Florence
Fraud Detection System
How to discover new possible tax evaders using the Oracle 12c Spatial and Graph option.

Claudia Corcione
Project Manager & Senior Ontologist

Silvia Naro
Software & Data Engineer
• Context overview – Tax Fraud Detection
• Technical solution
• System demonstration
How do Italian municipalities usually carry out Tax Evader Detection?

- Huge amount of data
- Heterogeneous data sources
- Data integration performed by humans
- Tax evader patterns detection performed by humans
How can we simplify the detection process?
Pinpoint suspects among the citizens

- Merging different data sources in a central knowledge base
- Exploiting the relationships between data
- Inferring implicit knowledge
- Embedding the domain knowledge inside an expert system
The solution: a Semantic Decision Support System

Relational Database → Abox RDF Graph → Semantic Database → Sparql

Tbox

FRAUD DETECTION system
Semantic Decision Support System: the architecture

File system

DBMS

Application Server

File
T-box
N-triples

ORACLE 12c

ORACLE WEBLOGIC

FRAUD DETECTION system

R2RML

RIQA ADAPTER

March 20, 2018
1. Study of the use case domain (legal aspects)
2. Study of the data sources (structure and contents)
3. Ontology design and logical rules creation
4. Mapping from the relational structure to the graph structure
5. Implementation of User-defined rules
6. Reasoning execution (the entailments generation)
7. Customization of Joseki Endpoint (with a set of pre-written queries)
Restricting the domain

To build a universal Tax Fraud Detection System...

...IT’S TOO EXHAUSTING!

Too many ranges, ontologies and rules

• Let’s focus on AIRE, the Official Register of Italian residents abroad
When can an AIRE Citizen be considered a possible evader?

When it is possible to detect a presence in the municipality territory

The Municipality of Florence knows how to do this...

...it owns the right hints: its DATA!
Use case: AIRE-Italians resident abroad
Possible reasons of suspicion

• High level suspect

School Attendance over 183 days/year for a registered AIRE minor citizen

• Medium level suspect

A registered AIRE citizen who presented a request for the occupation of public land

• Low level suspect

A registered AIRE citizen with a resident gas utility
Given an AIRE citizen $C$

- $C$ is a possible evader of level **Medium** if
- $C$ presented 2 or more construction requests in a year
insert into mdsys.semr_fraud_rule_15
values ('regola_15_2',

(?atto rdf:type fi:AttoNumeroAltoPratiche)
(?atto fi:relativoAlCittadino ?cittadino)
  (?atto fi:anno ?anno)
(?cittadino fi:haScheda ?scheda)
    (?scheda fi:annoInizioValiditaScheda ?annoiniziovalidita)
    (?scheda fi:annoFineValiditaScheda ?annofinevalidita)
(?scheda fi:statoAnagrafe "Iscritto A.I.R.E") ',
' (anno>=annoiniziovalidita and anno<=annofinevalidita)',
' (?cittadino fi:sospettatoPer fi:r15)' );

SEM_ALIASES(SEM_ALIAS('fi','http://www.firenze.fi.it/frodiFiscali.owl#'),SEM_ALIAS('rdf','http://www.w3.org/1999/02/22-rdf-syntax-ns#'),SEM_ALIAS('rdfs','http://www.w3.org/2000/01/rdf-schema#'));
FRAUD DETECTION system
Which are the pros?

- Exploitation of data already held by municipalities
- Centralized management of data
- Inclusion of all the necessary domain knowledge into a system
- System Reusability

Optimization of the suspects detection phase in the global process of Fraud Detection.
Thank you for your attention!

Q&A Time...
evodevo

Discovering, Together

www.evodevo.it