

A data quality framework applied to e-government metadata/conceptual models

A prerequisite to establish governance of interoperable e-services

Per Myrseth
Det Norske Veritas
Oslo, Norway

Co-authors:
Jørgen Stang, David Skogan, Vibeke Dalberg and Jim Yang

Semicolon II



Tax Administration: Data capture in electronic forms

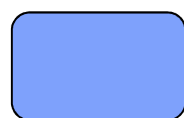


SERES: Metadata repository

CASE STUDY:

In 70+ types of data flows:
Measure the quality of the
metadata/ conceptual models

Citizen



Dataflow



Dataflow

Enterprise



Dataflow



**Tax
Administration**

Electronic forms: Tax Form, VAT form etc.

Metadata layers in the repository

Concepts and definitions

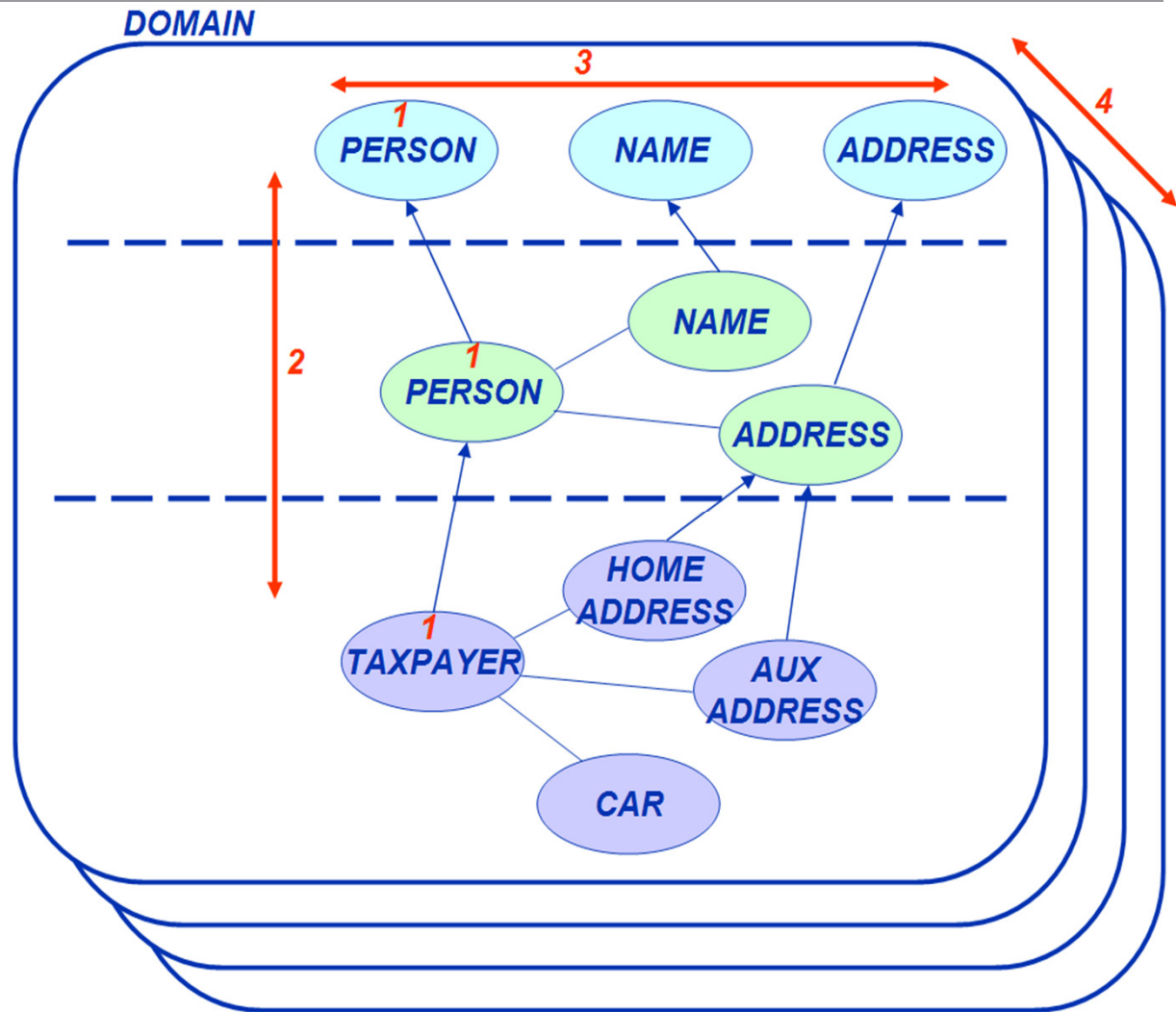
+ relations between terms

Structure

Reusable groups of concepts

Implementasjon

UML of a XML-Schema

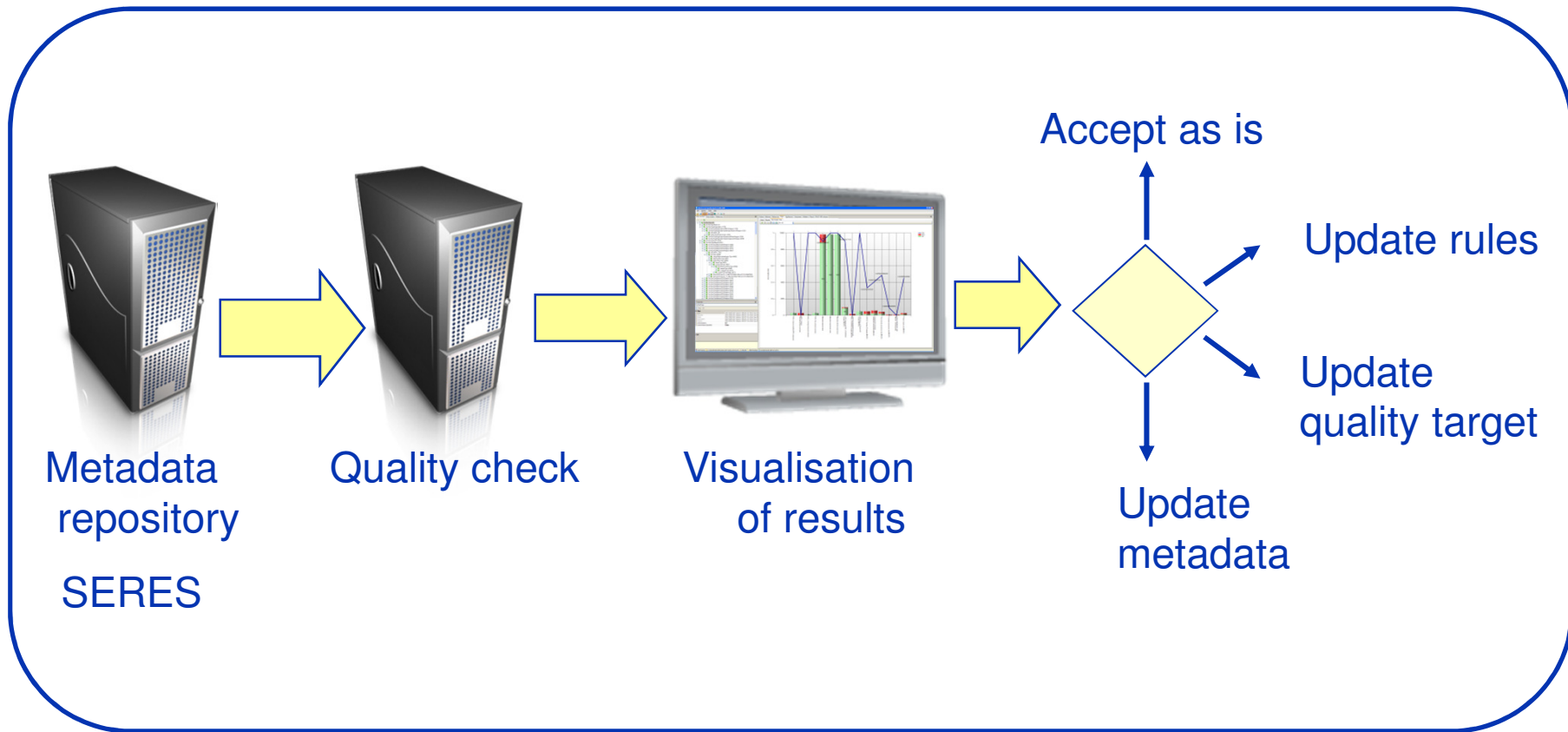


Automated quality metrics measured

	Metadata engineering rules	Rules measured
Concepts and definitions + relations between terms	5	0
Structure Reusable groups of concepts	23	0
Implementasjon UML of a XML-Schema	23	11+3

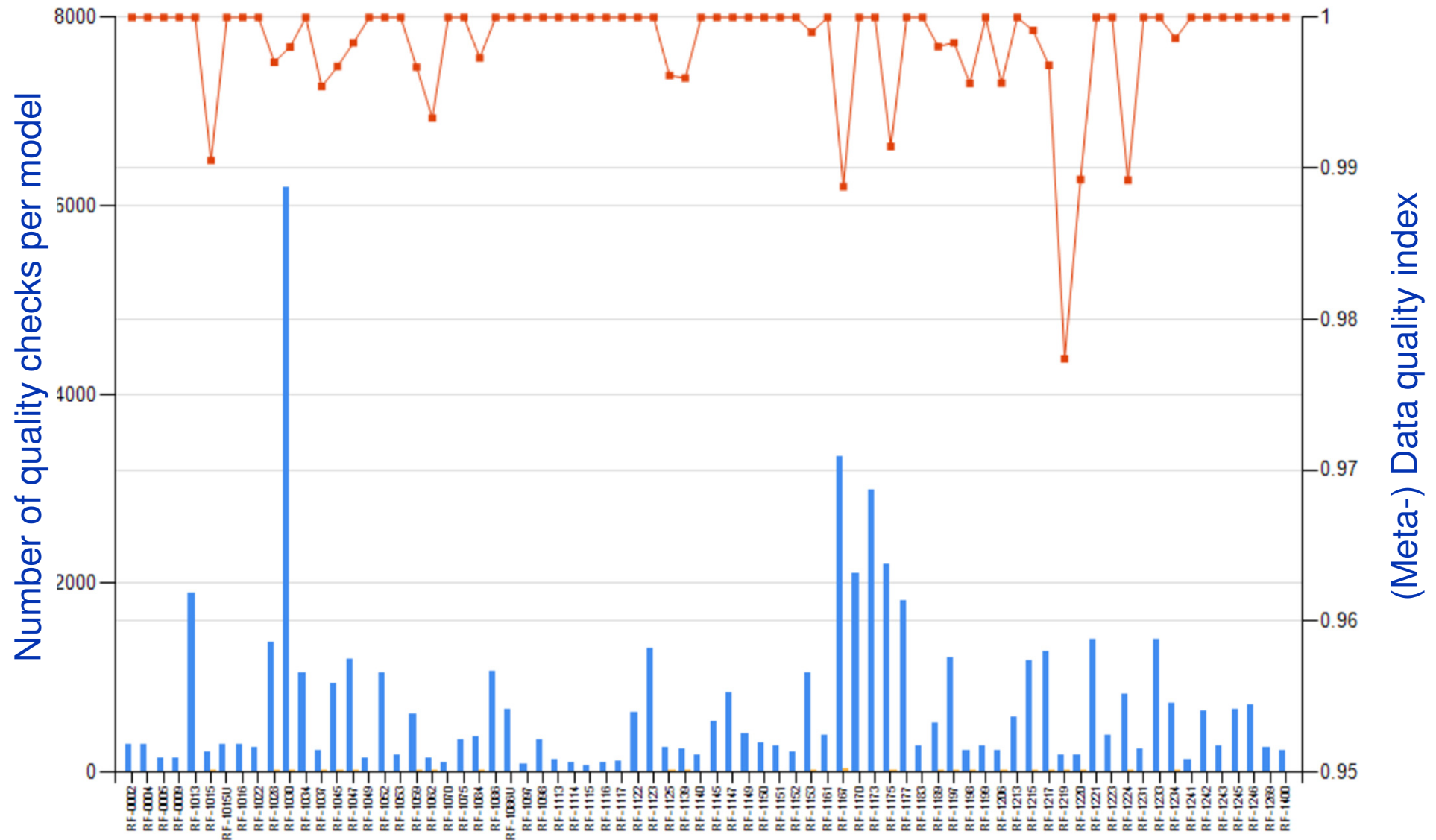
2

Case study setup



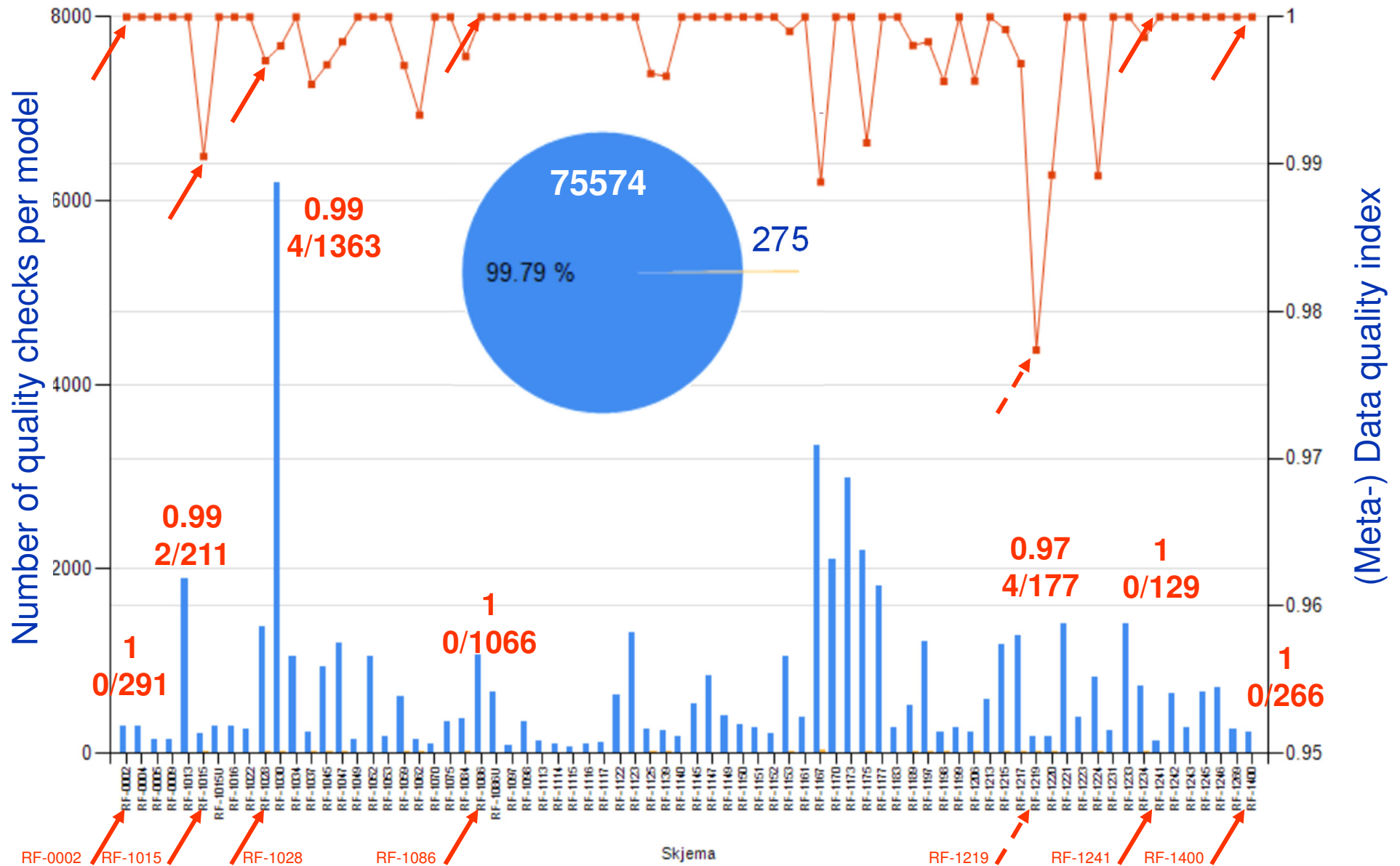
Measured result of metadata.

(Metadata content in repository is work in progress)



Measured result of metadata.

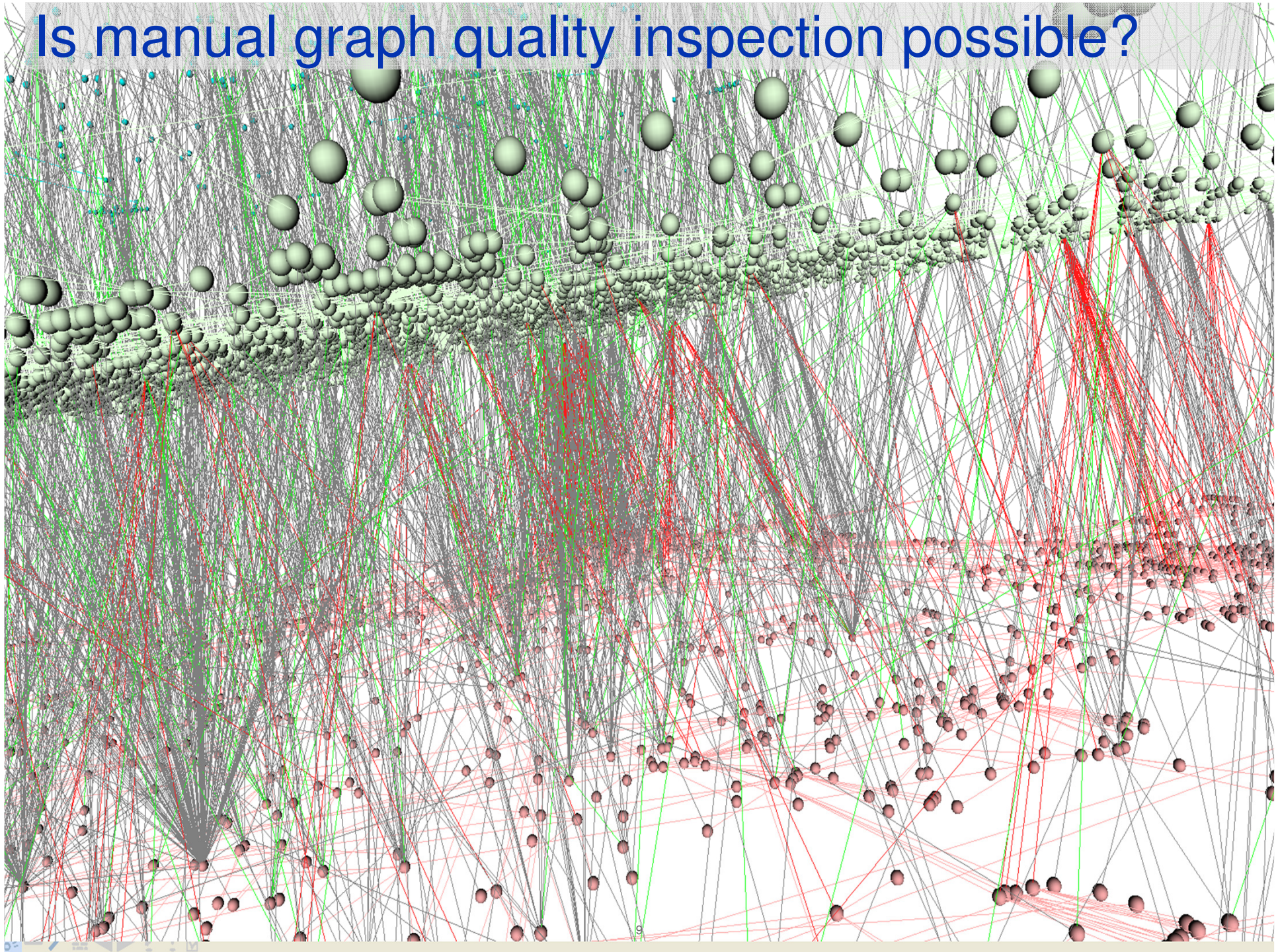
(Metadata content in repository is work in progress)



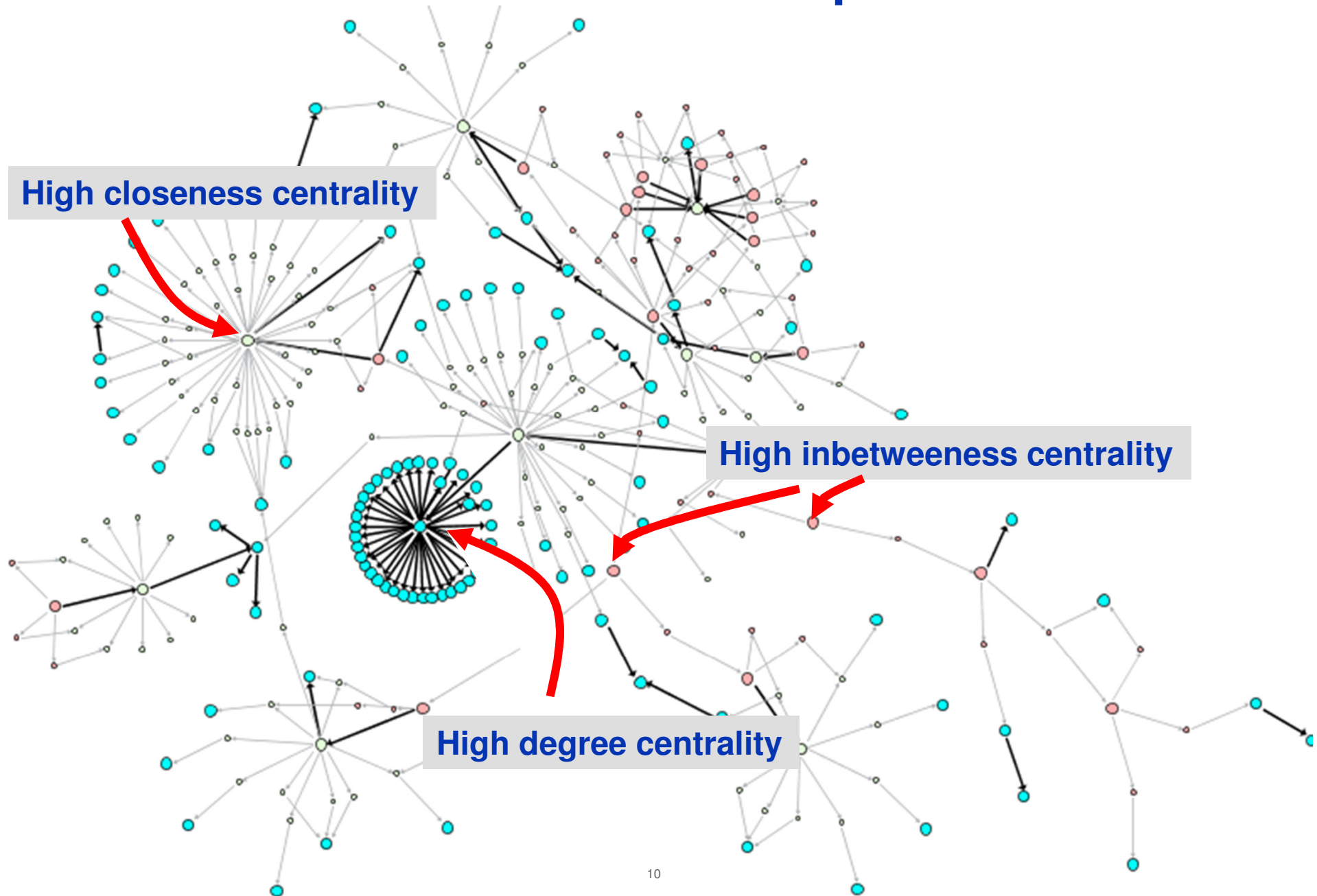
Framework: Categories of quality metrics for metadata/ conceptual models

	Syntax Schema	Semantics Meaning	Pragmatic Suited for use
Concepts and definitions + relations between terms			
Structure Reusable groups of concepts	✓		
Implementasjon UML of a XML-Schema	✓		

Is manual graph quality inspection possible?

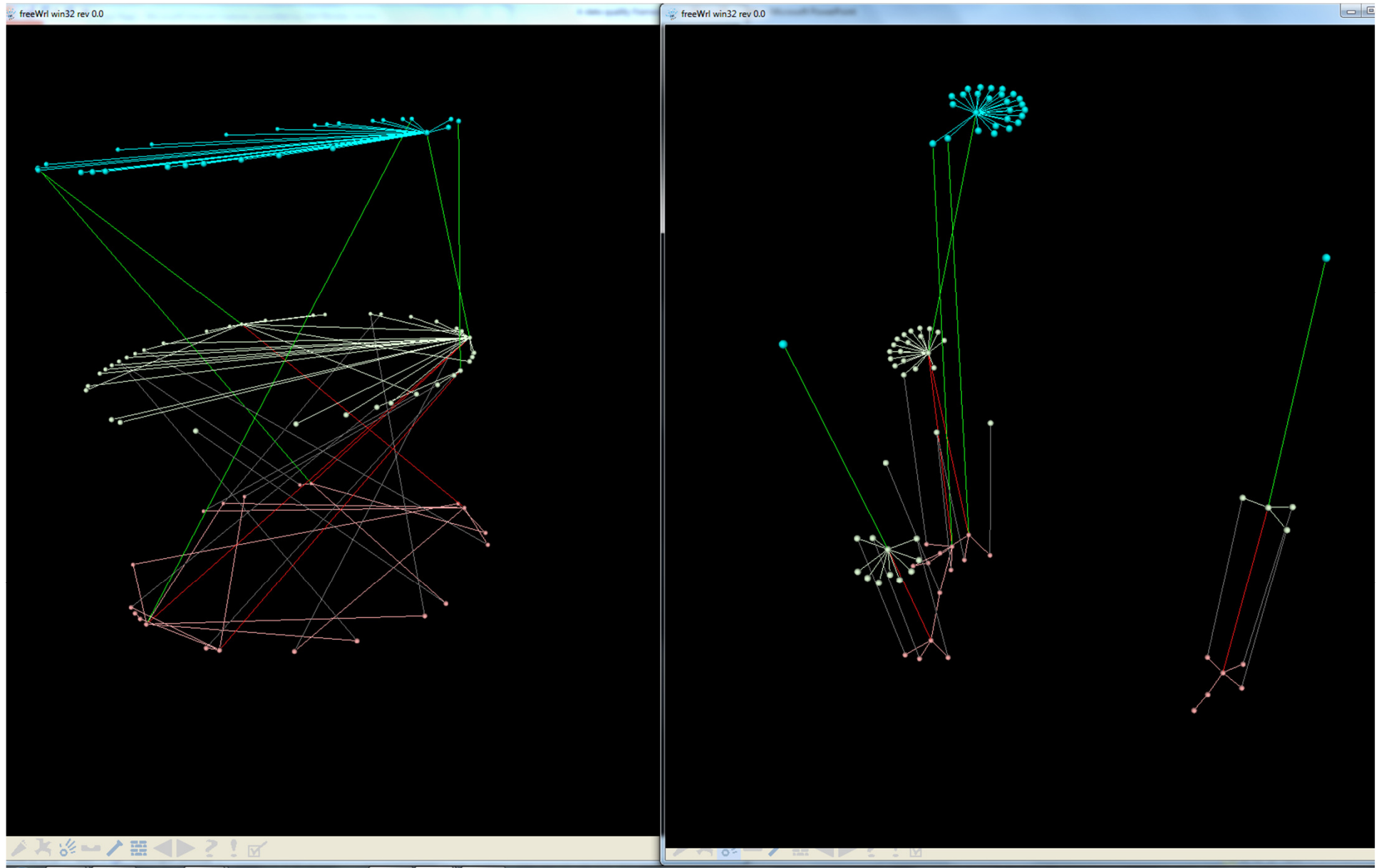


Visualization of metadata / concept models



3D visualization of metadata / ontologies

Below: two views of same model



Feedback loops using visual patterns

- Identify and correcting unwanted patterns
- Identify recommended patterns

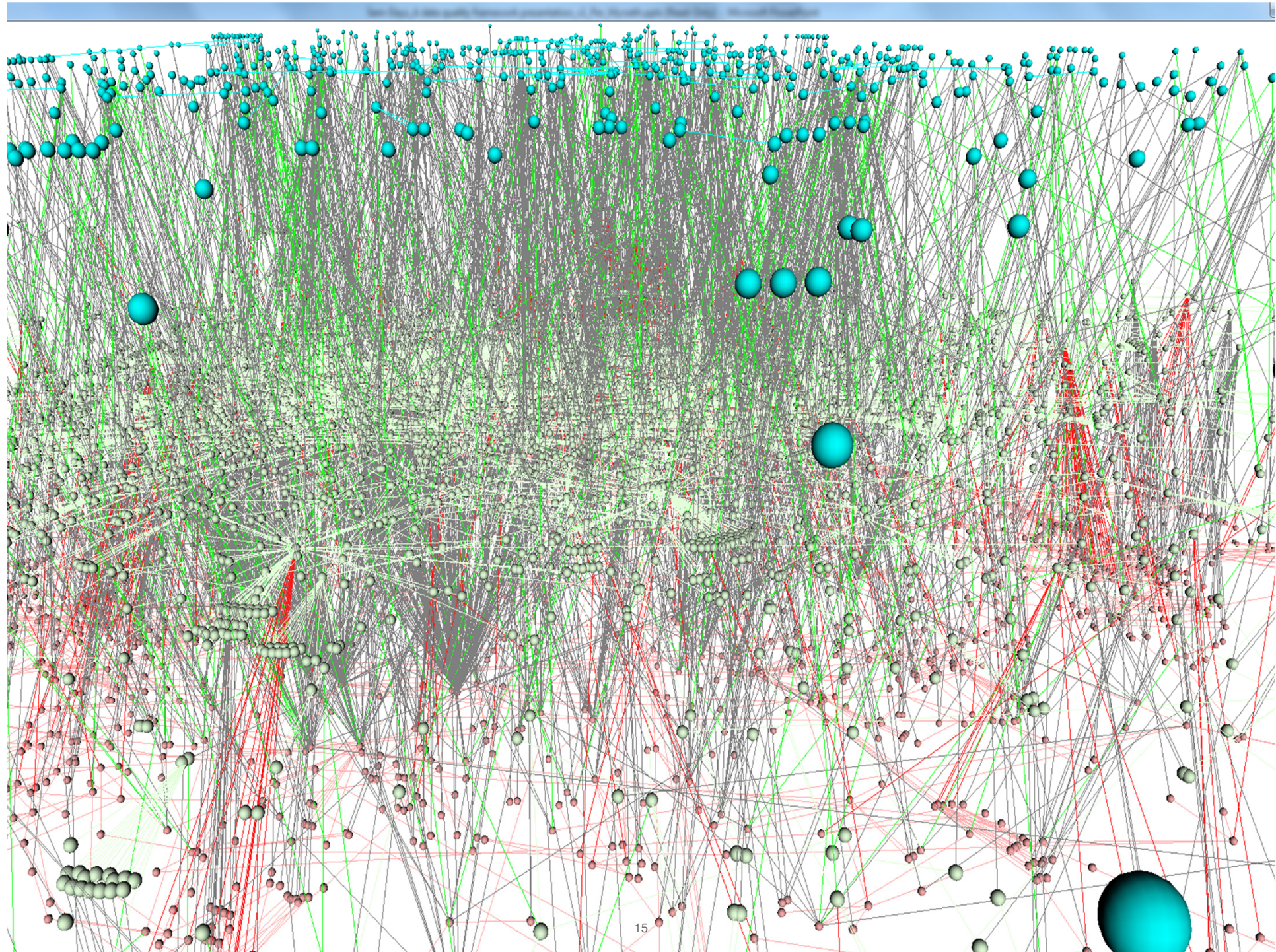
- Increased metadata quality
- New metadata / ontology engineering recommendations
- New automated quality validation rules
 - By calculating wanted and unwanted patterns
- Improved visual inspection guidelines

Framework: Categories of quality metrics for metadata/ conceptual models

	Syntax Schema	Semantics Meaning	Pragmatic Suited for use
Concepts and definitions + relations between terms	✓	✓	✓
Structure Reusable groups of concepts	✓		
Implementasjon UML of a XML-Schema	✓		

Background documentation:

- A data quality framework applied to e-government metadata.
 - Per Myrseth, Jørgen Stang and Vibeke Dalberg.
 - The International Conference on E-Business and E-Government (ICEE2011)
 - Shanghai 7. – 9. th of May 2011
 - Visualization of Complex Relations in E-Government Knowledge Taxonomies.
 - Per Myrseth, Jørgen Stang and David Skogan.
 - International Conference Information Visualization. 12, 13 - 15 July 2011, London.
 - A lifecycle process for engineering and maintaining concept models /ontologies
 - Norwegian title: Livssyklus prosess for utarbeidelse og forvaltning av begrepssystemer.
 - Vibeke Dalberg, Per Myrseth and Jim Yang
 - www.semicolon.no report februar 2011
-



Safeguarding life, property and the environment

www.dnv.com

Semicolon II

