unique (name, age),
      // conditional assignment
      if $p.xpath("appellation/text()") == "Mr" {
        gender = "male" + "narrator"
      }
    },
    create edge $e from $sn to $pn label "narrator" {
      alt = "story-teller"
      /* no more properties given */
    }
  }
  // should warn about $p hiding earlier binding
}
}
```

Zusammenfassung



- Konzept/Entwurf erstellt ✓
- Folgende Teile wurden implementiert:
- Filter ✓
- Graph-Generator ✓
 - Knoten
 - Kanten
- Exporter ✓
 - Csv

Ausblick



- Teile die nicht implementiert wurden Zeitbeschränkung :
- Wichtige Erweiterungen
 - Die Regelsprache einbinden
- Mögliche/zukünftige Erweiterungen:
 - Document stores JSON-Daten (Gibt es bereit Nachfolger für die Arbeit)
 - Key-values store
 - No-Sql Systemen
 - Relationale Strukturen

Daraus noch Graphen erzeugen.

- Einfache Erweiterungen
- Ausgabe der generierten Graphen
 - Gexf
 - Graphml
 - etc.

Universität
Rostock



Traditio et Innovatio

Danke für eure Aufmerksamkeit 🌸

Wenn ihr Fragen habt, könnt ihr diese jetzt stellen

