Investigator’s Toolkit
Ontology Summit 2017
15 March-2017
Overview

- Why we use ontologies
- Natural language processing
- Using the ontology in regulatory contexts
- Live demonstration
  - Delivering rapid situational awareness from source texts
It is not bulk collection that causes damage but the absence of tools that can extract the needle from the haystack.

French intelligence ‘could have prevented Paris attacks’

A former US intelligence director turned whistleblower tells MPs and peers that bulk collection of internet and phone data is putting lives at risk.

The French intelligence services could have prevented the November 2015 Paris terrorist attacks, a former US intelligence chief told MPs and peers, write Fiona O’Cleirigh and Bill Goodwin.

William Binney, a former technical director of the US National Security Agency, told a parliamentary committee that intelligence agencies are missing important data on terrorist attacks because they are overwhelmed with electronic data.
Merge Information in Real Time from Different Languages

Information extracted from 4 documents:
Ja'far Raed Qiddis, also known as Jafar, جعفر is a brother of al-Qaeda supporter Abdullah Raed Qiddis, and he is planning to fly to the US next week, probably for the purpose of suicide bombing.
"Think, by analogy, of individuals living in a series of tall closed towers, all erected over a common foundation.

When they try to communicate with one another, they shout back and forth, each from his own closed tower. It is difficult to make the sound penetrate even the nearest towers, and communication proceeds very poorly indeed.

But, when an individual goes down his tower, he finds himself in a great open basement, common to all the towers. Here he establishes easy and useful communication with the persons who have also descended from their towers".
An ontology is a hierarchal map of concepts and specific instances and their attributes that clearly defines each instance and its relationship with other instances. Many expressions in different languages can all “mean” the same ontological instance, which, in turn, may be linked to different concept parents under different conditions.
Ontology Hierarchy

That’s a great deal to make one word mean,” Alice said in a thoughtful tone. “When I make a word do a lot of work like that,” said Humpty Dumpty, “I always pay it extra.”

-Lewis Carroll, Alice in Wonderland
Functional features within the Investigator’s Toolkit

- Natural Language Processing
- Entity Extraction & Analysis
- Ontological Analysis
- Categorization
- Relationship Analysis
- Sentiment

View & Analyze insights
Natural Language Processing is the analysis of natural language to extract grammatical and syntactic features that enable automated understanding of the text.

Morphological analysis examines each “word” in the document in order to identify each individual lexeme, its tense, number, and affixes.

Syntactic and Parts of Speech analysis identifies parts of speech and concatenates lexical units in the text to larger meaningful units.
Categorization

First tier categorization allows the user to focus on the texts that really are important without the need to invent ambiguous "keywords" and crawling rules.

- Arts & Culture
- Cyber
- Finance
- Food & Recipes
- Health & Medicine
- History
- IED
- International Politics
- Islamic
- IT
- Legal
- Military
- Physical Sciences
- Politics
- Religion
- Social Issues
- Sports
- Terrorism
- Tourism
Ontology-based Categorization is language-independent and provides the user with a common denominator for deep categorization of texts in different languages. A text may receive multiple categories in different category clusters.

Political or Ideological Orientation
Jihadi-Salafi/Democrat/Socialist

Priority
High / Medium / Low

Type
Fatwa / Purchase Order/Letter/News article
Entity Extraction & Analysis

- Identification of words in the text as named entities: persons, groups, organizations, locations, addresses, URLs, dates, bank accounts, ideas, actions etc.

- Extraction of information implicit in the names such as gender, ethnicity, status, relationships and other attributes of the entity.

IntuScan™ goes beyond traditional named entity recognition. It matches names in the text that relate to the same entity, identifies the linguistic origin of the name, reverses it to its source orthography, applies cultural naming conventions and statistical models to generate name variants and discovers information such as ethnic origin, gender, religion/sect, status, family/tribal links etc.
Sentiment Analysis

Identification of the ontological instances linked to the lexical tokens from different languages.

Identification of the semantic parents of the ontological instances

Aggregation of references from different sentiment holders towards attributes of each object.

Aggregation of information regarding sentiment towards specific entities to a general aggregated sentiment towards their “parent” entities

“dishonest”/"malhonnête”/ לח יָשָר dishonest. “smart”/"bright"/"doué"/ = “מבריק” intelligent

dishonest ∈ Honesty-negative
intelligent ∈ Intelligence-positive

A: Y is dishonest = Honesty-negative
B: Y is corrupt = Honesty-negative
C: Y is smart = Intelligence-positive
D: Y is brilliant = Intelligence-positive
Agreement: Y is negative in honesty and positive in intelligence.

A: W is dishonest = Honesty-negative
B: X is corrupt = Honesty-negative
C: Y is disingenuous = Honesty-negative
D: Z is shifty = Honesty-negative
W, X, Y, Z ∈ K (ontologically pre-identified group)
Agreement: K is negative in honesty
Other Potential Uses and Opportunities

Defence
Security
Military
Border
Control

Legal
Discovery
and
Disclosure

Full
Commercial
exploitation
rights
including
Govt

Social
Media
Analysis

Political
Campaigns

- Law Firms, Organisations
- Internal Governance, M&A, PII Audit
- Corporate Intelligence
- Communications Surveillance

- Brand Agencies
- Advertising agencies
- Organisations

- PR Agencies
Psonify – The Investigator’s Toolkit

• **Technology – mining big data:** Our military precision engine enables us to find the signals through a the large amount of noise well beyond keywords and counting phrases. Our patented protected technology allows us to mine for the “context” and our sentiment engine turns it into quantifiable results which are automatically “pushed” to the customer.

• **Semantics:** We understand English, French, Spanish, Arabic, Russian and many more. This is not a translate tool but extraction of “meaning” and applies not just well written news articles but social media with its specific semantics.

• **Expert Knowledge:** Core to our semantic engine is a deep ontology of over 1.5million instances that contains a knowledge base across a number of domains including finance, law and regulations. Our team are dedicated experts across data science, linguistics, financial services and the traded markets.
Psonify: Our Three Pillars

Psonify has assembled a team of experts from various backgrounds to develop first class regulatory technology solutions.

**Regulation**
Expertise in Financial Services: Team Members worked in bulge bracket investment banks for extended period of time.

**Technology**
Expertise in Text Recognition Technology: an established track record in developing high precision semantic text analysis.

**Semantics**
Expertise in Ontology and Business Semantics: Team members co-chair various groups and develop significant parts of Financial Ontologies.
RegTech: Objectives

1. Interpret Regulation
   - show capability for multiple interpretations by different Clients/Client Types
   - show capability for a Master Interpretation

Achieved via the Content Manager with configuration individual configuration

2. Mine for Contextual Meaning and Identify Actionable Tasks (and Non-Actionable Tasks)
   - Duties – positive and negative
   - Permissions = FYI
   - On whom – Entity, Departmental and Individual
   - Identifying the N/A’s

Achieved via the existing Machine Learning capability within the Platform based on individually customized Financial Services Ontology

Requirement: Full Ontology implementation
Commercial Application: Apply the Workflow

Challenge x 3
Where to draw the line
Who can support us
How do we share

Client
Terminological Dictionary

Content Manager

Ontology SME:
Analyze, Understand, Action

Regulatory Interpretation:
Analyze, Understand, Action

Psonify
Linguistic Processing
RegTech-Proof of Concept: Example

- **Example:** Aya v. 5 vs MiFIR Article 6 Post-Trade Transparency

- Line by Line Interpretation of a Regulation demonstrated in the Context of a line by line interpretation of the Koran

- These capabilities are within the Platform. Some configuration is required in the front end and in the Content Manager together with the full deployment of the slots for the Financial Services Ontology.

- In no way will this customisation affect the higher precision and recall rates which are offered with this military grade tool.

**Actionable Tasks in Example**

**Potential Actionable Tasks in Regulatory Framework**

- Duty
  - Duty-at-entity-level
  - Duty-at-department-level
    - Client-reporting-department
    - Regulatory-reporting-department
    - ...
- Prohibition
  - Duty-at-entity-level
  - Duty-at-department-level
  - ...
- Permission
On-Screen Example

"Dealing with Idol Worshippers - KORAN"

Koran Aya 5
Once the Sacred Months are past, (and they refuse to make peace) you may kill the idol worshipers when you encounter them, punish them, and resist every move they make. If they repent and observe the Contact Prayers (Salat) and give the obligatory charity (Zakat), you shall let them go. GOD is Forgiver, Most Merciful.

MiFIR Article 6
Post-trade transparency requirements for trading venues in respect of shares, depositary receipts, ETFs, certificates and other similar financial instruments
1. Market operators and investment firms operating a trading venue shall make public the price, volume and time of the transactions executed ...

MiFIR Article 7
1. Competent authorities shall be able to authorise market operators and investment firms operating a trading venue to provide for deferred publication of the details of transactions based on their type or size.

MiFIR Article 10, 11, 20 ...

1. Master Interpretation and Multiple Interpretations

Shiite Interpretation
Sunni Interpretation
Radical Interpretation
All types Interpretation

Bank A: Standard Interpretation
Bank B: Permission for deferred publication (Article 7)
Asset Manager C: ...
All Clients/Client Types: Master Interpretation

2. Capturing the Actionable Rule

radical-interpretation:
• acquisition-of-weapons-means
• permission-indiscriminate-killing
• permission-annul-treaties-with-infidels

department-client-reporting:
• duty-publication-price
• duty-publication-volume
• duty-publication-time

moderate-interpretation:
• duty-restraint-prohibition-excess
Psonify: Thank you

Simon Davidson

Simon Davidson, Founder and CEO of Psonify is a seasoned business professional with a unique blend of legal, regulatory, big data and commercial skills. A former investment banker at Credit Suisse and corporate lawyer at Slaughter & May in London, Simon has spent the last 20 years working in the financial services and corporate services sector. Simon has held a number of board seats in the fintec and market structure area including Markit.

Contact:

simon@psonify.com
+44 7956 437 442
Disclaimer

These materials have been provided to you by Psonify Forensic Ltd in connection with an actual or potential mandate or engagement and may not be used or relied upon for any purpose other than as specifically contemplated by a written agreement with Psonify Forensic Ltd. In addition, these materials may not be disclosed, in whole or in part, or summarized or otherwise referred to except as agreed in writing by Psonify Forensic Ltd.

This material is not investment research or a research recommendation for the purposes of Financial Services Authority rules as it does not constitute substantive research or analysis.

The information provided is not intended to provide a sufficient basis on which to make an investment decision. It is intended only to provide observations and views of the said Psonify Forensic Ltd personnel. Observations and views of Psonify Forensic Ltd personnel may change at any time without notice.

Information and opinions presented in this material have been obtained or derived from sources believed by Psonify Forensic Ltd to be reliable, but Psonify Forensic Ltd makes no representation as to their accuracy or completeness. Psonify Forensic Ltd accepts no liability for loss arising from the use of this material. The information used in preparing these materials was obtained from or through you or your representatives or from public sources. Psonify Forensic Ltd assumes no responsibility for independent verification of such information and has relied on such information being complete and accurate in all material respects. To the extent such information includes estimates and forecasts of future financial performance (including estimates of potential cost savings and synergies) prepared by or reviewed or discussed with the managements of your company and/or other potential transaction participants or obtained from public sources, we have assumed that such estimates and forecasts have been reasonably prepared on bases reflecting the best currently available estimates and judgments of such managements (or, with respect to estimates and forecasts obtained from public sources, represent reasonable estimates). This document may contain extracts from various media/news sources, which are provided for your information only. Psonify Forensic Ltd makes no representation as to the accuracy or validity of these extracts and does not endorse the content thereof.

These materials were designed for use by specific persons familiar with the business and the affairs of your company and Psonify Forensic Ltd assumes no obligation to update or otherwise revise these materials.

Nothing contained herein should be construed as tax, accounting or legal advice.

The information contained herein is correct as at the date of publication and is subject to change without notice.