

STATUS OF THE APACHE ODF TOOLKIT (INCUBATING)

FOSDEM
FEB. 3, 2018

Svante.Schubert@cib.de



PROTOTYPEFUND.DE

- Funding Open Source Software
- 50.000 € over 6 months
- 51 days left (next round)
- 8 rounds to come till 2020
- Alone or as group
- Germany as place of residence



OPENDOCUMENT FILE FORMAT STANDARD DOES MATTER A LOT

- Open Source is **not** enough..
- (Office) Interoperability goes a long way...
- Freedom of choice of ODF application



ODF TOOLKIT (INCUBATING) HISTORY

- Started around 2005
- Putting server solutions into one place
- From StarOffice / Sun & IBM Team
- Used as backend of Sun Weboffice
- Donated with OpenOffice to Apache by Oracle



ODF Validator

This service checks conformance of ODF documents based on their OpenDocument Format specification. It does not cover all conformance criteria, yet (see [implementation details](#)).

ODF Version:

auto-detect ▼

Logging:

verbose ▼

Choose ODF documents for validation:

Keine ausgewählt

This service is provided to you by [The Document Foundation](#)



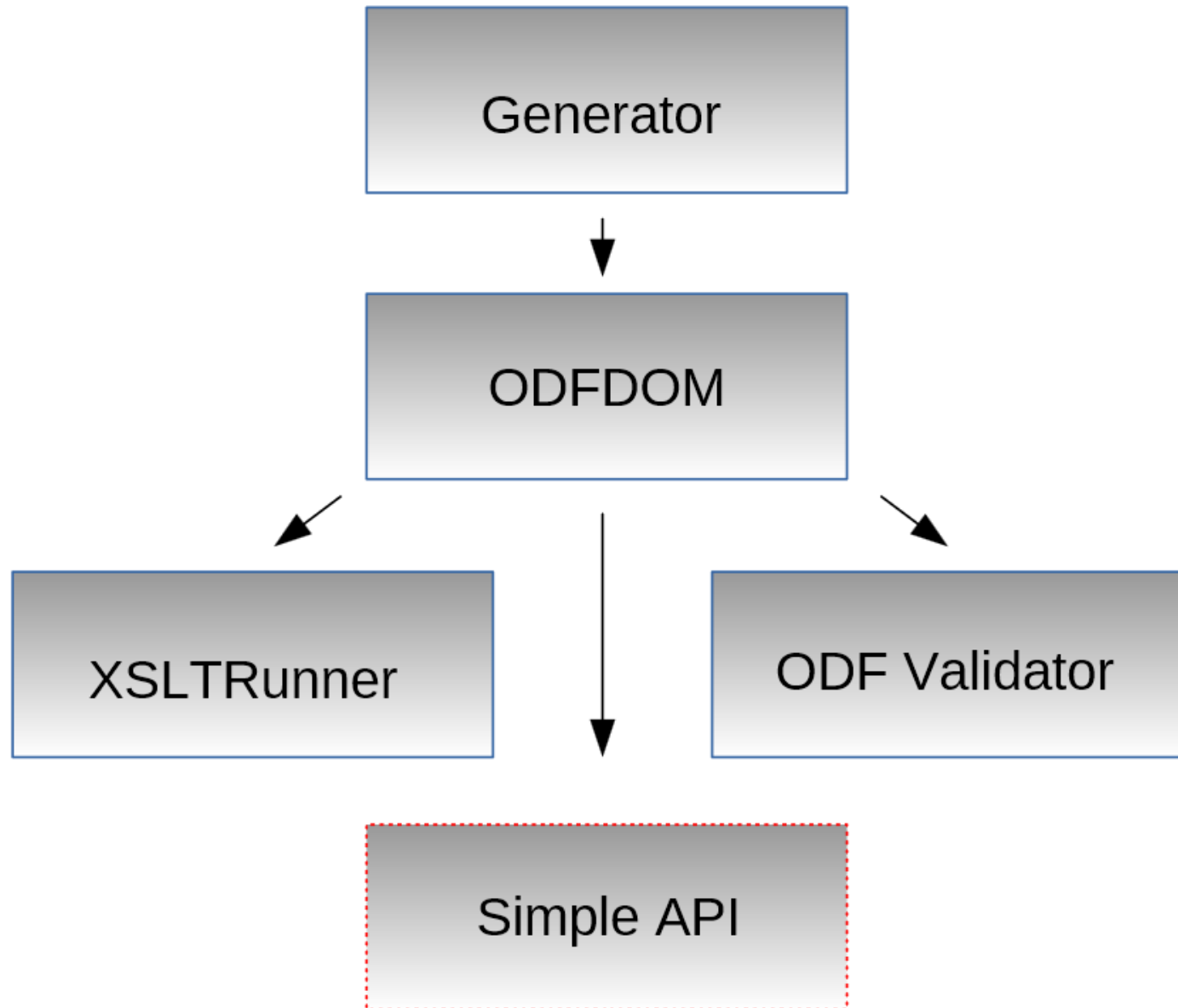
LibreOffice
The Document Foundation

This service does not cover all conformance criteria of the OpenDocument Format specification. It is not applicable for formal validation proof. Problems reported by this service only indicate that a document may not conform to the specification. It must not be concluded from errors that are reported that the document does not conform to the specification without further investigation of the error report, and it must not be concluded from the absence of error reports that the OpenDocument Format document conforms to the

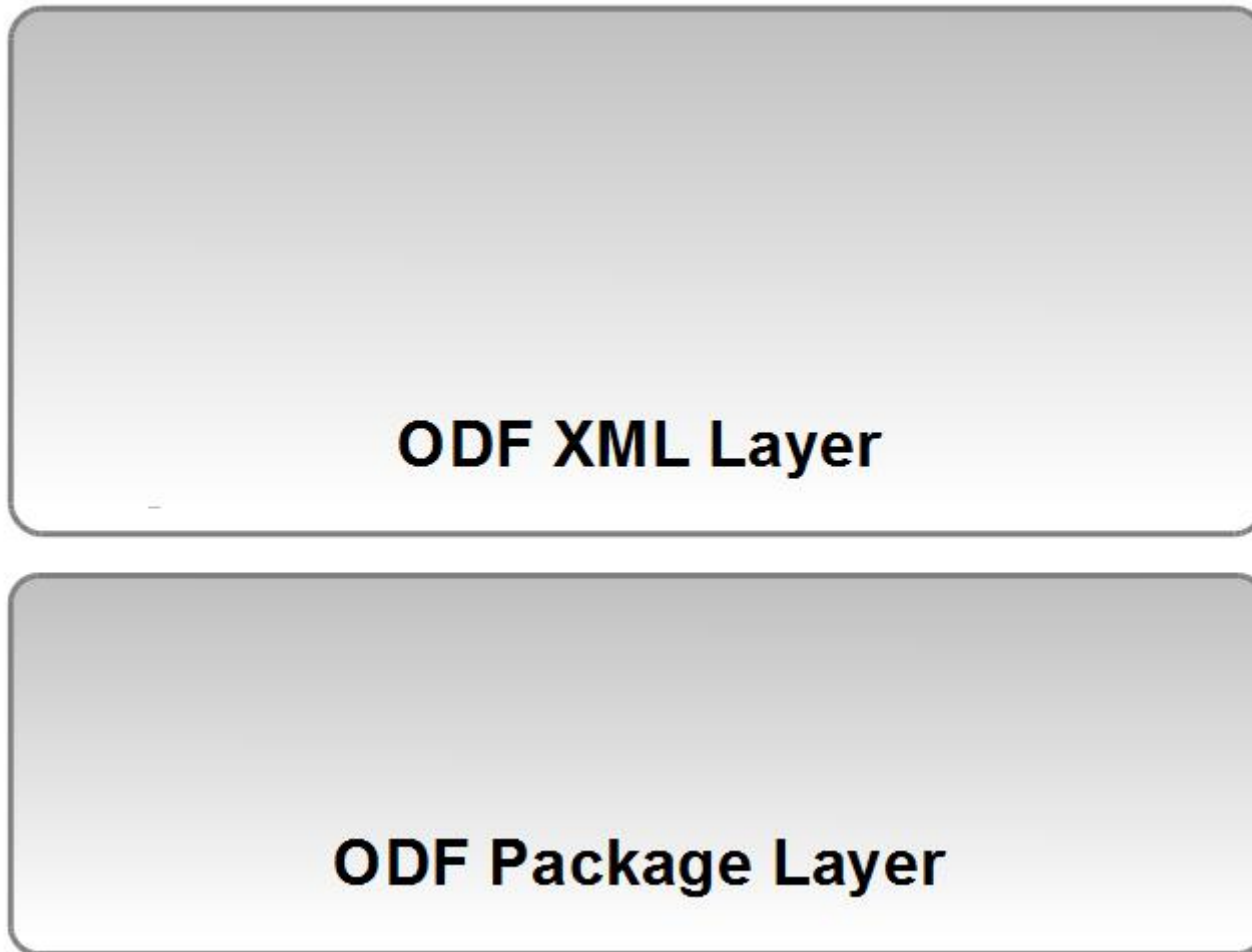
OpenDocument Format specification.

[Impressum \(Legal Info\)](#) | [Privacy Policy](#)

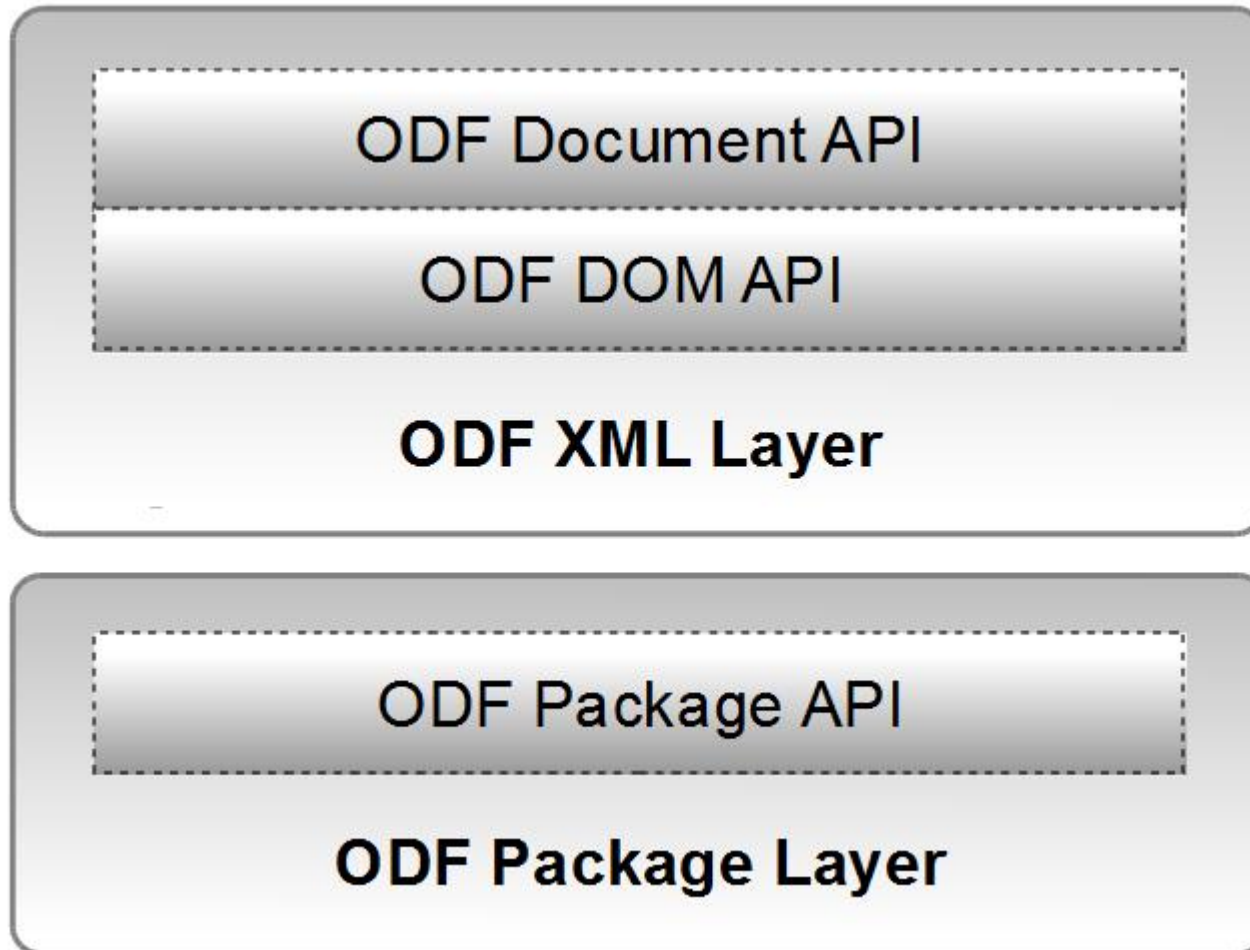
ODF TOOLKIT (INCUBATING) ARCHITECTURE



ODFDOM ARCHITECTURE



ODFDOM ARCHITECTURE



OPENDOCUMENT FILE FORMAT STANDARD DOES MATTER A LOT

ODF 1.2 consists of 3 parts

- Part 1 specifies XML Model
- Part 2 specifies Formula
- Part 3 specifies Package Format (ZIP)



ODF TOOLKIT GENERATOR

ODF XML SCHEMA – USABILITY NIGHTMARE

```
<define name="table-table">
  <element name="table:table">
    <ref name="table-table-attlist"/>
    <optional>
      <ref name="table-title"/>
    </optional>
    <optional>
      <ref name="table-desc"/>
    </optional>
    <optional>
      <ref name="table-table-source"/>
    </optional>
    <optional>
      <ref name="office-dde-source"/>
    </optional>
    <optional>
      <ref name="table-scenario"/>
    </optional>
    <optional>
      <ref name="office-forms"/>
    </optional>
    <optional>
      <ref name="table-shapes"/>
    </optional>
    <ref name="table-columns-and-groups"/>
    <ref name="table-rows-and-groups"/>
    <optional>
      <ref name="table-named-expressions"/>
    </optional>
  </element>
</define>
<define name="table-columns-and-groups">
  <oneOrMore>
```

ODF 1.2 documents:

- 598 XML Elements
- 1300 XML Attributes

ODF TOOLKIT GENERATOR

ODF XML SCHEMA 2 CODE

- **First try:**

- Using XSL Transformation filling Java Templates
- !JAXP (Java XML Binding works only for W3C Schema)

- **Currently:** Improved redesign by software reuse:

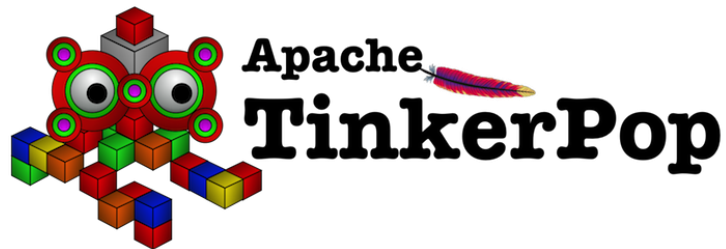
- Reusing parser from Multi Schema Validator (MSV)
- Reusing template engine from Apache Velocity
- Desired ODF information collected in Lists & Maps

- **Future:** Better using Graphs..

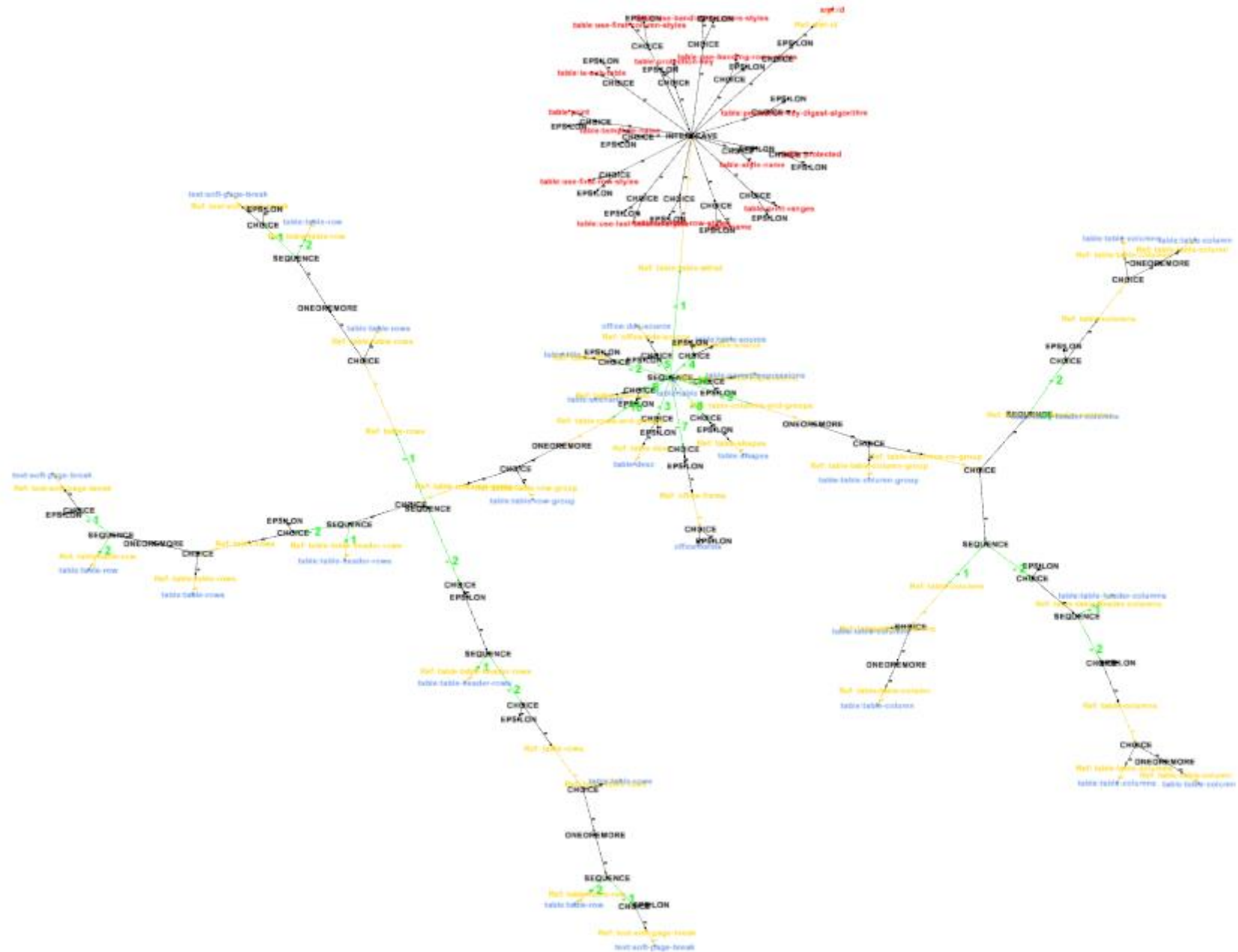
ODF TOOLKIT GENERATOR

ODF XML SCHEMA 2 CODE

- Multi Schema Validator Model dumped to TXT
- Mapped with ANTLR grammar to GraphML model
- GraphML model refactored Apache TinkerPop
- Apache TinkerPop compiled to
- Using Apache TinkerPop Graph Scripting Language Gremlin

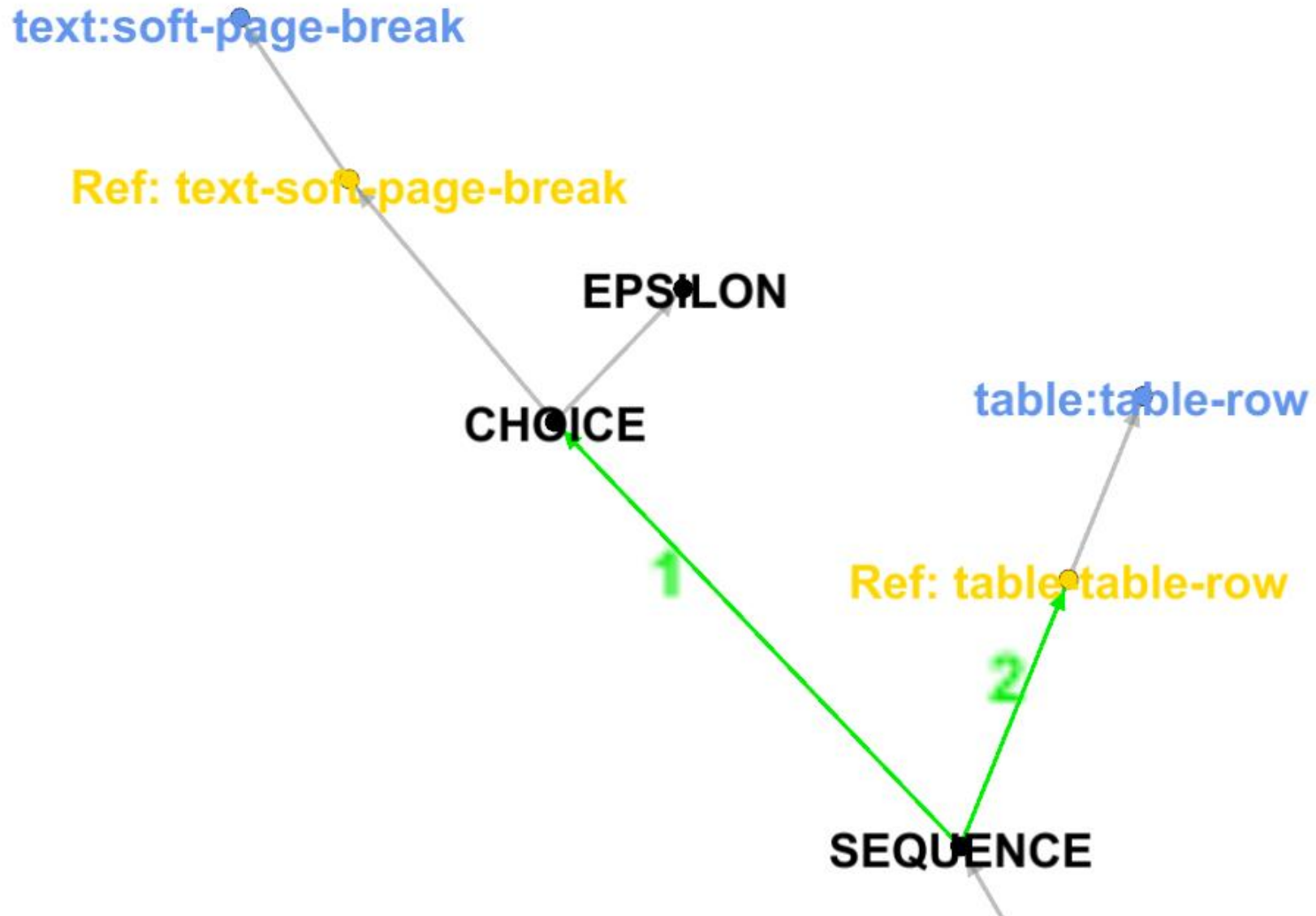


ODF GRAPH – TABLE GEPHI VIEWER



ODF GRAPH – TABLE

GEPHI VIEWER – REFACTORING / SIMPLIFYING



ODF TOOLKIT GENERATOR

ODF XML SCHEMA 2 CODE

- **Yet unresolved XML schema pattern for generation**
 - Schema CHOICE / SEQUENCE => no manual coding
 - Parent with multiple children with ID => Lazy Map
 - REF / ID are not connected => automate connection
 - Annotation with Metadata (User Changes, visible/layout relevant, metadata)
-

- **Collaboration** – multiple users working offline/async
 - Problem always the merge
 - Question always: What you have you changed?
 - Answered by **User Changes**
-

ODF TOOLKIT - COLLABORATION

USER CHANGES

- **User changes** implicitly a standard
 - User expectation of changes across office apps
 - 3 Types: insert / modify / delete
 - **User Components:**
 - Paragraph, image, table, character, ...
-

ODF TOOLKIT - COLLABORATION PROTOYPE PROMISE



- ODF Toolkit plus “changes support” just a black box:
 - ODT being mapped to a sequence of changes
 - Changes allowed to be merged back to ODT
-

ODF TOOLKIT - COLLABORATION CHANGES

- If **user changes** are a de-facto standard:
 - Enhancing the ODF RelaxNG XML with changes info?
 - Generating Changes Code?
Everything is deterministic!
 - **How?**
-

ODF TOOLKIT - COLLABORATION CHANGES



To be continued...
