



# THOMSON REUTERS BOLD

## Quantitative Finance @WORK

Douglas McKinley  
Tor Vergata Rome 04/05/2018

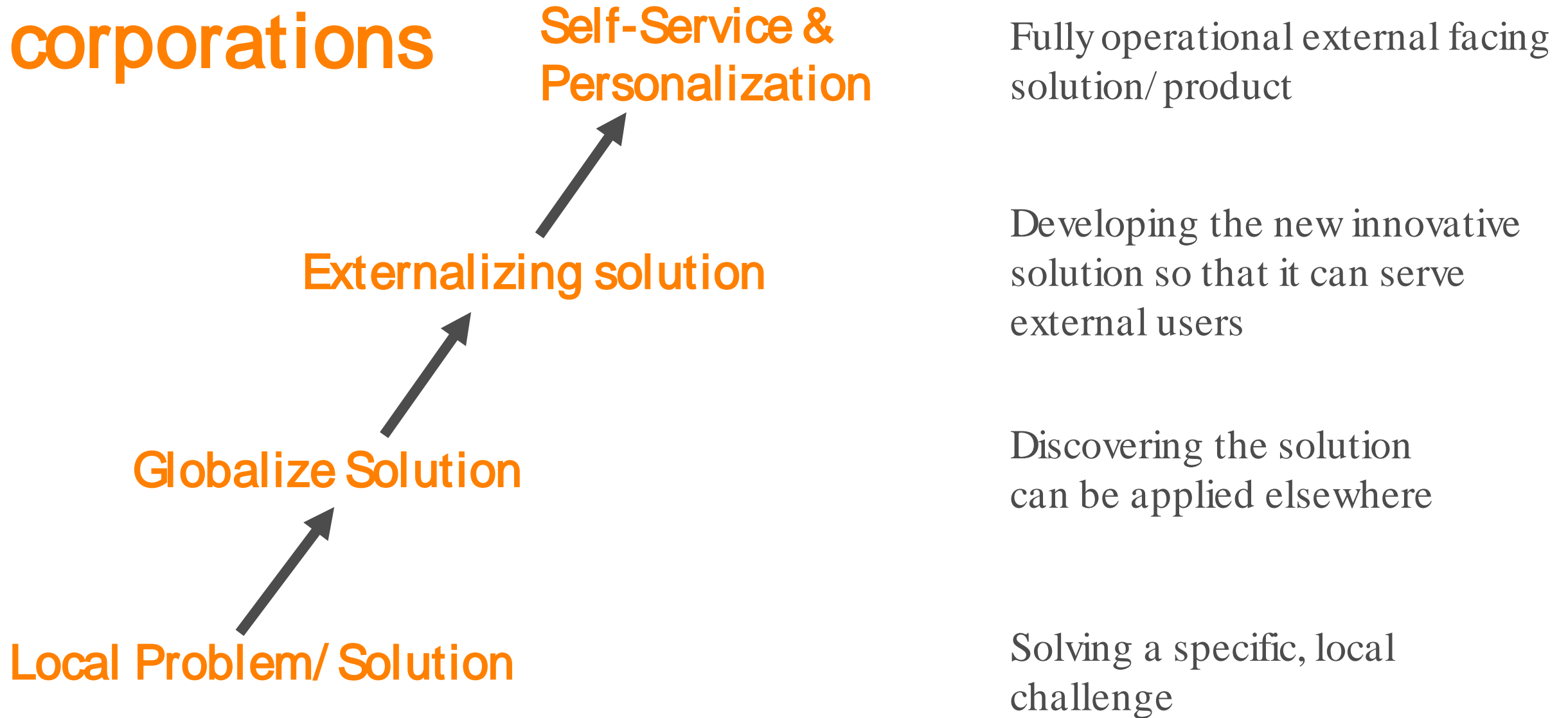
The intelligence, technology and human expertise  
you need to find trusted answers.



the answer company™  
**THOMSON REUTERS®**

How does  
innovation start?

# Innovation – follows a certain path within large corporations



(big) Data is an  
example

# Standard challenges with (big)data...

- 1. Extraction**
- 2. Connection**
- 3. Navigation**

# Data Data Everywhere - **the key is to link the data**

*"Any [Chief Data Officer] really ought to be able to ask a question that involves connecting data across the organization, be able to run an [enterprise] effectively, and especially to be able to respond to unexpected events. Most organizations are missing this ability to connect all the data together."*

Sir Tim Berners-Lee, The inventor of the World Wide Web and one of Time Magazine's 100 Most Important People of the 20th Century

# Exploit signals in News

News flow and sentiment are important sources of signals in quantitative stock selection and systematic trading. However news and social media data is unstructured and there is a lot of it. Is it positive or negative? New, old or recycled? Relevant to a company?

500,000 text news stories per day

7,000+ Text News Sources

540,000 video clips and growing

80+ video sources

1,600 pictures per day

34 languages

# Unstructured to Structured - News Analytics Output

News Item Id	Regional Timestamp	Co	PermID	Attrib	Item Type	Headline	Rel	Sent	Pos	Neut	Neg
CEN34MqtG_161019	2016-10-19T01:05:21.044Z	TCKb.TO	4295861257	MKW	ARTICLE	Teck to Acquire 100% of Teena/Reward Zinc Project<TCKb.TO>			0.70	0.18	0.11
FWN1CO0CV_161019	2016-10-19T07:13:03.655Z	WFC.N	8589934175	RTRS	ALERT	WELLS FARGO & CO <WFC.N> :FBR CUTS TO MARKET PERFORM FROM OUTPERFORM	1.00	-1	0.06	0.13	0.82
FWN1CP0EI_161019	2016-10-19T09:46:23.203Z	TWTR.N	4296301199	RTRS	ALERT	TWITTER INC <TWTR.N> : LOOP CAPITAL RAISES TO HOLD FROM SELL - THEFLY.COM <	1.00	1	0.60	0.32	0.09
ZHN0BUV0T_161019	2016-10-19T19:45:00.828Z	AVP.N	4295903466	RTRS	ARTICLE	NYSE ORDER IMBALANCE <AVP.N> 205700.0 SHARES ON SELL SIDE<AVP.N>	0.50	0	0.18	0.59	0.23
L8N1D96VH_161109	2016-11-09T04:50:11.802Z	VLLP.PA				French and Benelux stocks-Factors to watch on Nov. 9<CNAT.PA><EDF.PA><VLLP.PA>	0.24	-1	0.07	0.14	0.79
L4N1DA09T_161109	2016-11-09T08:50:57.806Z	RIO.AX	4295856917	RTRS	ARTICLE	BREAKINGVIEWS-Rio Tinto trips into bribery mine shaft in Africa<RIO.AX><R					
L4N1DA09T_161109	2016-11-09T10:09:30.182Z	RIO.AX	4295856917	RTRS	ARTICLE	BREAKINGVIEWS-Rio Tinto trips into bribery mine shaft in Africa<RIO.AX><R					
L8N1D95GR_161109	2016-11-09T17:55:19.976Z	GS.N	4295911963	RTRS	ARTICLE	EXCLUSIVE-Goldman Sachs considers Frankfurt move over Brexit - sources<GS.N>	1.00	-1	0.11	0.29	0.60
FWN1DA0JB_161109	2016-11-09T23:28:01.245Z	CTS.N	4295903630	RTRS	ALERT	CTS CORP <CTS.N>: B. RILEY STARTS WITH BUY; TARGET PRICE \$22<CTS.N>	1.00	1	0.55	0.44	0.01
IFR20h64D_161116	2016-11-16T01:44:11.485Z	SGXL.SI	4298007743	IFR	ARTICLE	UPDATE: Rickmers misses coupon payment, suspends t	0.20	-1	0.06	0.13	0.82
FWN1DG129_161116	2016-11-16T06:07:49.218Z	CRDI.MI	4295875726	RTRS	ALERT	UNICREDIT SPA <CRDI.MI>: HSBC RAISES TARGET PRICE TO 2.85 EUROS FROM 2.83 EU	1.00	1	0.56	0.43	0.01
L8N1DH37L_161116	2016-11-16T11:32:23.557Z	RDSa.L	4295885039	RTRS	ARTICLE	BUZZ-Shell announces nearly 400 job cuts in Scotland<RDSa.L>	1.00	-1	0.06	0.13	0.82
FWN1DH0M5_161116	2016-11-16T15:03:11.091Z	NDAQ.O	8589934167	RTRS	ALERT	NASDAQ HALTS DRYSHIPS INC.<NDAQ.O>	1.00	0	0.18	0.81	0.01

Strong Positive Sentiment

Story updated 89 mins after previous version

Low Relevance

Strong Negative Sentiment

1st Mentn	Tot Sentcs	# of COS	# Sent Wds/Tkns	Tot Wds/Tkns	Volume Count	Novelty Count	Exchange Action	Broker Action	Price Target	Market Commentary	Topic Codes
1	21	1	430	490	1 1 15 25 33	0 0 0 0	UNDEFINED	UNDEFINED	UNDEFINED	FALSE	AMERS BMAT CA CMPNY MEMI MIN MINE MTAL NAMER NEWR
1	1	1	13	13	30 50 81 202 261	0 0 8 20 25	UNDEFINED	DOWNGRADE	UNDEFINED	FALSE	AMERS BANK BISV BLR BSVC CMPNY FINS INVM INVS NAMER RCH US WLTH
1	1	1	13	13	2 3 7 22 39	0 0 0 0	UNDEFINED	UPGRADE	UNDEFINED	FALSE	AMERS BLR CMPNY ITSE NAMER RCH SWIT TECH TMT US
0	2	1	17	17	0 0 0 0	0 0 0 0	IMBALANCE	UNDEFINED	UNDEFINED	FALSE	AGA AMERS CMPNY NAMER NCYC PEPS PPRO US
7	49	3	36	289	3 8 8 8 8	0 0 0 0	UNDEFINED	UNDEFINED	UNDEFINED	TRUE	EUROP EZC FINS FR INVD INVS NL OILQ REP S
1	42	2	859	912	22 24 38 40 42	2 2 2 2	UNDEFINED	UNDEFINED	UNDEFINED	FALSE	AFR AMERS ANV ASIA ASXPAC AU AUNZ BACT BASMTL BMAT BR BRIB BRV BRVF
1	43	2	871	924	15 23 39 41 43	3 3 3 3	UNDEFINED	UNDEFINED	UNDEFINED	FALSE	AFR AMERS ANV ASIA ASXPAC AU AUNZ BACT BASMTL BMAT BR BRIB BRV BRVF
1	34	1	858	909	11 12 32 33 59	2 2 2 2	UNDEFINED	UNDEFINED	UNDEFINED	FALSE	AMERS BACT BANK BISV BRXT BSVC CEEU CEN CLJ CMPNY DE DIP ECB EU EURC
1	2	1	14	14	0 0 0 0	0 0 0 0	UNDEFINED	INITIATE	INITIATE	FALSE	AMERS BLR CMPNY ELCO INDG INDS MACH NAMER RCHUS
8	13	1	72	238	1 2 5 6 18	0 0 0 0	UNDEFINED	UNDEFINED	UNDEFINED	FALSE	ASEAN ASIA ASXPAC BACT BISV CDM CMPNY CORPD DBT DBTR EMRG EUB EXCA
2	1	1	17	17	1 15 56 85 113	0 1 2 4 4	UNDEFINED	MAINTAIN	INCREASE	FALSE	BANK BISV BLR BSVC CMPNY EUROP EZC FINS IT RCH WEU
3	1	1	108	108	5 19 41 49 73	0 0 0 0	UNDEFINED	UNDEFINED	UNDEFINED	FALSE	BACT BISV CMPNY COM ENER EUROP FINS GB GBS JOB LAYOFS NRG OILG REAL
1	1	1	5	5	0 0 0 0	0 0 0 0	HALT	UNDEFINED	UNDEFINED	FALSE	AMERS BISV BLR CMPNY EXCA FINS INVS NAMER US

Company mentioned in second half of the story

Broker Action w/ Price Target in different direction



# BOLD Solutions

Big Open Linked Data Solutions is a range of data integration, structured and unstructured data management tools and a powerful set of analytical tools built upon the Thomson Reuters' expertise and foundation technologies in content management.

# Intelligent Tagging – Structuring the Unstructured

## Unstructured Content

Content uploaded from news articles, blog postings, proprietary data, catalogs, social media, and more.



Extract, classify, and tag metadata

### ACME Corp. Announces Big Expansion

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Maecenas porttitor congue massa. **ACME Corp.** magna sed pulvinar ultricies, purus lectus malesuada libero, sit amet commodo magna eros quis urna. Nunc viverra imperdiet enim. Fusce est. Vivamus a tellus.

Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Proin **ACME Corp.** Mauris et orci. Aenean nec lorem. In porttitor. Donec laoreet nonummy augue.

Suspendisse dui purus, scelerisque at, vulputate **ACME Corp.** nunc. Mauris eget neque at sem venenatis eleifend. Ut nonummy. Fusce aliquet pede non pede. Suspendisse dapibus lorem pellentesque magna. Integer nulla.

Donec **ACME Corp.** ligula. Donec hendrerit, felis et imperdiet euismod, purus ipsum pretium metus, in lacinia nulla nisl eget sapien. Donec ut est in lectus consequat consequat. Etiam eget dui. **ACME Corp.** volutpat. Sed at lorem in nunc porta tristique.

Proin nec augue. Quisque aliquam tempor magna. Pellentesque habitant morbi tristique senectus et netus **ACME Corp.** fames ac turpis egestas. Nunc ac magna. Maecenas odio dolor, vulputate vel, auctor ac, accumsan id, felis. Pellentesque cursus sagittis felis.

Pellentesque porttitor, velit lacinia egestas auctor, diam eros tempus arcu, nec vulputate augue magna vel risus. Cras non **ACME Corp.** adipiscing rhoncus. Vivamus a mi. Morbi neque. Aliquam erat volutpat. Integer ultrices lobortis eros.

Donec blandit feugiat ligula. **ACME Corp.** hendrerit, felis et imperdiet euismod, purus ipsum pretium metus, in lacinia nulla nisl eget sapien. Donec ut est in lectus consequat **ACME Corp.** Etiam eget dui. Aliquam erat volutpat. Sed at lorem in nunc porta tristique.

Proin **ACME Corp.** Quisque aliquam tempor magna. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas.



Companies



People



Relationships



Topics



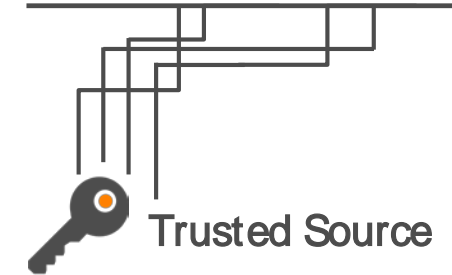
Locations



Events

## Structured Metadata

Our unique identifiers leverage the deep knowledge in Thomson Reuters professional data, creating metadata to enrich your own content – and also mapping it to Thomson Reuters content to give the best of both worlds.



The Thomson Reuters key advantage is assigning unique identifiers, or PermIDs, which go beyond keywords, returning the right connections you'd otherwise miss.

# PermID – A Barcode for Information



**Currency**  
Canadian Dollar  
PermID: 500140



**Instrument**  
TR Ord Shares  
PermID: 85909928696



**Quote**  
Primary Ticker – TRI  
Primary Exchange – TSX  
Primary RIC – TRLTO  
PermID: 55838860337



**Asset Class**  
Ordinary shares  
PermID: 300281



**Geography**  
Canada  
PermID: 100052



**Organization**  
Thomson Reuters Corp  
PermID: 4295861160



**TR Industry Classification**  
Professional Information Services (NEC)  
PermID: 4294951759

Unique

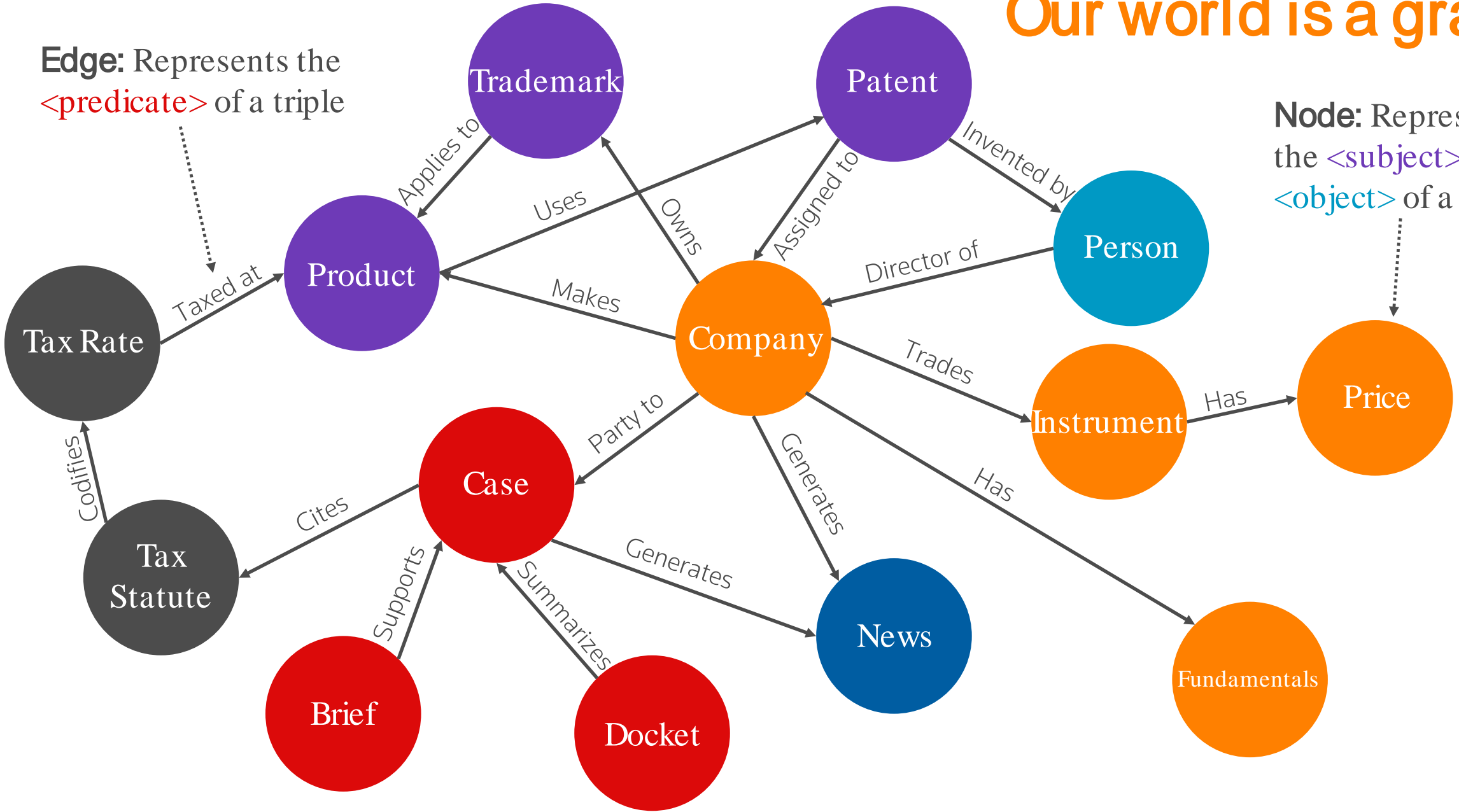
Permanent

Key identifier

# Our world is a graph

**Edge:** Represents the **<predicate>** of a triple

**Node:** Represents the **<subject>** or the **<object>** of a triple



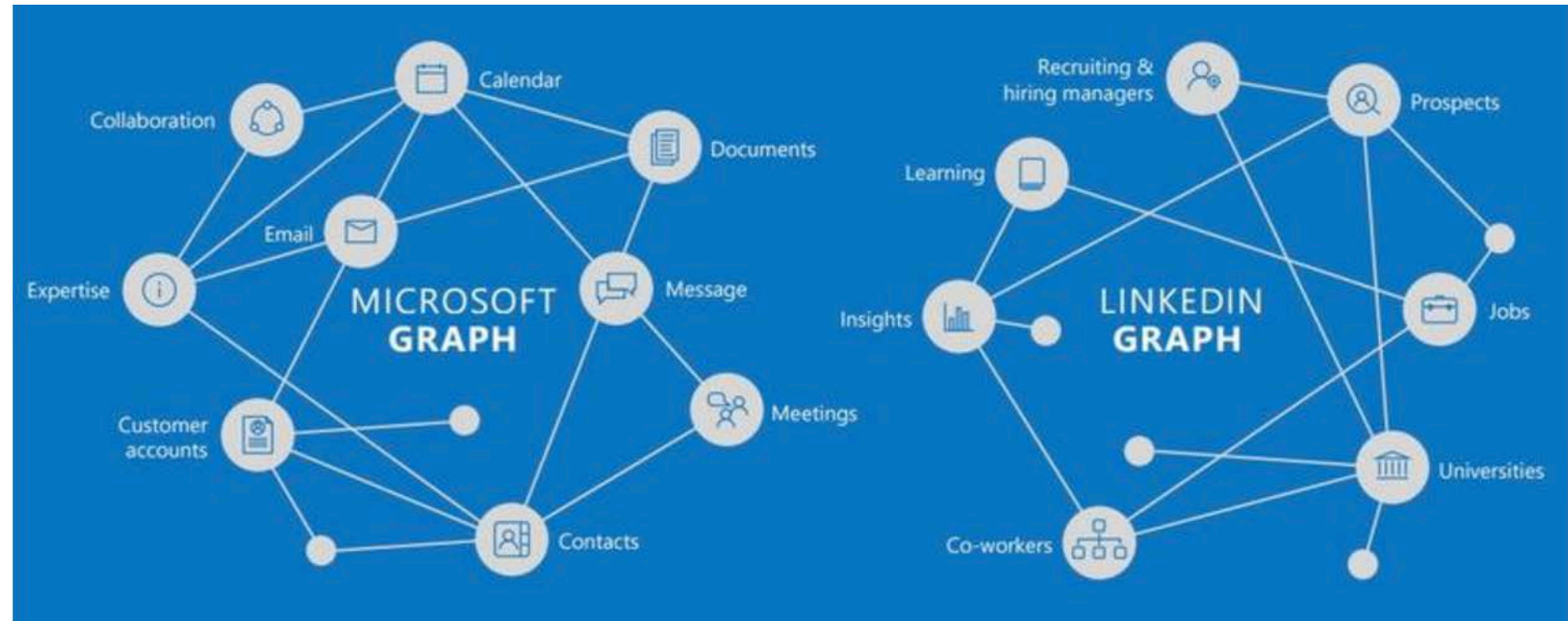
# Knowledge Graphs – Famous Examples

## Building The LinkedIn Knowledge Graph

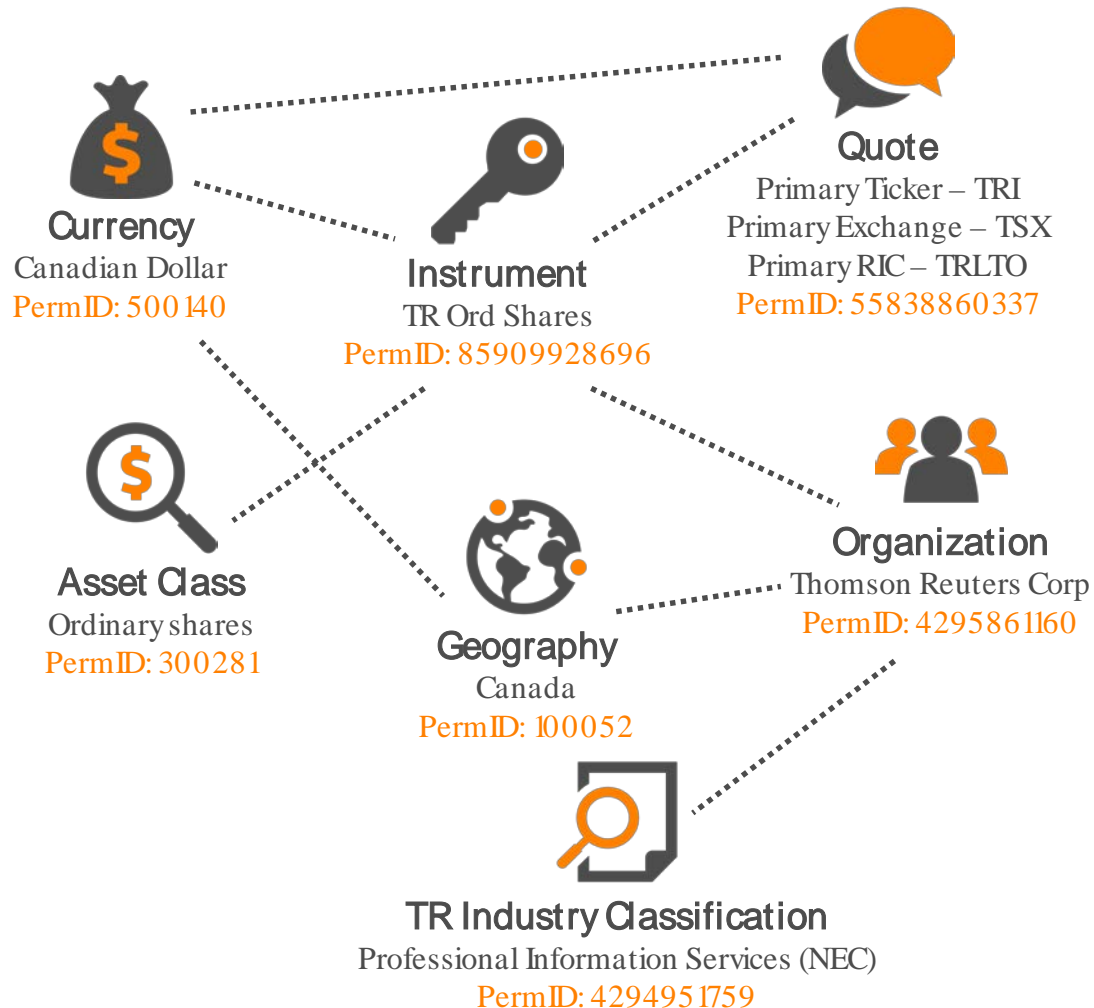
LinkedIn’s knowledge graph is a large knowledge base built upon “entities” on LinkedIn, such as members, **jobs, titles, skills, companies, geographical locations, schools**, etc.

These entities and the relationships among them form the **Ontology** of the professional world as LinkedIn see its

Used in its **recommender systems, search, monetization and consumer products**, and **business and consumer analytics**.



# TR Knowledge Graph – A Wealth of Knowledge



## Connect the dots

- Capturing relationships and entities from TRIT and Thomson Reuters Data
- Aggregating facts about entities over time
- Rely on a rich and standardized ontology of information

# TR Data Fusion – Graph Analytics

- Data Fusion integrates the Knowledge Graph with your data, allowing you to visualize and explore relationships
- Graphs can grow very quickly and become unreadable
- We have developed tools to filter out the noise based on a relevance score

The screenshot displays the Thomson Reuters Data Fusion interface. The top navigation bar includes 'Paths' and 'Entities'. The main search area shows a query for paths from 'United States (Country)' to 'Switzerland (Country)', excluding entities with more than 1,000 connections. The interface is divided into several panels:

- Left Panel:** Controls for the graph search, including 'Paths' (set to 100), 'LEVELS' (set to 2), 'LAST UPDATED' (set to 'ANY TIME'), 'SHORTEST PATHS' (with 'Shortest paths only' selected), and 'Entities' (set to 100). It also shows 'TYPES' with 'Deal' selected.
- Center Panel:** A large, dense graph visualization showing a complex network of nodes and edges, representing the relationships between entities.
- Right Panel:** A detailed view of an article titled 'RPT-Dell software deal flexes Elliott's newly developed buyout muscle (DI,UL) (EMC,N)'. It includes a 'News' section, a 'Company' dropdown set to 'VMware Inc', and a 'CONFIGURE WEIGHTS' section with various filters like 'Activity', 'Economic sector', 'Industry group', etc. Below the article text is a 'KEYWORDS' section.
- Far Right Panel:** A 'GRAPH' view showing a simplified network of nodes and edges, with a 'SCORE SUMMARY' section below it. The summary includes 'Relevance to article: High (6.2%)' and 'THE ARTICLE IS ABOUT:' followed by a list of entities and their relationships.

# Extract signal & insights - From unstructured content noise

## Unique barcode identifier for information

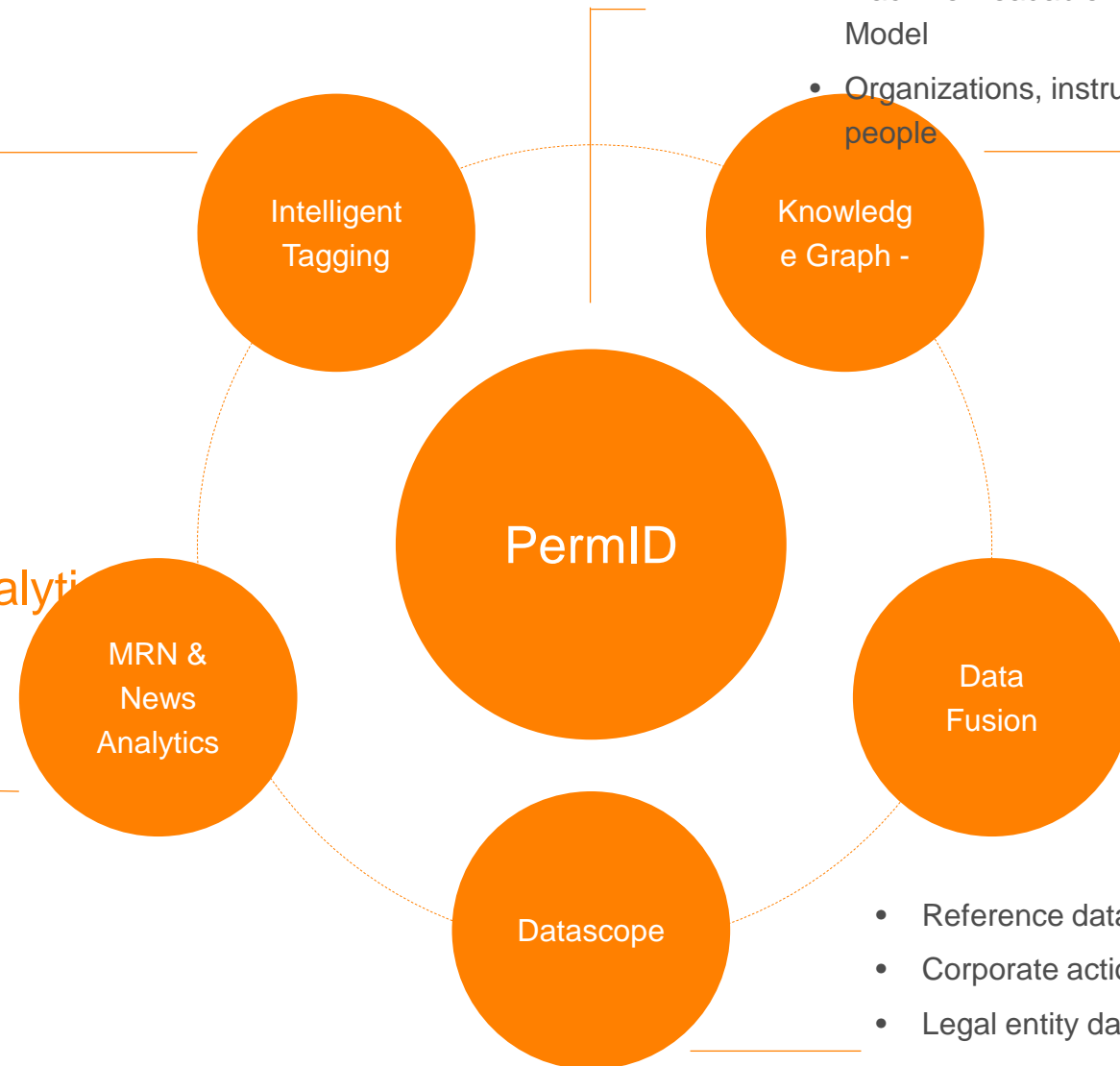
- Machine Readable Identifier for TR Information Model
- Organizations, instruments, funds, issuers and people

## Stored relationships of data

- Comprehensive relationship-linked information ecosystem of different TR content sets
- Tens of billions of triples
- News metadata, Supply Chain, Org Authority, People Authority, Officers & Directors
- RDF (resource description framework) format

## Big Data Ingestion & Analytics

- Ingest, stitch map and index vast disparate datasets
  - Uncover relationships
  - TR Knowledge Graph
  - Proprietary data
- Reference data
  - Corporate actions
  - Legal entity data
  - End-of-day/intra-day pricing
  - Evaluated pricing services



PermID

Intelligent Tagging

Knowledge Graph -

Data Fusion

Datascope

MRN & News Analytics

## Unstructured data analytics

- The fastest, easiest, and most accurate way to tag the people, places, facts, and events in your content to increase its value, accessibility, and interoperability.
- Appends the PermID to extracted content – unique to the financial industry

## Thomson Reuters News Analytics

- Sentiment, relevance and novelty scoring of Reuters & 3rd party regulatory news
- 47000 companies
- 40+ commodities topics
- Document-level scoring for macro, political and general news
- Real-time. Archive back to 2003.



Use the (free) PermID demo  
PermID.org

# Grazie! Domande?



didmckinley

[www.linkedin.com/in/douglasmckinley](http://www.linkedin.com/in/douglasmckinley)