



The
University
Of
Sheffield.

Knowledge Capturing, Retrieval and Reuse

Prof. Fabio Ciravegna

Director of Research and Innovation in the Digital World
University of Sheffield

and

Professor of Language and Knowledge Technologies
Department of Computer Science
University of Sheffield

fabio@dcs.shef.ac.uk

<http://www.dcs.shef.ac.uk/~fabio/>



Outline

- Semantic Web and Knowledge Management
 - ▶ From the Web to Corporate Applications
- Discovering Structured and Unstructured Information and Knowledge
- Structuring unstructured data
 - ▶ From Data to Information and Knowledge
 - ▶ Extracting Information from Texts
- Retrieving Knowledge and Information
 - ▶ Semantic Search
- Visualising Knowledge and Information
- Reusing Information



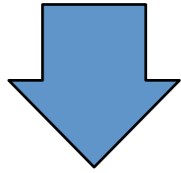
Knowledge Management

- The main goal of knowledge management is to allow companies ways to reappropriate of their knowledge
 - ▶ Which is generally implicit in
 - the mind of the employees
 - the procedures
 - Millions of documents
 - ▶ but rarely formalised explicitly.
- So in short the goal of knowledge management is to enable reuse of knowledge and know-how by the appropriate people at the right time in the right form

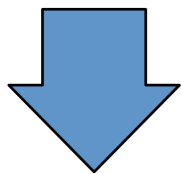


Why Manage Knowledge?

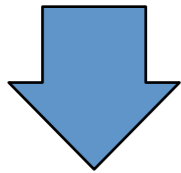
reuse



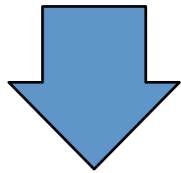
retrieval



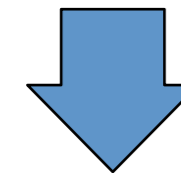
sharing



capture



modelling



acquisition

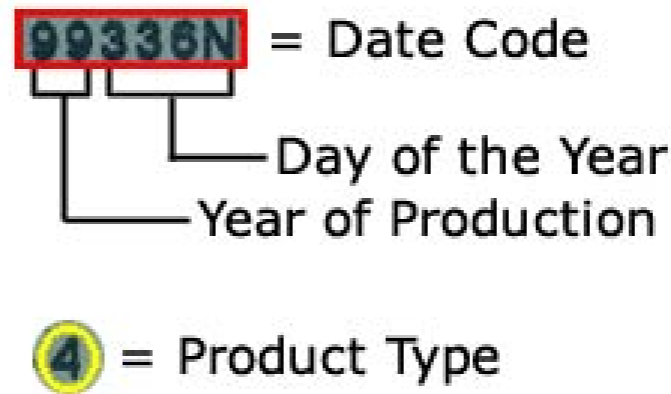
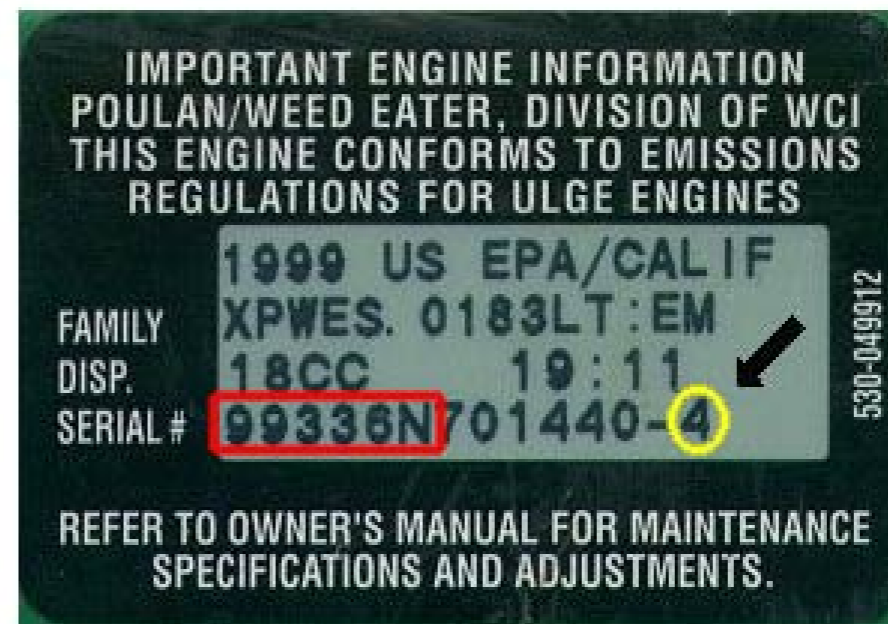
- To enable easy timely and effective **reuse** of knowledge
 - ▶ We need to: enable **retrieval**
- To enable retrieval
 - ▶ we need to: enable **sharing**
- To enable sharing
 - ▶ we need to: **capture** knowledge
- To enable capture:
 - ▶ We need **modelling** the domain and process in an appropriate way
- To enable Modelling:
 - ▶ We need **acquisition of** domain and process knowledge

jet engines are moving towards complete serialisation

- every piece has a serial number (excepts nuts and bolts)
- the history of each part is recorded
 - e.g. part transferred between engines



© Rolls-Royce plc



- a jet engine can produce ~1 Gbyte of vibration data per hour of flight;
 - if irregularities are found, part of the data can be stored
 - reports can be written (event reports)
 - pictures can be taken

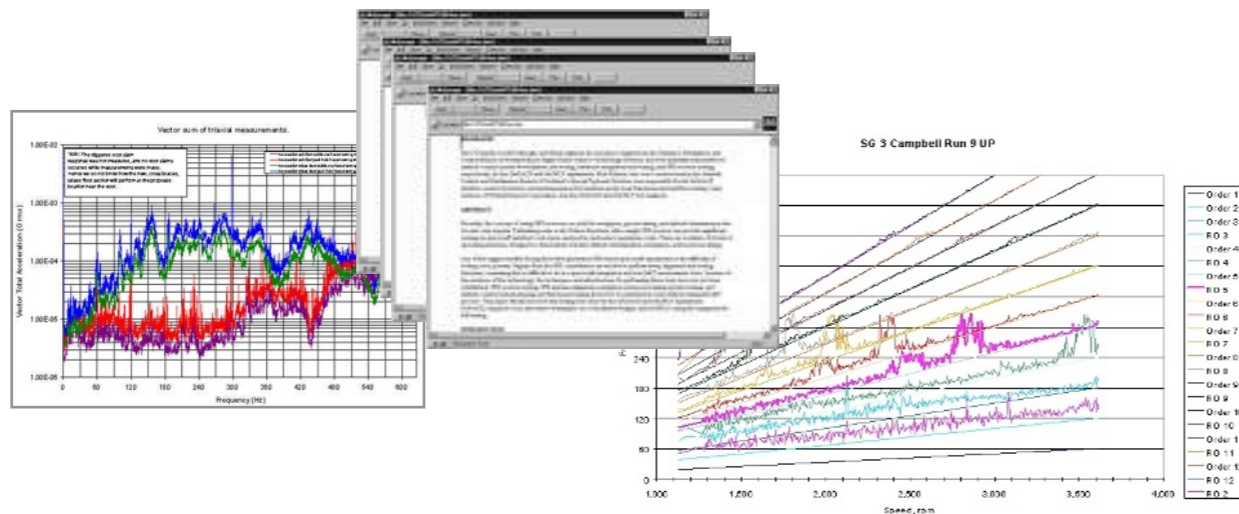


image © www.rolls-royce.com

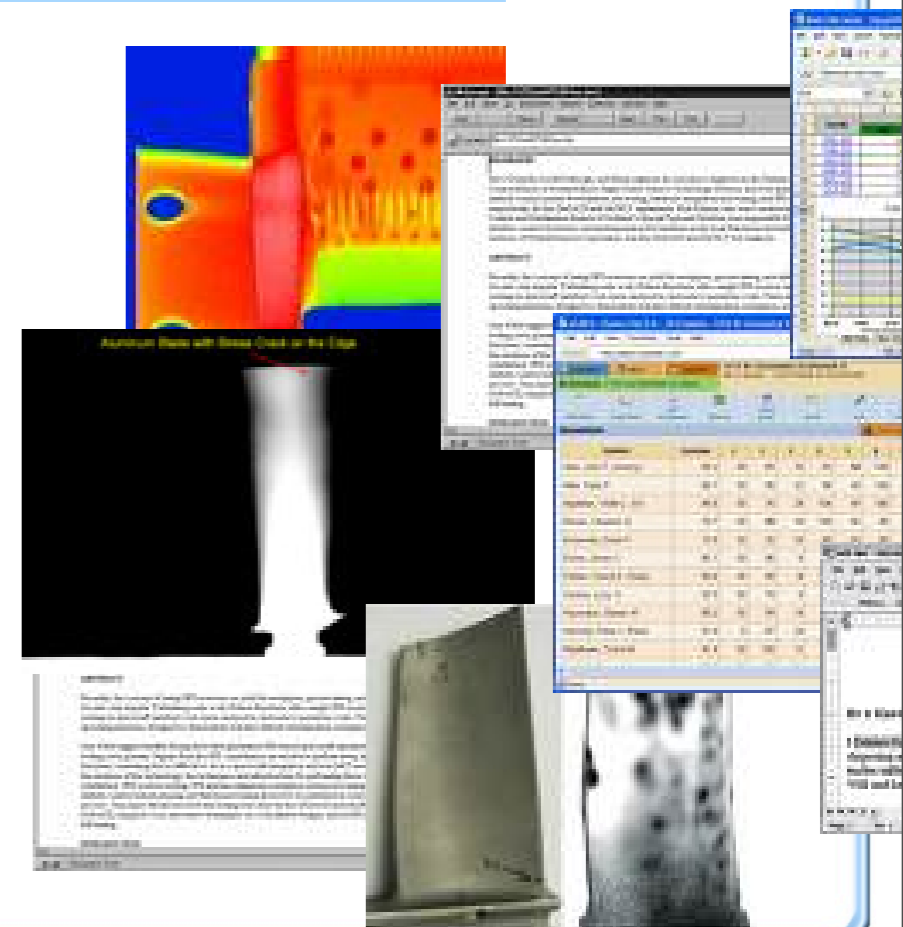


When engine is serviced (e.g. overhaul)

- financial information is produced.
- if issues are found,
 - pictures are taken
 - reports are written
 - engine is tested



image © Rolls-Royce plc



- If problem is recurring (or suspected so)
 - a problem resolution group is established
 - existing evidence is retrieved
 - further evidence is collected
 - a learned lesson is generated
 - same problem is investigated across models

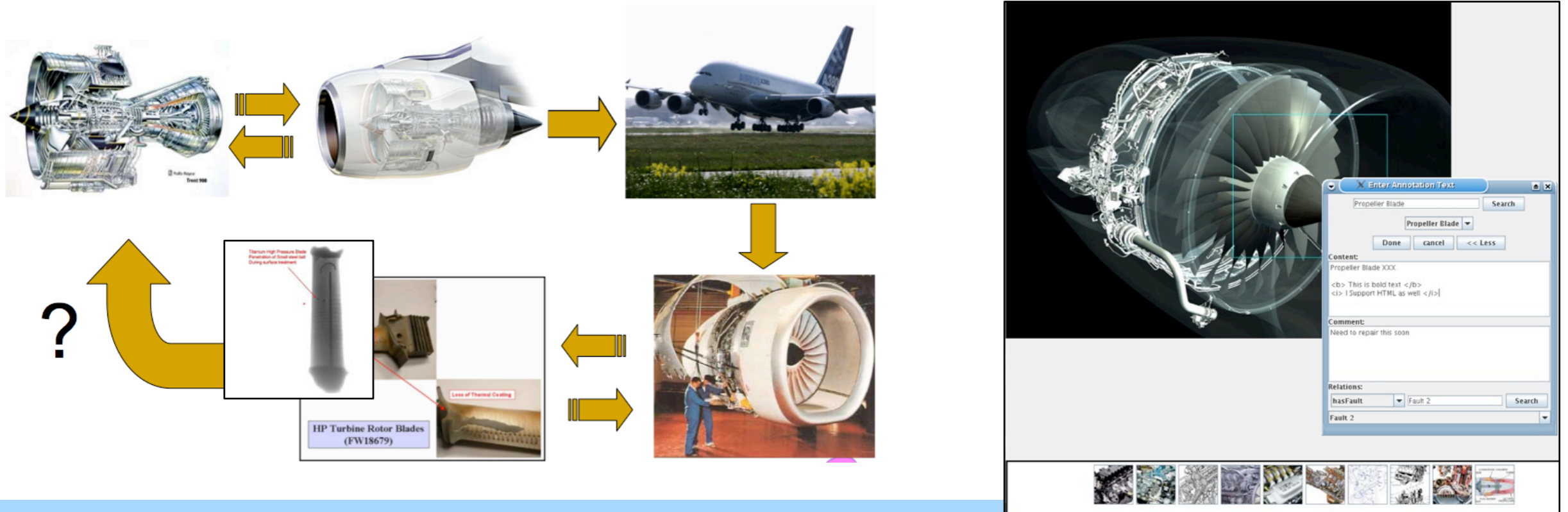


Different repositories represent different communities point of view!!

images © www.rolls-royce.com

Document Type
AROC <u>proforma</u>
AROC results
Development
EHM data
Emails
ONWING emails
Images
Lab findings
Monitoring Requirements
Presentations
Procedures
RCP
Risk Assessment
Solution Reports
Technical Reports
<u>TS&O Reports</u>

- Lifecycle “folder” will easily sum up to several Terabytes
- Folder will contain highly interrelated information stored in different media



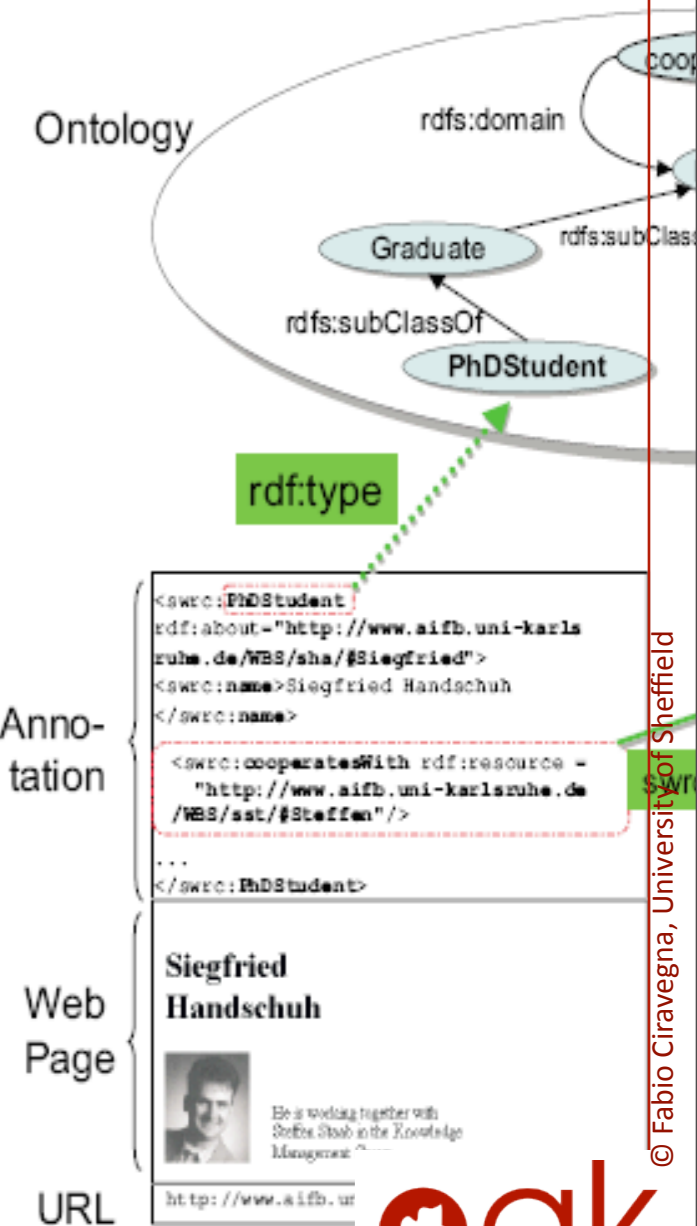
- Goal for Knowledge Management:
- Making information available independently from
 - Data format (structured/unstructured)
 - The archive
- Making it available for automatic processing
- Making it easily accessible and manageable despite its size



Ontologies and KM

Motivations for use:

- ▶ To represent an organisation's general view on the domain
 - How does the organisation work?
 - What is the organisation's official dictionary?
- ▶ As a middle layer to connect information from different information sources
 - The Web of data (as opposed to Web of documents)
- ▶ To represent communities' views of domains
 - e.g. marketing dept, customers, design and service departments have different views of the same products.
- ▶ Ontology mapping to navigate information sources
 - Mapping enables seamless communication among different worlds



Capturing Knowledge from the Semantic Web

Crawling the Semantic Web

Sindice

Giovanni Tummarello, Renaud Delbru, and Eyal Oren
[Sindice.com: Weaving the Open Linked Data](http://Sindice.com)

6th International Semantic Web Conference and the 2nd
Asian Semantic Web Conference, Busan, Korea Dates:
November 11 - 15 , 2007

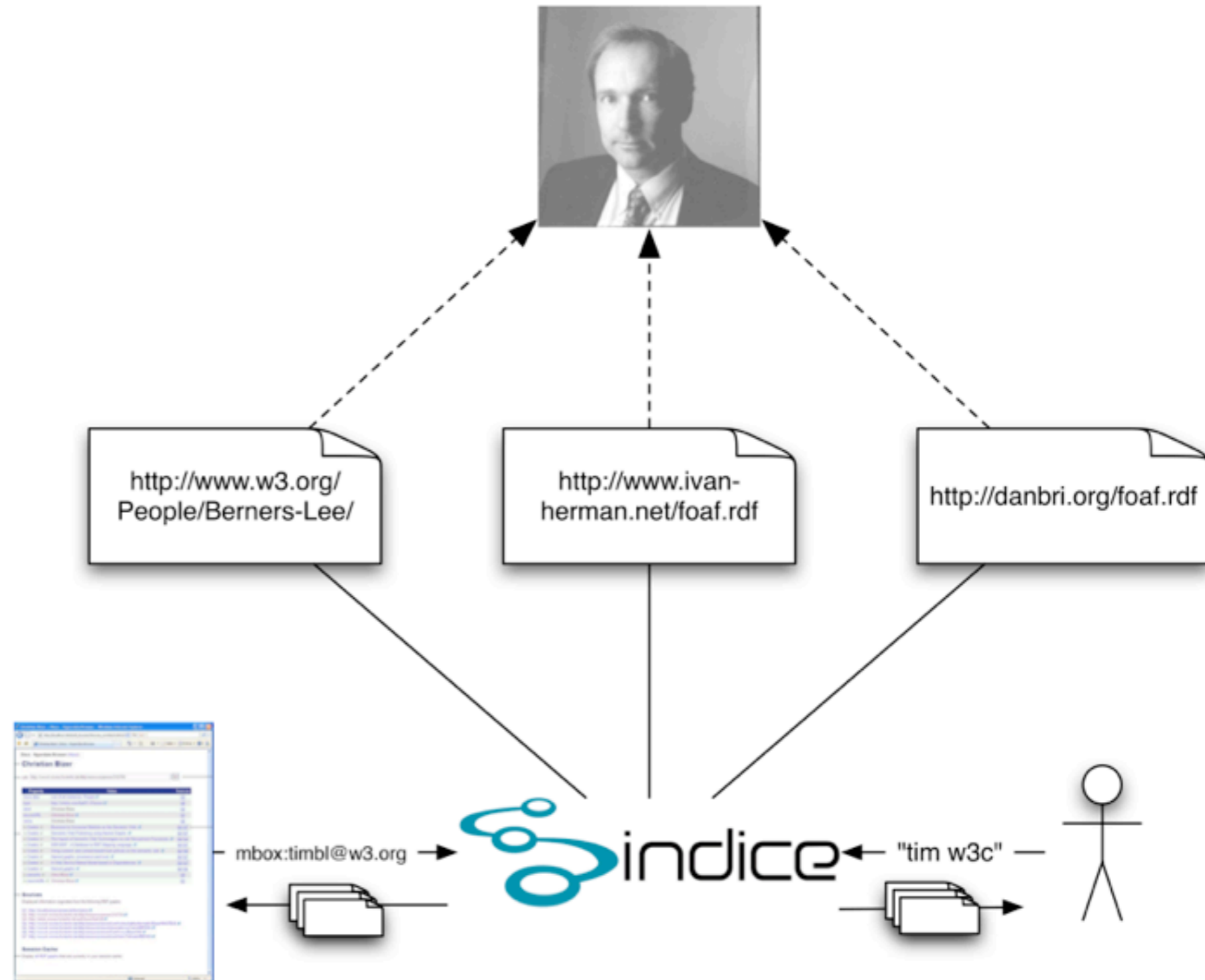
Indexing Triples across Sites

- The Semantic Web can be seen as a large knowledge-base
 - formed by sources that serve information as RDF files or through SPARQL endpoints.
- A fundamental feature of the Semantic Web is that the graphs are decentralised:
 - it has no single knowledge-base of statements
 - instead anyone can contribute statements by making them available in a public web space
 - These sources might have nothing in common,
 - but by using shared identifiers (URIs) and shared terms, their information can be merged to provide useful services to both humans and software clients.

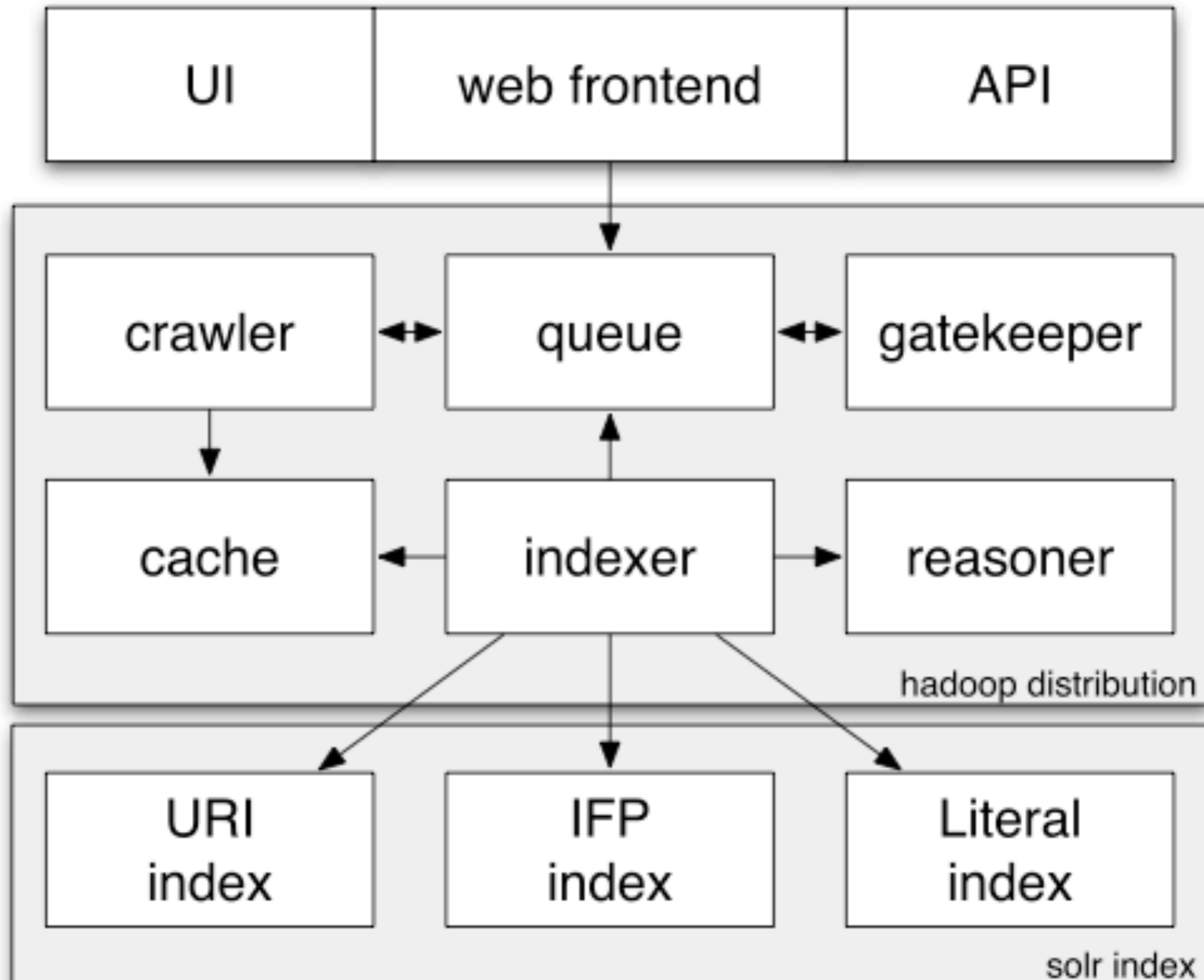
- Sindice crawls the Semantic Web and indexes the resources encountered in each source.
- A simple API offers to Semantic Web application developers
 - the ability to automatically locate relevant data sources
 - integrate the data from these sources into their applications
- Sindice collects RDF documents from the Semantic Web and indexes these on
 - resource URIs,
 - IFPs: inverse functional properties
 - Keywords
- It offers a user interface through which human users can find these documents, based on keywords, URIs, or IFPs

Example

- Example RDF documents mentioning Tim Berners-Lee, either by using his URI or by using IFPs that uniquely identify him.



Sindice: Architecture



Architecture (ctd)

- The crawler autonomously harvests RDF data from the Web and adds it to the indexing queue
 - It can be pinged (through the human interface or the API) to parse new documents, these are also added to the queue.
- The gatekeeper evaluates each entry in the queue and decides
 - whether, and with which priority, we want to index it,
 - based on whether we have seen the document before, its last modification date, its content digest, etc.
- The indexer extracts URIs, IFPs and keywords from each document and adds these to their respective index
- During lookup, the interface components
 - passes the queries to the relevant index,
 - gathers the results,
 - generate the required output such as HTML pages

Indexes

- The URI index contains an entry for each resource URI
 - It returns the list of document URLs where this resource occurs
- The IFP index uses the uniquely identifying pair (property, value) as index key.
 - It returns the list of document URLs where this pair occurs
- The literal index contains an entry for each token (extracted from the literals in the documents), again pointing to a list of document URLs
 - Just for text literals (no numbers!)

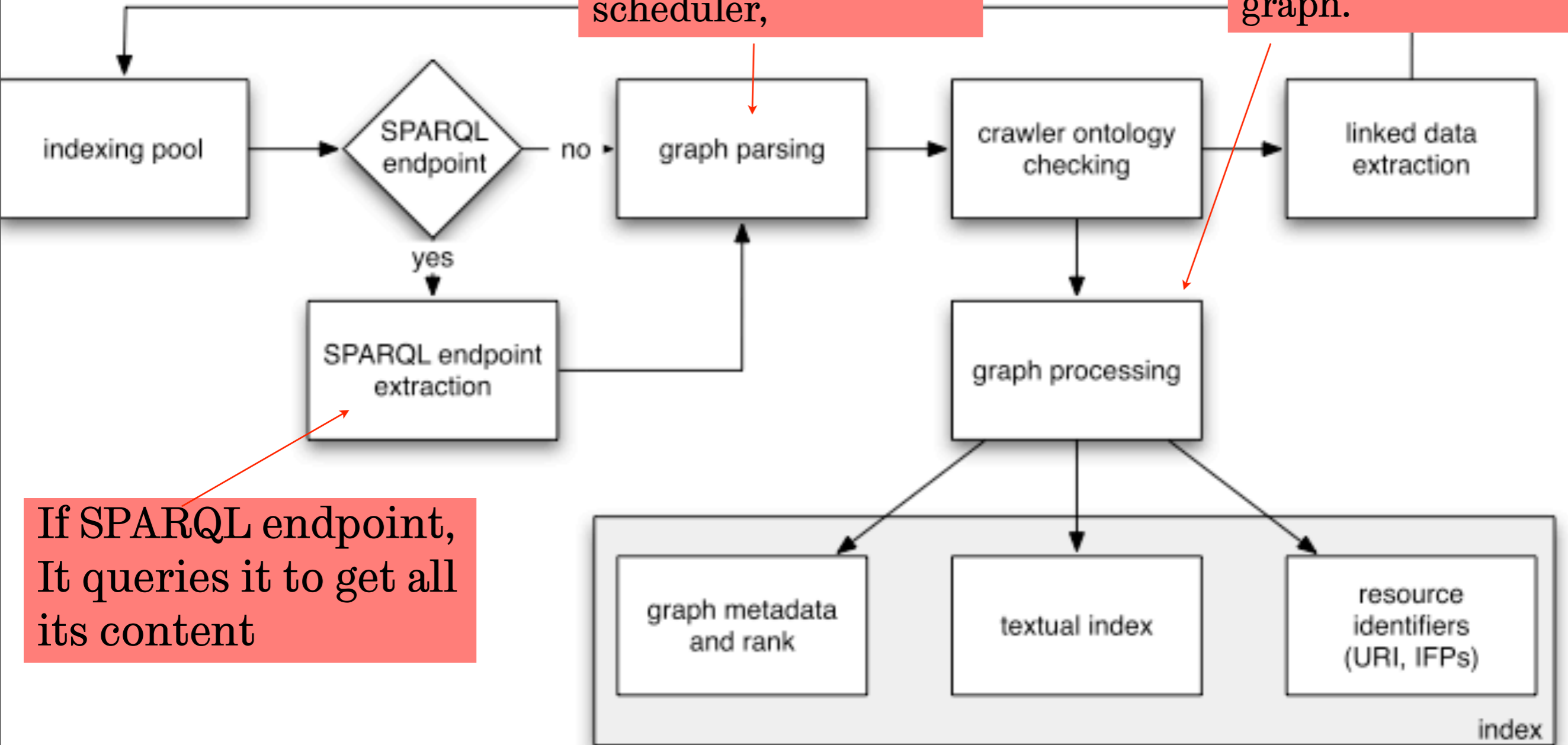
Index Size for a Billion Triples

- The index size of the crawl on the simple persistent hashtable of URI occurrences was around 2.5GB for 3.2 million URIs.
 - Given the scale-invariance of the URI/URL ratio we can extrapolate from this data and estimate to need around 785 bytes per resource;
 - Indexing a billion unique resources would thus require around **785GB**, an ordinary capacity for commodity harddisks

Indexing Pipeline

Extracts all URIs from the graph and injects them into the scheduler,

The graph processor then extracts and indexes the full-text and all resource identifiers in the graph.



If SPARQL endpoint, It queries it to get all its content

Querying pipeline

- Index retrieval
 - The query is looked up in the inverted index,
 - implemented either as an on-disk hashmap or in an information retrieval engine.
 - The list of results is cached for later reuse (refreshed daily)
- Ranking phase
 - Results are ranked according to various metrics
 - Hostname: we prefer sources whose hostname is the same as the resource's hostname, For example, we consider that more information on the resource *http://eyaloren.org/foaf.rdf#me* can be found at the source *http://eyaloren.org/foaf.rdf*
 - External rank: we prefer sources hosted on sites which rank high using traditional Web ranking algorithms.
 - Relevant sources: we prefer sources that share rare terms (URIs, IFPs, keywords) rather than common terms with the requested terms. This relevance metric is comparable to the TF/IDF relevance metric (Frakes and Baeza-Yates, 1992) in information retrieval.

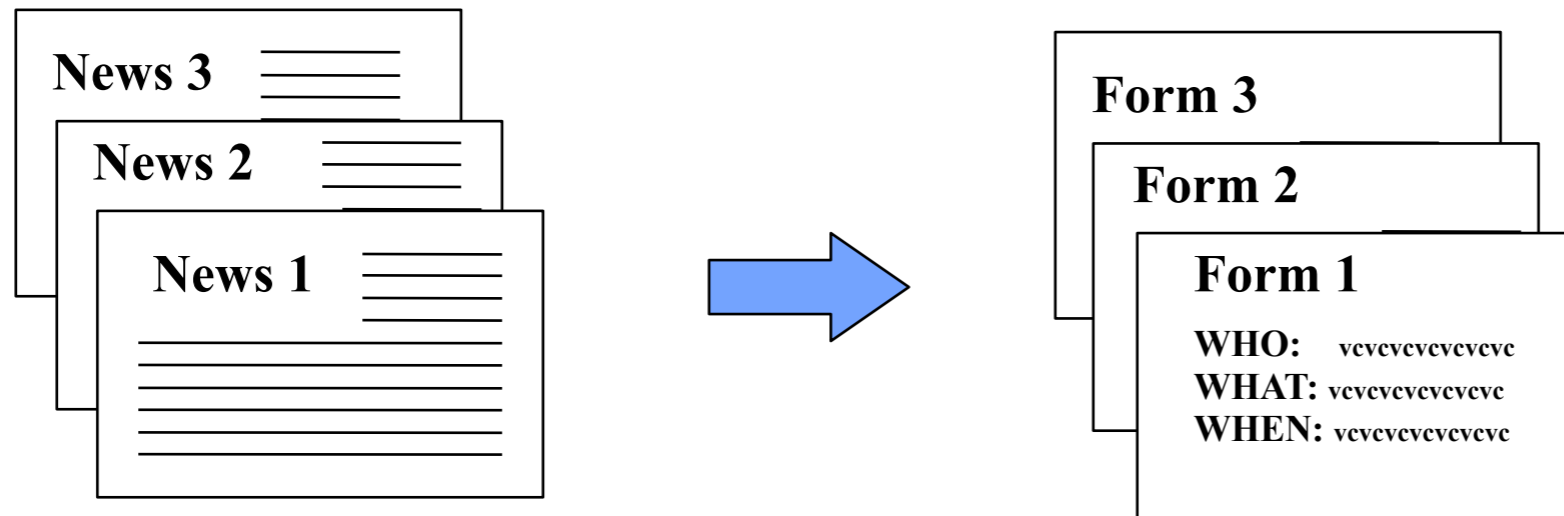
Querying (ctd)

- Result generation
 - Sindice can export results into different syntaxes, such as the HTML Web interface, RDF, XML, JSON, and plain text.

Extracting Information from Unstructured Documents



Information Extraction



- Automatically extracting pre-specified information from natural language texts
 - salient facts about pre-specified types of events, entities or relationships.
- Populating a structured information source from a semi-structured, unstructured, or free text, information source.



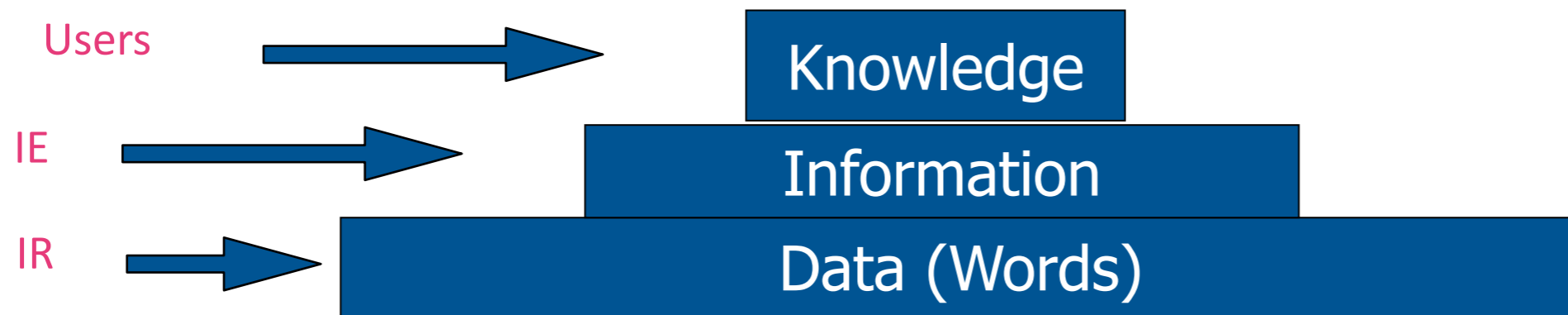
Why Texts and IE?

- Textual documents are pervasive (e.g. Web)
 - Contained knowledge cannot be queried
 - Q: How many cases of swine flu have been identified in the UK in the last three months that involve children under 5 years old
 - therefore knowledge cannot be
 - Used by automatic systems
 - Easily managed by humans
- IE can identify information in documents
 - e.g. to populate a database/ontology
 - e.g. to annotate documents
- Method: some forms of language analysis



IE Vs Information Retrieval

	IR	IE
Task	Data Indexing	Information Extraction
Returns upon User Query	Relevant Documents	Relevant Information
Query Generality	Full	Limited to target information



IE tasks

WASHINGTON, D.C. (October 5, 1999) -
nQuest Inc. today announced that P Paul Jacobs, former
Vice-President of E-Commerce at SR SRA International
has joined the company's executive management team
as president.

Named Entities

Relation Extraction

Event Extraction



IE tasks

WASHINGTON, D.C. (October 5, 1999) -

nQuest Inc. today announced that **Paul Jacobs.**

Vice-President of E-Commerce at **SRA International** has joined the company's executive management team as president.

Company: nQuest Inc.

Date: today

InPerson: Paul Jacobs

InRole: president

Company: SRA International

OutPerson: Paul Jacobs

OutRole: Vice-President of E-Commerce

Named Entities

Relation Extraction

Event Extraction





Tasks

- Entity Extraction/Ontology Population
 - Terminology Recognition
 - Classic NER
- Event Extraction
 - Field/Table Extraction
 - Relation Extraction
 - Cross-Media Extraction



The University
Of
Sheffield.

Document Type	Term	Tables	Entity	Relat	X-Media
AROC <u>proforma</u>	X	X	X	X	
AROC results					
Development	X		X	X	X
EHM data					
Emails	X	X	X		X
ONWING emails	X		X		
Images					X
Lab findings	X	X	X	X	X
Monitoring Requirements			X	X	X
Presentations	X		X		X
Procedures	X				X
RCP	X	X	X	X	
Risk Assessment	X	X	X		
Solution Reports	X				
Technical Reports	X	X	X	X	X
TS&O Reports	X	X	X		X

A Table by Jose' Iria



Requirements for IE

- Coping with large scale, in terms of size of corpus, of ontology and KB
- Efficiency: the system must be efficient
 - both in terms of speed and memory
- Ability to focus on information and knowledge that is infrequent
 - in contrast with redundancy based methods which extract the most frequent information



Req (ctd)

- Ability to use the background knowledge provided
 - By users (e.g. gazetteers)
 - By other media
- Portability:
 - Across corpora
 - Across domains
- Robustness:
 - being able to gracefully cope with unexpected situations
 - E.g. broken html



Approaches at a Glance

- Coping with technical terminology
 - Complex terms and URI identification
- Entity Recognition
 - Few labeled data, a lot of unlabelled
 - Problem: scalability
 - Problem: portability across corpora sub-domains
- Coping with tables
- Cross-Media
- Annotation:
 - Can we annotate?



Coping with Documents Formats

How do we

Represent documents across formats

Correlations across media?

Lei Xia, Jose Iria: An Approach to Modeling Heterogeneous Resources for Information Extraction, Proceedings of the Sixth International Language Resources and Evaluation (LREC'08), Marrakech, Morocco, May 2008

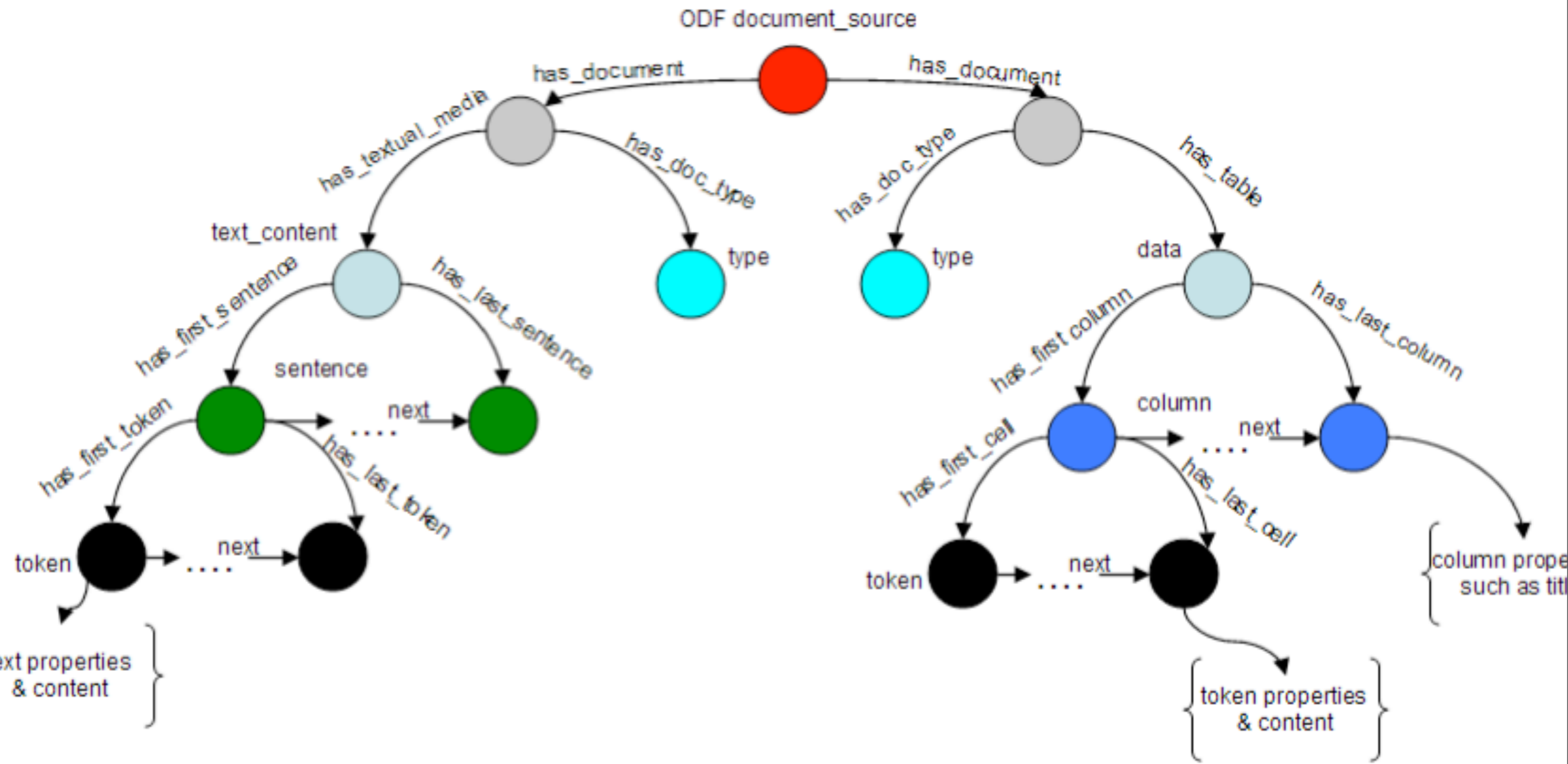


Coping with Multiple Formats

- Multiple document formats:
 - OpenDocument, Microsoft Office's, HTML, PDF, etc.
 - Carrying a mixture of
 - textual content,
 - metadata about the text (e.g. style information),
 - images, tables and other media objects
 - Carrying relational information
 - Valuable features to an IE algorithm



Modelling ODF





Preparing ODF Docs

- The algorithm is divided into two stages,
 - Document is split into coarse-grain-blocks according to document headings and sections
 - Using e.g. explicit formatting (`<head>`, `<P>`, `<style>`, `<bold>`, `<size>`, `<underline>`, etc).
 - Further analysis generates finer-grain-blocks of text and other nearby content types
 - Sliding-window of predefined size
 - Also for images -- see next slides

Entity Extraction

Using semi-supervised Methods for
Entity Extraction

Jose' Iria: Automating knowledge capture in the aerospace domain,
in Proceedings of the fifth international conference on Knowledge
capture, Redondo Beach, California, 2009



Goal

- Mimic domain experts in classifying pieces of text

Summary^{#60}

After the IP Compressor certification test vehicle Engine x failed to complete all of its planned testing before being removed from the test bed, a strain gauged IP compressor was fitted to engine x. The purpose of this test was to complete the unscheduled VSV test programme, assess the mechanical integrity of the engine and provide data for the engine certification planned for x. *2 test cond.*

This report covers data acquired from the IPC Rotors.

Conclusion And Recommendations

Excessive rotating and / or pseudo stall induced stresses were witnessed on rotor 1 for many of the ungauged VSV configurations tested. Running an engine in these conditions should be avoided to maintain integrity. *phen.* *test cond*

There were no significant Non Integral responses seen during any of the gauged unscheduled running or nominal schedule running. *phen* *component*

Stages 2- 8 all had acceptable levels of stress for all of the VSV configurations tested. The kulites fitted to stages 2 and 3 did detect stall patterns on some of the more extreme VSV configs but this excitation did not latch on to any of the modes of vibration. *test con*

The maximum speed cleared on all strain gauge selections was x% N2. *test cond has speed range.*

The allocated strain gauge selections prevented valuable detailed analysis being done. For future tests, involving Experimental Vibration earlier in the process should avoid these issues. *mode has freq (range)*

The blade to blade modal frequency variation combined with the close proximity of the modes will make it very difficult to incorporate any form of automated post processing system for tests of this type.

Generally modal frequencies were higher than predictions and further work by the Capability Acquisition group is needed to understand this.

Additional keywords ¹⁹⁹	Retention category A
------------------------------------	--------------------------------





Extraction by Boundary Classification

- Boundary: virtual separator between any two tokens
- [...] in <location>London</location> this week [...]
- Binary Classification ↑ ↑ ↑
 - set of binary classifiers classify boundaries as start/end of relevant text fragments
 - positive instances for start/end classifier for a given type become negative instances for all other classifiers



Encoding Features for Learning

- Binary features
- Concatenation of feature type, feature value, and position with respect to boundary

[...] in <location>London</location> this week [...]



"-1_str_in"

"1_str_this"

"1_pos_DT"

"1_orth_alllow"

"2_str_week"

...



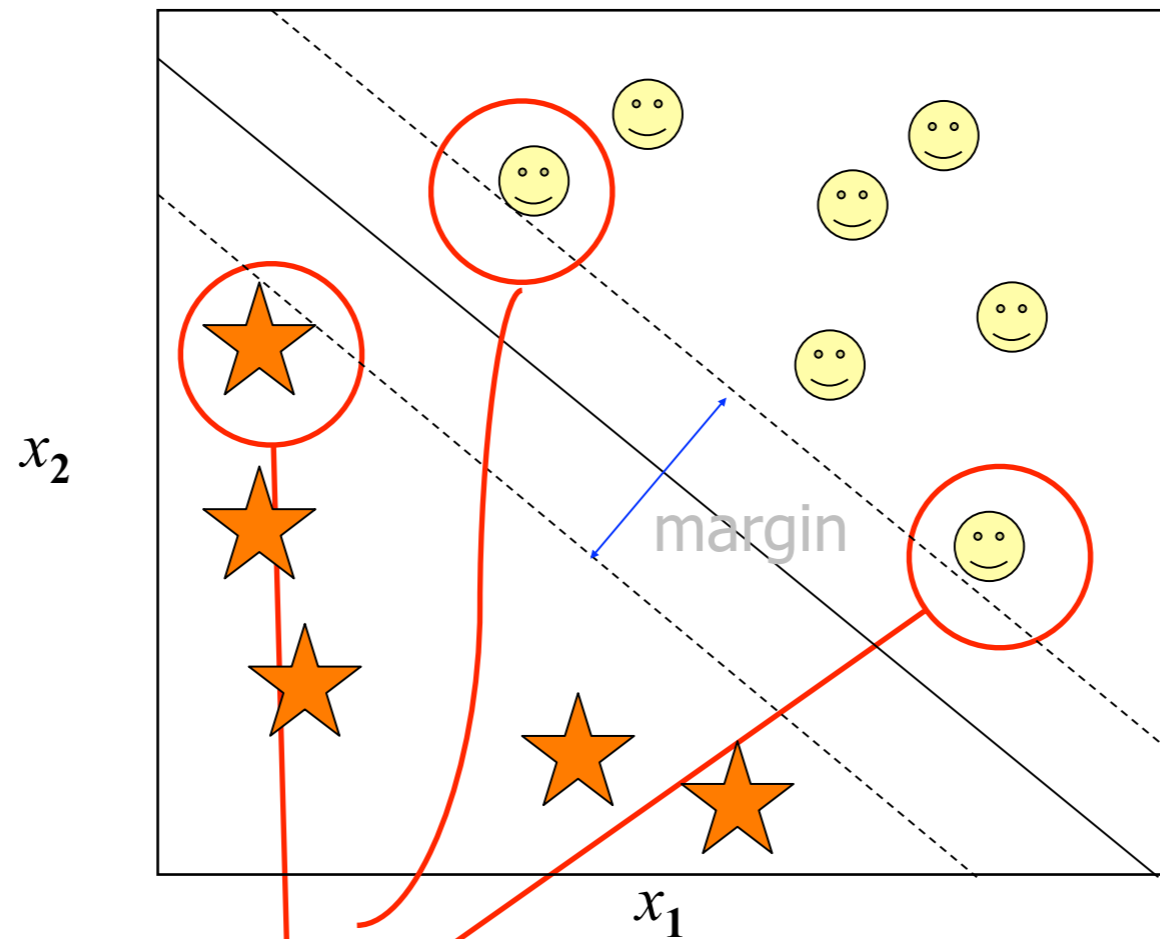
Reducing Annotation Needs

- Exhaustive annotation and training is very complex and difficult
- Solution:
 - Use of semi-supervised methods for learning



Baseline: SVM

- Margin-based classifier, identifies support vectors in the data
- Optimisation procedure has quadratic complexity
- Requires complete annotation



"Support vectors"



Label Propagation

- Graph-based semi-supervised methods define a graph where
 - Nodes are both the labelled and unlabelled examples in the dataset
 - Edges (may be weighted) reflect the similarity of examples
- In label propagation,
 - Known labels are used to propagate information through the graph in order to label all nodes.
 - Goal is to learn a labelling function that:
 - Is close to the given labels on the labelled nodes,
 - Is smooth on the whole graph



C-SVC algorithm

- Starting with nodes $1, 2, \dots, l$ labelled with their known label (1 or -1) and nodes $l+1, \dots, n$ labelled with 0, each node starts to propagate its label to its neighbours, and the process is repeated until convergence

Compute kernel matrix K

Compute the diagonal degree matrix D by $D_{ii} \leftarrow \sum_j K_{ij}$

Initialize $Y^0 \leftarrow (y_0, \dots, y_l, 0, \dots, 0)$

Iterate

1. $Y^{(t+1)} \leftarrow D^{-1}KY^t$

2. $Y^{(t+1)} \leftarrow Y_l$

until convergence to $Y^{(\infty)}$

Label point x_i by the sign of $y_i^{(\infty)}$

- Complexity $O(kn^2)$
 - for a sparse graph where each data point has k neighbours

It can be seen as optimising a cost function that can be expressed as an average over the training examples

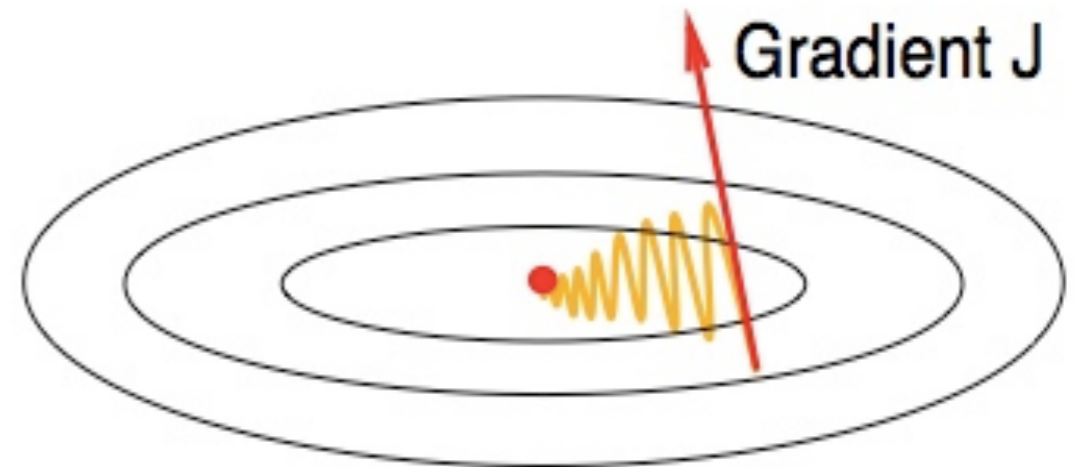


Large-scale: Stochastic Gradient Descent

- Try and exploit the availability of lots of data
- Iterative procedure, linear complexity to convergence

Iterate

$$\bullet w_{t+1} \leftarrow w_t - \eta \frac{\partial E_n(f_{w_t})}{\partial w}$$





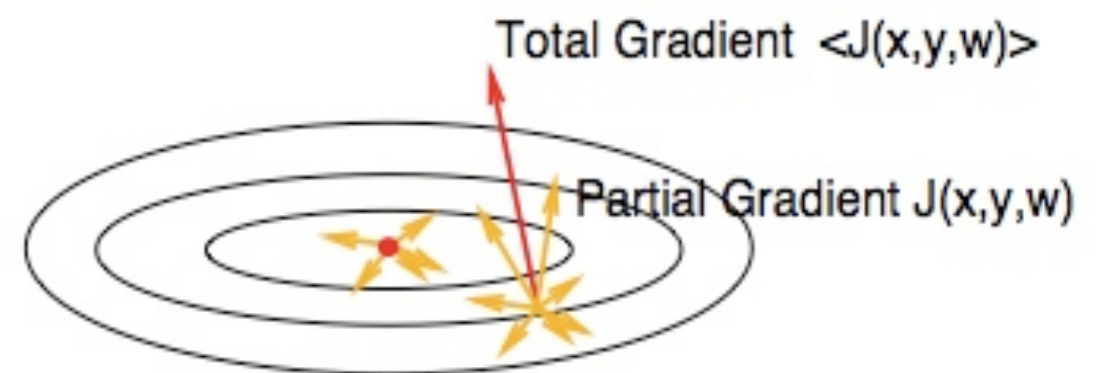
Stochastic Gradient Descent

- In stochastic (or "on-line") gradient descent, the true gradient is approximated by the gradient of the cost function only evaluated on a single training example
- The parameters are then adjusted by an amount proportional to this approximate gradient

Iterate

- Draw random example (x_t, y_t) .

- $w_{t+1} \leftarrow w_t - \frac{\eta}{t} \frac{\partial \ell(f_{w_t}(x_t), y_t)}{\partial w}$





Experiments

Runes, plus tokenizer, pos etc. provided the preprocessing (annotation) of the corpora

- 70 documents
- 661,117 words
- 14 tags
- over 2500 annotations

Engine	
Engine Module	HP Compressor
Engine Serial Number	HP Turbine
Module Serial Number	IP Turbine
Customer Number	LP Turbine
Document Title	Tube
Document Date	Groove
Observed Damage	Ring



Validation Methodology


- Test three algorithms:
 - SVM, Graph Label Propagation, Stochastic Gradient Descent
- General test conditions:
 - Using a linear kernel
 - 5 fold cross-validation for svm and svm-sgd
 - Same underlying framework, just change algorithm
- Datasets generated for learning:
 - One per concept start/end
- Evaluation measure: F-measure
 - Exact matching on the predicted and gold standard boundaries



Results (cont.)

- F-measures obtained are virtually the same
 - Graph label propagation obtains slightly better results
 - SVM and SGD differ mostly in recall
- However, training times vary dramatically
 - SVM-SGD ideal for application scenarios where on-the-fly analysis is required

Algorithm	Average Precision	Average Recall	F-measure	Training Time
SVM	0.75	0.65	0.70	4m 58s
Label Propagation	0.77	0.66	0.72	2h 11m
Stochastic Gradient Descent	0.75	0.62	0.68	28s

$O(kn^2)$
 $O(n)$

 Organisations, Information and Knowledge



The
University
Of
Sheffield.

Terminology Recognition in the Aerospace Domain

Jonathan Butters and Fabio Ciravegna:
Authoring Technical Documents for Effective Retrieval
17th International Conference on Knowledge Engineering and Knowledge Management
Lisbon October 2010



Rolls-Royce

Runner up at the
Director of Research's Creativity Award 2009





Terminology Recognition

“Low Pressure Turbine Stage 2 Rotor Blade”

“LP2 Blade”

“FK42164”

“LPT 2 Blade”

“72-41-12”

“T800 LP Turbine Blade Stage 2”

“Turbine Blade”

“72-41-12-400”

“Blade, Turb l2”

“Blade, LPT”

“TurbinneBladee”

“FK12548”

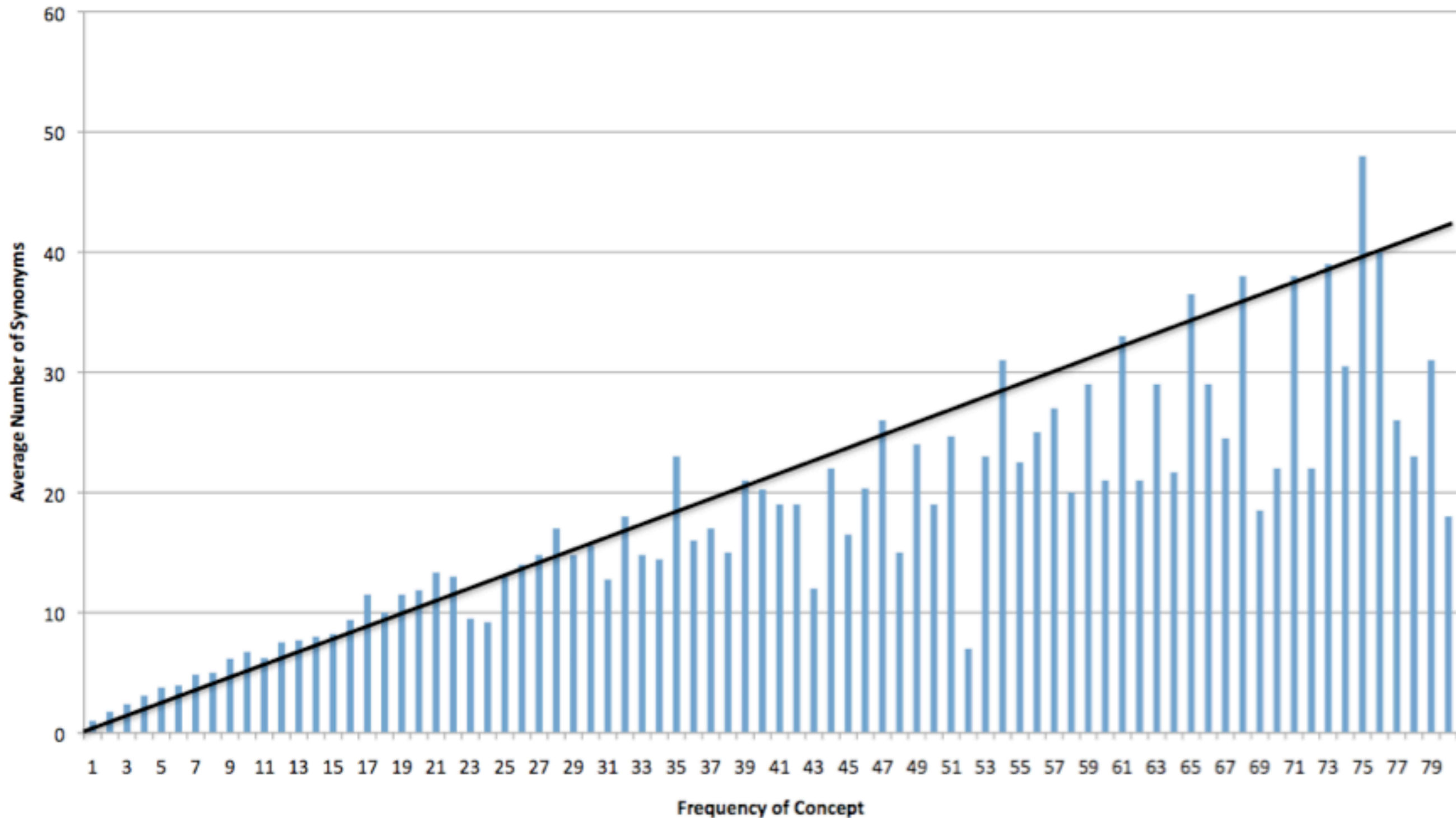


- Task of reducing a term to a URI



A Pervasive Issue

Synonym Increase with Concept Frequency





Term Variation

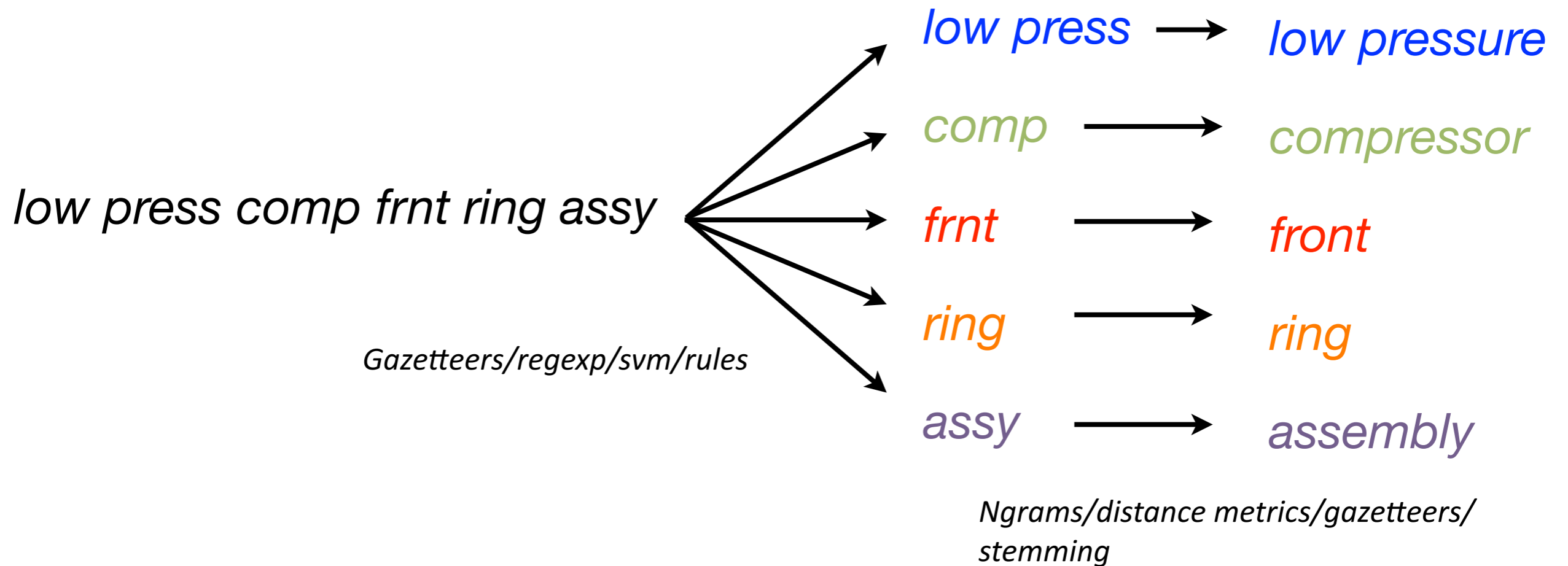
- Aerospace Terms Typically Comprise Long Strings of Nouns – [Lehrberger & Kittredge, 1982]

- “*low press comp frnt ring assy.*” This is just a string of characters!
- [LP], [COMPRESSOR], [FRONT], [RING], [ASSY] These are the ‘sub-concepts’
- 5 3 7 3 5 These are the numbers of terms that can represent each sub-concept

- That gives 1575 combinations (not counting word order!)
 - E.G: “front ring assembly, m34”
- A Gazetteer list of 1575 terms for one master term is NOT practical!
- So, our Approach..



Approach: stage 1&2

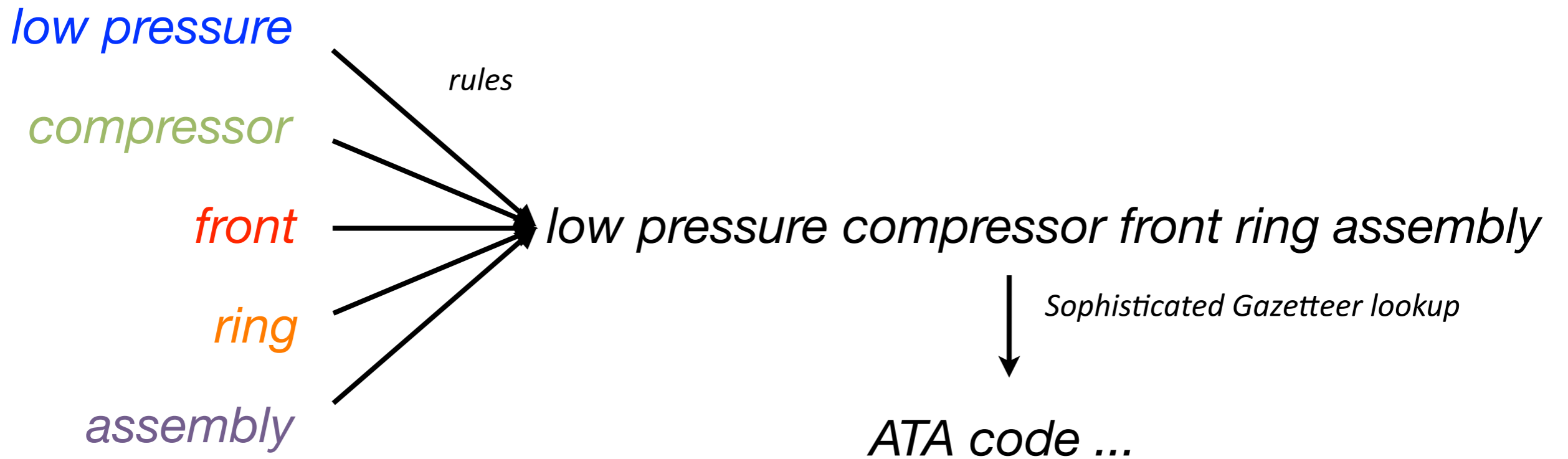


- All possible solutions carried forward in parallel
- A new recognition at each stage makes the algorithm cycle

Patent Pending



Approach: stage 3&4



All possible solutions carried forward in parallel

Patent Pending

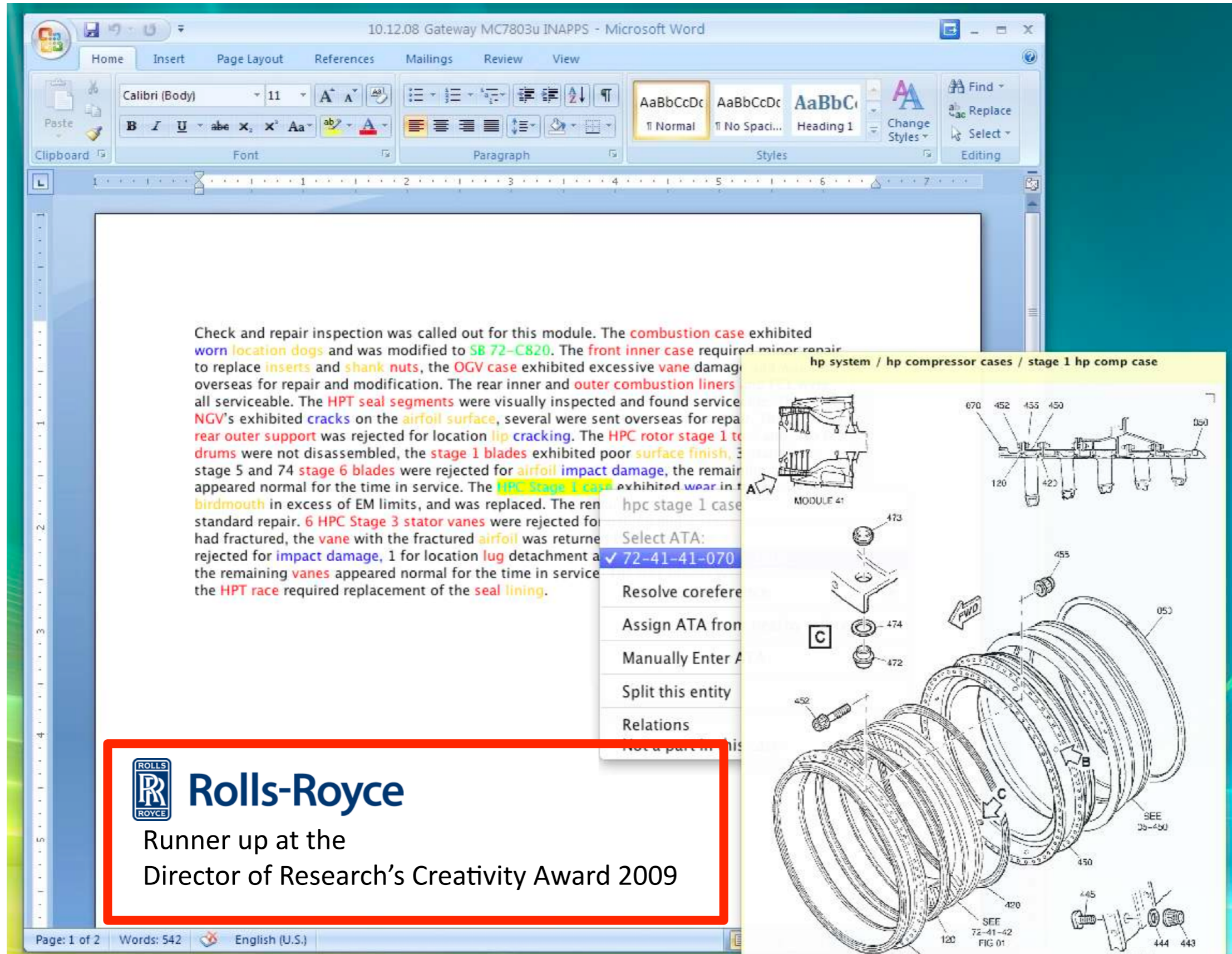
Comparative Evaluation

- 88,213 reports marked with ATA100 number
- Data set 2 comprised 4,394 complete documents randomly selected from across 6 corpora

	Corpus A			Corpus B		
	Pre	Rec	F1	Pre	Rec	F1
TF-IDF	12.00%	8.54%	9.98%	14.62%	7.32%	9.76%
Termex	41.69%	18.02%	25.16%	49.82%	21.30%	29.84%
C-Value	52.87%	34.86%	42.02%	62.40%	41.85%	50.10%
TR	69.03%	97.12%	80.70%	92.77%	98.30%	95.45%

	Corpus C			Corpus D		
	Pre	Rec	F1	Pre	Rec	F1
TF-IDF	16.33%	5.83%	8.59%	13.59%	6.21%	8.52%
Termex	41.34%	25.34	33.34%	51.43%	22.73%	31.53%
C-Value	60.29%	39.93%	48.04%	64.76%	43.86%	52.30%
TR	94.49%	98.10%	96.26%	85.14%	94.03%	89.36%

The RR Application



10.12.08 Gateway MC7803u INAPPS - Microsoft Word

Home Insert Page Layout References Mailings Review View

Clipboard Font Paragraph Styles Editing

Calibri (Body) 11

Find Replace Select

Check and repair inspection was called out for this module. The **combustion case** exhibited **worn location dogs** and was modified to **SB 72-C820**. The **front inner case** required minor repair to replace **inserts** and **shank nuts**, the **OGV case** exhibited excessive **vane damage** overseas for repair and modification. The rear inner and **outer combustion liners** all serviceable. The **HPT seal segments** were visually inspected and found serviceable. **NGV's** exhibited **cracks** on the **airfoil surface**, several were sent overseas for repair. **rear outer support** was rejected for location **lip cracking**. The **HPC rotor stage 1 teeth** and **drums** were not disassembled, the **stage 1 blades** exhibited poor **surface finish**, 3 stage 5 and 74 **stage 6 blades** were rejected for **airfoil impact damage**, the remainder appeared normal for the time in service. The **HPC stage 1 case** exhibited **wear** in the **birdmouth** in excess of EM limits, and was replaced. The remaining **HPC stage 1 case** required standard repair. **6 HPC Stage 3 stator vanes** were rejected for **impact damage**, 1 for location **lug detachment** and the remaining **vanes** appeared normal for the time in service. The **HPT race** required replacement of the **seal lining**.

hp system / hp compressor cases / stage 1 hp comp case

MODULE 41

670 452 455 450 120 420 120

473 474 472

455 053 450 420 445 444 443

SEE 72-41-42 FIG 01

SEE 72-41-42

SELECT ATA:

- 72-41-41-070
- Resolve corefer
- Assign ATA from
- Manually Enter A
- Split this entity
- Relations

ROLLS ROYCE

Rolls-Royce

Runner up at the Director of Research's Creativity Award 2009

Page: 1 of 2 Words: 542 English (U.S.)



The
University
Of
Sheffield.

Table Extraction



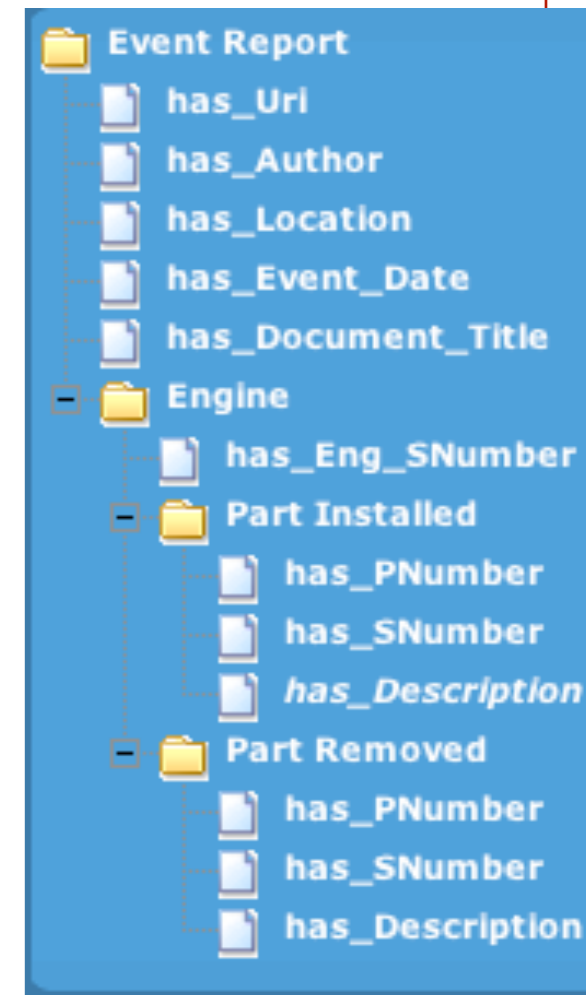
Tables

- Are pervasive
- Carry implicit relations
 - That need making explicit
- Can be created at any time by anyone
 - Semantics must be discovered



Annotating Documents

- Automatic extraction of information from event report
 - 18,000 documents analysed
 - Mainly Forms implemented in Word
- Metadata generated according to an ontology developed by Aberdeen U
- Automatic extraction of metadata and indexing of documents





An Experiment on Event Report for Jet Engines

Event Report No. COMPANY2698 /

1263 Prepared By: Richard Williamson Originated/Revised on: 13 March 2004

Event Report Data

engine type

company

Aircraft

LN144

Event Date: 12-Mar-04 Engine S/N: 51179 Flight Regime: Unknown Hazard Type:
 Aircraft Regn: 9V-SRA Installed Posn: Right Location: SIN No Hazard
 Airframe Hours: 20779 Engine TSN/CSN: 14242 / 4014 Event Type: Operational
 Airframe Cycles: 5609 Engine TSF/CSF: 6249 / 1814 Event Category: Basic

Reactions to Event:

Primary: None ABTO Speed
 Secondary: None (Knots): N/A Operational Effect: No Effect SERAPH Symptom Codes:
 Third: None Delay Time (mins): N/A NREP NREP NREP
 Fourth: None Fuel Dumped?: No

EICAS Messages (If Any):

Maintenance Messages (If Any):

Parts/Components Removed or Installed (If Any):

On/Off	Part Number / Serial Number	Part Description	Hours / Cycles	Qty	Destiny / Disposition	Pull Category / Pull Code
Installed	9-217-62 Y487	FUEL FLOW TRANSMITTER		1	SE - Serviceable	
Removed	921762 Y403	FUEL FLOW TRANSMITTER		1	R4 - Return to Manufacturer US - Unserviceable	U - Unplanned I - Inspection/Investigation

Description of Event:

a short sentence





Examples of tables in Event Reports

module/accessory details			
<u>item</u>	<u>part number</u>	<u>s/n removed</u>	<u>s/n installed</u>
	p39-401revf	04-0721257 <u>tsn/csn: 268/106</u>	04-1012229 tsn/csn:0/0

Part numbers
04-0721257 <u>tsn/csn: 268/106</u> off
04-1012229 tsn/csn:0/0 on

<u>s/n removed</u>	04-0721257 <u>tsn/csn: 268/106</u>
<u>s/n installed</u>	04-1012229 tsn/csn:0/0

Parts/Components Removed or Installed (If Any):					
On/Off	Part Number / Serial Number	Part Description	Hours / Cycles	Qty	Destiny Disposit
Installed	FK30840		11129 TSN 1954	1	
	RGG12340				
Installed	FK21221 EC092		11652 TSN 2119	1	
Installed	FK30840		11129 TSN 1954	1	
	RGG12501				
Installed	FK30840		11129 TSN 1954	1	
	RGG12208				
Installed	FK30840		11129 TSN 1954	1	
	RGG12391				





Applying information extraction

- AktiveMedia to annotate texts
- SVM to train and extract (supervised)
- IE captures all the information in tables
 - 99% of the information captured (recall=99)
 - 98% of proposed information is correct (precision=98)

As said using semi-supervised would not have made much difference

	POS	ACT	CORR	WRONG	MISSED	PREC	REC	F1
airport	120	120	120	0	0	100	100	100
has_airframe_cycles	104	104	104	0	0	100	100	100
has_airframe_hours	104	104	104	0	0	100	100	100
has_author	120	120	120	0	0	100	100	100
has_engine_serial_number	120	120	120	0	0	100	100	100
has_engine_type	120	120	120	0	0	100	100	100
has_event_date	120	120	120	0	0	100	100	100
has_event_report_no	356	358	356	2	0	99	100	100
has_part_description_installed	120	113	111	2	9	98	93	95
has_part_description_removed	120	133	120	13	0	90	100	95
has_part_number_installed	120	113	111	2	9	98	93	95
has_part_number_removed	120	133	119	14	1	89	99	94
TOTAL	1644	1658	1625	33	19	98	99	98

Porting across Domains

Jing Jiang: Domain Adaptation in Natural Language Processing
PhD Thesis, University of Illinois at Urbana-Champaign, 2008



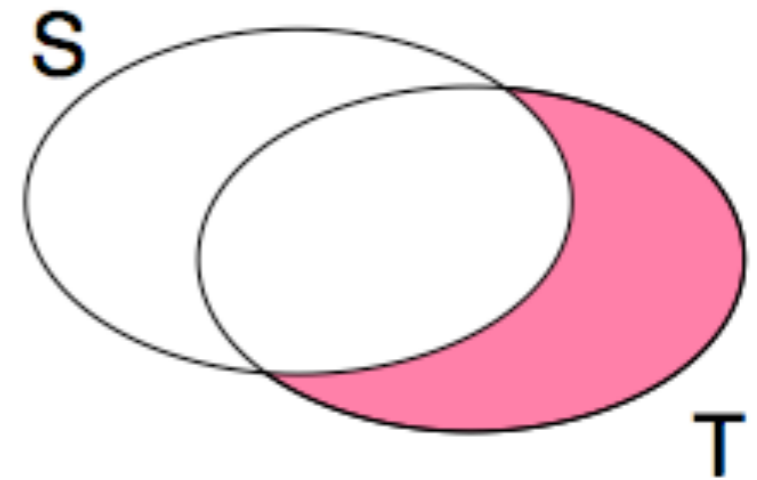
Large no. of Corpora

- Given the number of corpora under consideration
 - it will be very difficult (if not impossible) to process each from scratch.
- Requirement:
 - ability to process one corpora and then
 - port the learned models (with minor work) to other corpora.



Transfer Learning

- The domain difference comes from some special characteristics
 - in the target domain or
 - in the source domain or
 - in both
- In transfer learning the goal is to use training data from a related domain, along with training data from the target domain, to train the target classifier





Domain Adaptation

- In domain adaptation
 - labeled data from source domain is used to train a model that maximises accuracy in a target domain
 - for which we only have unlabelled data available
- Previous work focuses on document classification on academic datasets
 - Our experiments concern the entity recognition task over real-world data
 - Corpora cover the same domain
 - but are distinct enough to be considered covering different sub-domains



Approach

- Extending existing learning models according to new evidence obtained from previously unseen corpora from the same domain
- A bootstrapping approach that iteratively refines the learned models on the new corpus,
 - using as starting point a context-independent model derived from the initial corpus
 - and exploiting user interaction and terminology recognition at each iteration step





- Context-independent patterns yield low-precision and variable recall
 - Coupled with additional techniques designed to raise both precision and recall can address requirements of target domain
 - Terminology recognition can help raise recall

Information Integration

70



Information Integration

- Facts from different sources need to be integrated
 - To connect information/knowledge across docs
 - Assign unique URI
 - To solve discrepancies and ambiguities
- Steps
 - Unique instance identification (for entities)
 - Record linkage (for events)
- Information Integration strategies
 - Generic
 - Distance metrics (Chapman 2004)
 - Using Web bias
 - Statistical matching
 - Application specific
 - Rules



SimMetrics

- Library of distance metrics released as open source
- <http://sourceforge.net/projects/simmetrics/>
- >20,000 downloads since end of 2004
- Most downloaded distance metrics library on the Web
 - for strings and records
- Hundreds of applications
- Developed by Sam Chapman, University of Sheffi

Anastasiosyal.com

So what's the buzzword today!?

<< SQL Server script to auto-create indexes on all Foreign Key Columns

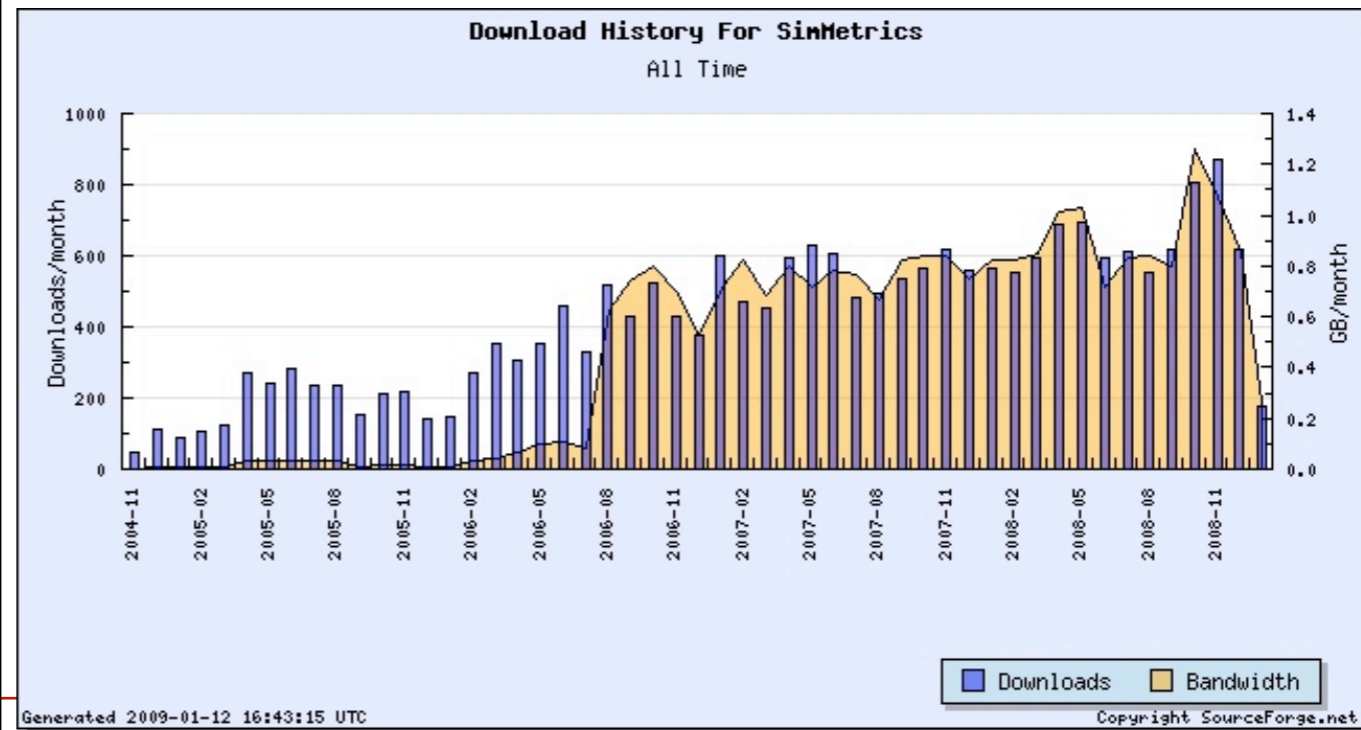
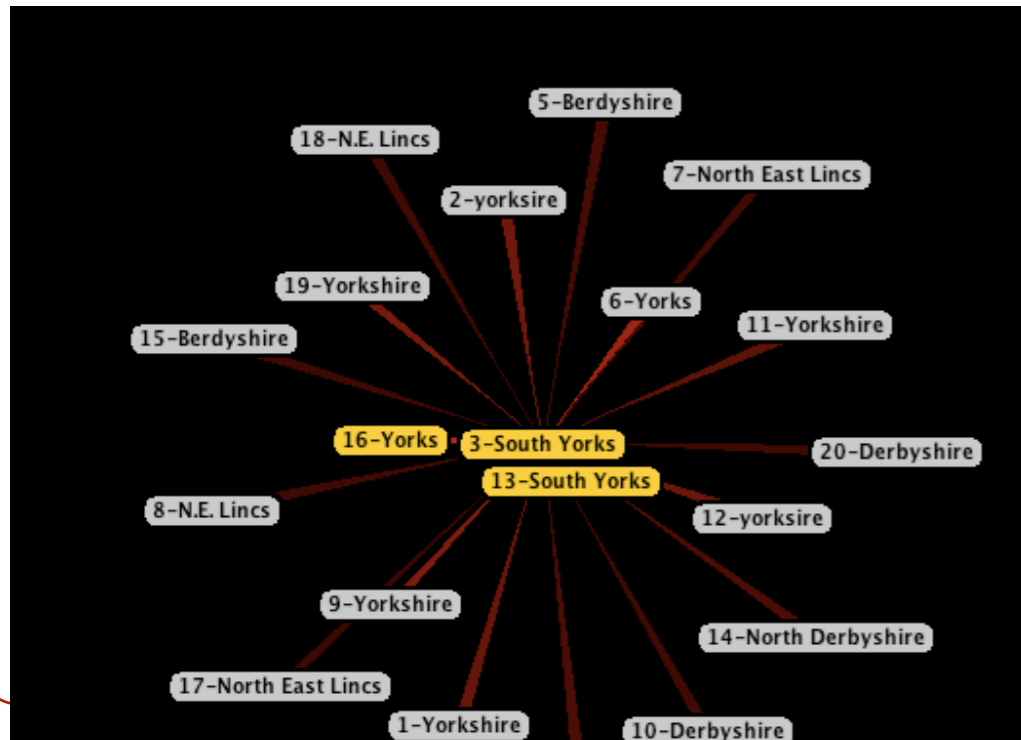
Beyond SoundEx - Functions for Fuzzy Searching in MS SQL Server

Sunday, January 11, 2009 12:08 AM, Filed Under [SQL C#](#)

In this post:

1. [SoundEx in Sql Server](#)
2. [SimMetrics](#)
3. [Adding string Metric functions in MS Sql Server](#)
4. [Evaluating metric accuracy and comparing Metrics](#)
5. [Conclusion + code](#)

Quite often we come across a requirement where we may need to perform some sort of fuzzy string grouping or data customer records of a database by identifying records that are similar but not necessarily exactly the same (due to spelling not successfully group such data. We will need to employ what is commonly referred to as a distance algorithm or a string are.





The University Of Sheffield.

Armadillo: Historical Data Mining

<http://www.hrionline.ac.uk/armadillo/>

The Marine Society Registers

The Westminster Historical Database

Eighteenth Century Fire Insurance Policies

Prerogative Court of Canterbury Wills

The Proceedings of the Old Bailey

AHDS Deposits

St. Martin's Settlement Exams Index
WESTCAT

Collage image database
Guildhall Library

Harben's Dictionary of London

John Strype's "Survey..."

Metropolitan London in the 1690s
IHR

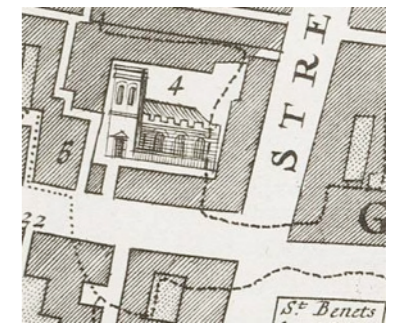
Selected Criminal Records
PRO

<http://www.motco.com>

House of Lords Journals
BOPCRIS



THE PROCEEDINGS ON THE KING'S Commission of the Peace AND Oyer and Terminer, and Goal-Delivery of Newgate, held for London and COUNTY of Middlesex, at Justice Hall in the City of London, On Wednesday, Thursday, and Friday, being the 16th, 17th, and 18th, January, in the Ninth Year of His MAJESTY'S REIGN.



Arts & Humanities Research Council

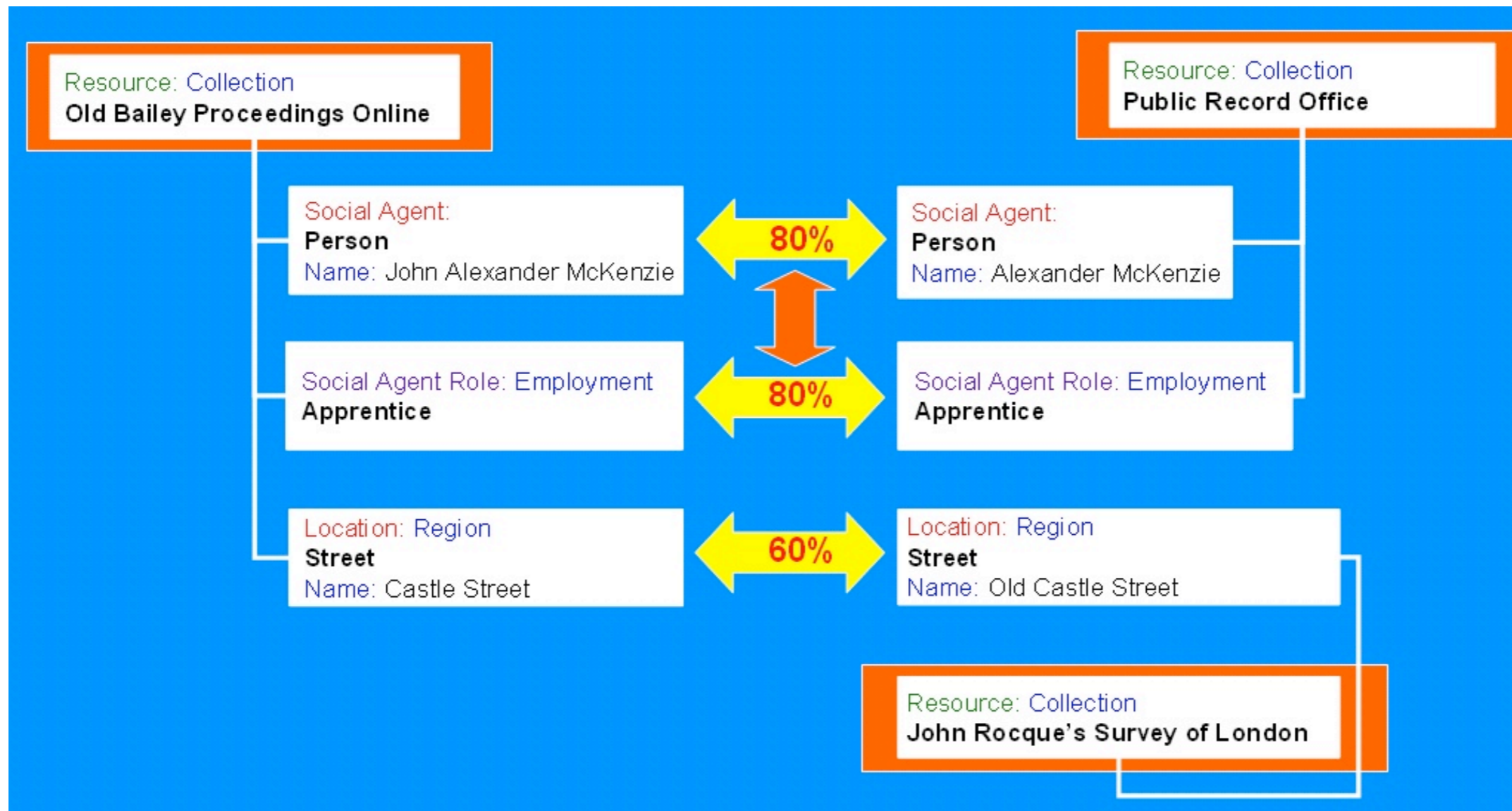


Organisations, Information and Knowledge



Information Integration

Armadillo: Historical Data Mining





Department of W&P Appl

• Goal: identifying the most likely record matching a NL description

– I want to become (formation)

- A plumber (job)
- In South Yorkshire (geography)

– Given misspellings and lack of facility with computers

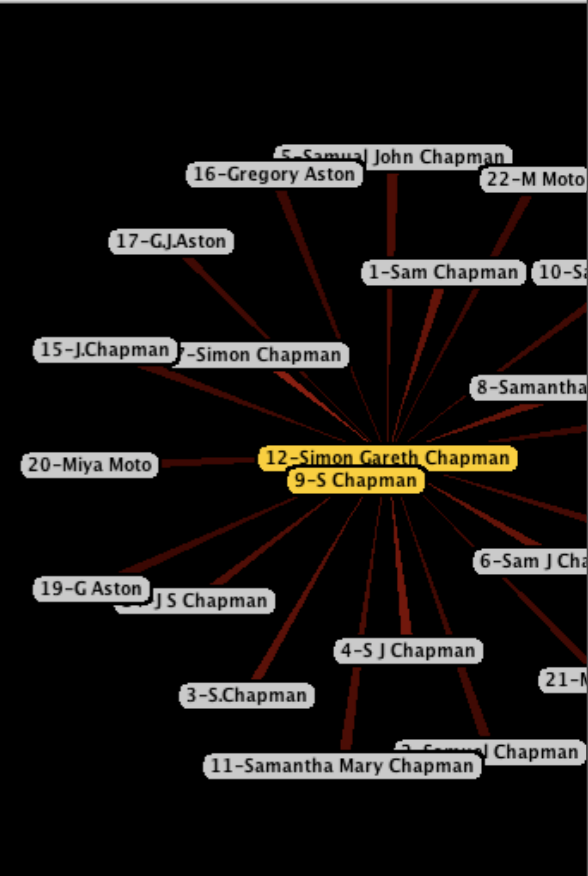
- “become plumba in Sothyorks”

SimilarityTester

A	B	C	D
1	Sam Chapman	26th July 1974	Computer Science
2	Samuel Chapman	26 / 07 / 74	Comp Sci
3	S.Chapman	26 July 74	Researcher
4	S J Chapman	26 July	Computer Sci
5	Samual John Chapman	26 / 07 / 1974	Research Student
6	Sam J Chapman	26 July 1974	Computer Science Student
7	Simon Chapman	31 July 1969	Monster Trucks
8	Samantha Chapman	19th June 1974	Personal Assistant
9	S Chapman	31st July 1969	Monster Truck Driver
10	Sam. M. Chapman	19 / 06 / 74	P A
11	Samantha Mary Chapman	19 - 06 - 1974	Personel Assistant
12	Simon Gareth Chapman	31st july	Stunt Driver
13	John Chapman	11 - 01 - 80	Performance Artist
14	J S Chapman	Eleventh of January 1980	Modern Artist
15	J.Chapman	11 / 01 / 80	Performer of Art
16	Gregory Aston	03 / 02 / 59	unknown ?
17	G.J.Aston	3rd Feb 1959	Headmaster of KEVIS

Cutoff Ratio 0.6

12 vs 13 | Average = 0.012772216
 12 vs 16 | Average = 0.006714133
 12 vs 17 | Average = 0.011662246
 12 vs 18 | Average = 0.0075900513
 12 vs 19 | Average = 0.007684646
 12 vs 20 | Average = 0.0070008663
 12 vs 21 | Average = 0.011981451
 12 vs 22 | Average = 0.008990145
 Total Time: 91ms
 Total Tests: 1320 Metric Tests
 Total Time: 0.068939395ms per test





- DU360: adaptive information integration for products/vendors from different suppliers and EU standards
 - Several large customers in detail commerce (supermarkets)



Sam Chapman and Fabio Ciravegna: K-Spend: Semantic Web Technologies for Spend Analysis. Sam Chapman and Fabio Ciravegna
Industry Track, ISWC 2010
To be presented on Thursday, November 11, 10:30Am - 12:30 Pm
Industry Track: Session 1
Location: Room 3B

Retrieval

Semantic Search

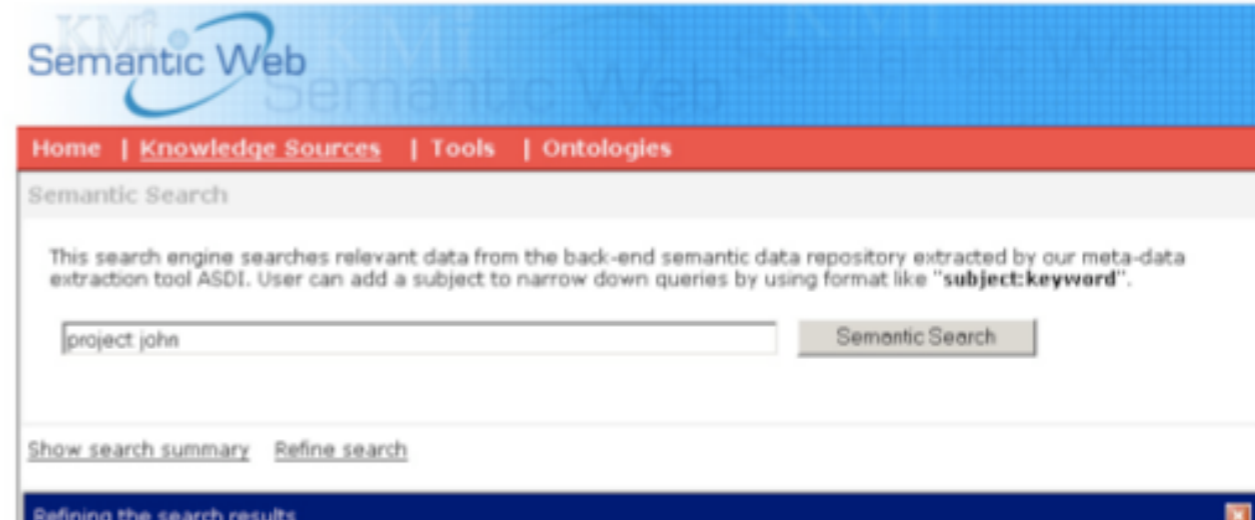
77



Approaches to Semantic Search

- Keyword-based approaches

- ▶ Require translating all the keywords in order to perform the query
 - ▶ E.g. SemSearch



- View-based approaches

- ▶ Based on querying by building visual graphs
 - E.g. Falcon





Search Strategy (ctd)

- A natural language approach

- ▶ E.g. Aqua



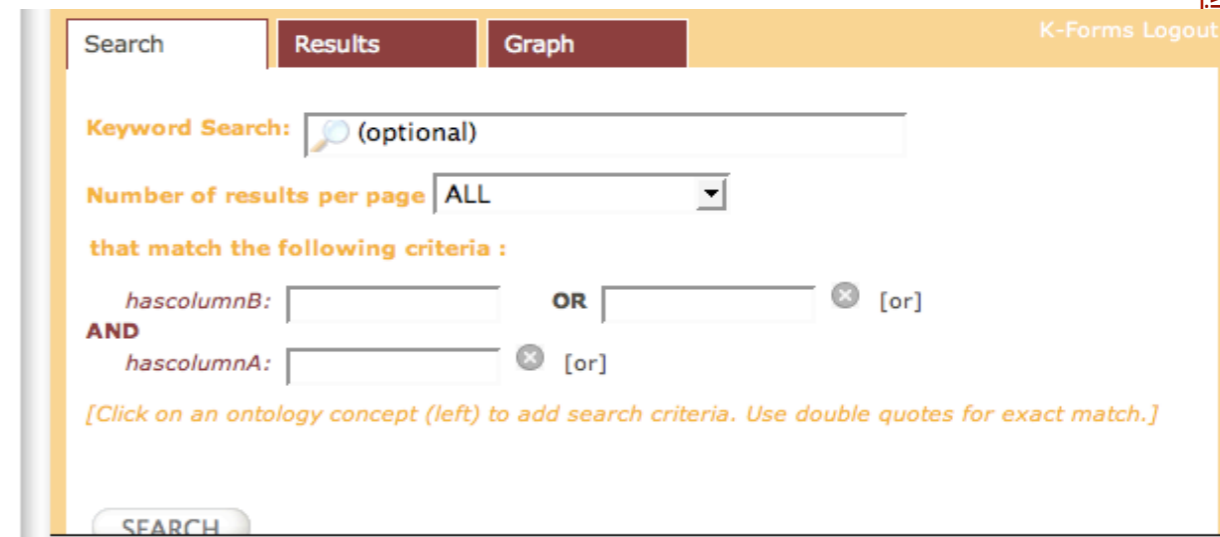
- Form-based approaches

- ▶ e.g. k-search



Available Reports

- test
- ▼ hastable
- ▼ table
 - hascolumnB
 - hascolumnA





Ontology-based Querying: Issues

- Metadata can cover just part of the material of interest to the users
 - ▶ The information not annotated using metadata is irretrievable
 - How many topics can we model with Information Extraction?
 - ▶ 21 topics/ 14 topics partially or not covered by annotations
 - given size of corpus there is no way that manual annotations are added
- Often the use people will do of information is impossible to foresee
 - communities organise forms for themselves
 - some information not structured
 - text fields
 - Sometimes Information is impossible to retrieve reliably using automatic methods
 - If automatic means are used, often some parts of the knowledge is beyond the current technical capability



Issues and Solutions

- **Ontology can be extended**
 - ▶ **But increases effort in indexing**
 - Equivalent to extending metadata in SDM
 - ▶ **But it is impossible to foresee all uses of information**
 - Ontology will always be insufficient somehow
- **Information Extraction can be used to reduce burden of annotation**
 - ▶ **But some parts are irretrievable**



Hybrid Search

- Keywords and ontology-based search can be mixed within the same query
 - ▶ Pure ontology-based searching
 - When metadata covers information precisely
 - ▶ Keyword-in-context of annotation
 - To match strings in text annotated with semantics (textual form fields)
 - ▶ e.g. “fuel” is matched only on snippets of texts annotated as removed parts
 - ▶ General Keyword querying
 - For searching on the document/form as a whole

The screenshot shows the 'know' search interface. On the left, under 'Available Reports', there is a tree structure with 'test', 'hastable', and 'table'. Under 'table', there are two ontology concepts: 'hascolumnB' and 'hascolumnA'. On the right, the search interface has tabs for 'Search', 'Results', and 'Graph'. The 'Search' tab is active. It features a 'Keyword Search' input field with the placeholder 'keyword here'. Below it, a dropdown menu shows 'Number of results per page' set to 'ALL'. The search criteria are displayed as: 'that match the following criteria :'. The criteria are: 'hascolumnB: blah' OR 'burble' (with a close button and '[or]'), and 'AND hascolumnA: blahblah' (with a close button and '[or]'). A note at the bottom says: '[Click on an ontology concept (left) to add search criteria. Use double quotes for match.]'

- Enables querying documents using hybrid search
- Enables quantification of unstructured information
- Currently applied at Rolls-Royce, University of Sheffield and several other organisations

The screenshot displays the K-Search web application interface. On the left, there is a sidebar with the 'k.now' logo and a section for 'Available Reports' listing 'test', 'hastable', and 'table'. The main content area is divided into several sections:

- Search Results:** Shows a 'Keyword Search' field with the text 'key', a 'Number of results per page' field, and a list of results. The results include document titles like 'part2_240400ER_RMR.htm' and 'part2_130401ER_RMR.htm' with brief descriptions.
- Ontology Tree:** A hierarchical tree structure on the left side of the main content area, listing various categories such as 'Event Report', 'Affected Engine', 'Location', 'Part Installed', and 'Part Removed', each with sub-items like 'Part Number', 'Serial Number', and 'Description'.
- Event Report Data:** A detailed view of an event report for 'FS0' on a 'Boeing 777-300'. It includes fields for 'Event Date' (09-Nov-01), 'Aircraft Regn' (9V-SYB), 'Airframe Hours' (574), 'Engine S/N' (51127), 'Installed Posn' (Right), 'Flight Regime' (Hazard Type), and 'Location' (No Hazard).
- Pie Chart:** A circular chart titled 'Pie Chart' showing the distribution of search results. The chart is divided into several segments, each representing a different category and its percentage of the total results. The categories and their percentages are:

Category	Percentage
front combustion outer case	23%
"fan cowl door, lh"	15%
common nozzle assembly	15%
"fan cowl door, rh"	7%
electronic unit protection box	7%
vigv	7%
thrust reverser rh c-duct	7%
lpc front case	7%
fooo	7%



Querying across Ontologies

- **Distributed interconnected resources**
 - ▶ Can be queried across via interconnected ontologies
 - Searching metadata rather than text
 - ▶ Retrieving information independently from the store/media
 - ▶ Enables querying resources using my ontological view
 - largely independently from the view used originally to create it



Searching Across

K-Search (for IPAS) - K-Now Ltd. - Mozilla Firefox

File Edit View History Bookmarks Wired_Marker Tools Help

http://localhost:8080/k-search/

K-Search (for IPAS) - K... Gmail BBC BBC NEWS | News Fro... Internet Banking: HSB... DevX: Semantic Web Z... Main Page - Witwiki Meeting Room Bookin... Semantic Search - Fac...

Disable Cookies CSS Forms Images Information Miscellaneous Outline Resize Tools View Source Options

k.now

Ontology Perspective(s)

- Service Event Report
 - Report Creation Date
 - Report Number
 - File Name
 - Report Author
 - Referred Service Event
 - Service Event
 - Event Description
 - Flight Regime
 - Airframe Cycles
 - Operational Effect
 - Fuel Dumped
 - Event Type
 - Delay Time
 - Airframe Hours
 - Engine Installed Location
 - Event Date
 - Event Category
 - Removed Part
 - Component
 - Part Number
 - Part Description
 - Installed Part
 - Component
 - Part Number

Search Results Graph

Service Event Report Event Summary Report ERMS T800 Technical Variance Module Bulletin

Criteria : Part Description = Fan blade

Document	Part Description
DXB034-01_250201AR	Fan blade
EVR-CES-18-05_080405AR	Fan blades(26 off)
MAS594_130301AR	ANNULUS FILLER - FAN BLADE
BAW-316_290503AR	FAN BLADE ANNULUS FILLER
MAS264_051099AR_issue2	ANNULUS FILLER - FAN BLADE
ER-BKK-923_111002AR	Fan Blade
ER-BKK-922_101002AR	Fan Blade
EVR-CES-008-04_210704MISC	Fan Blades



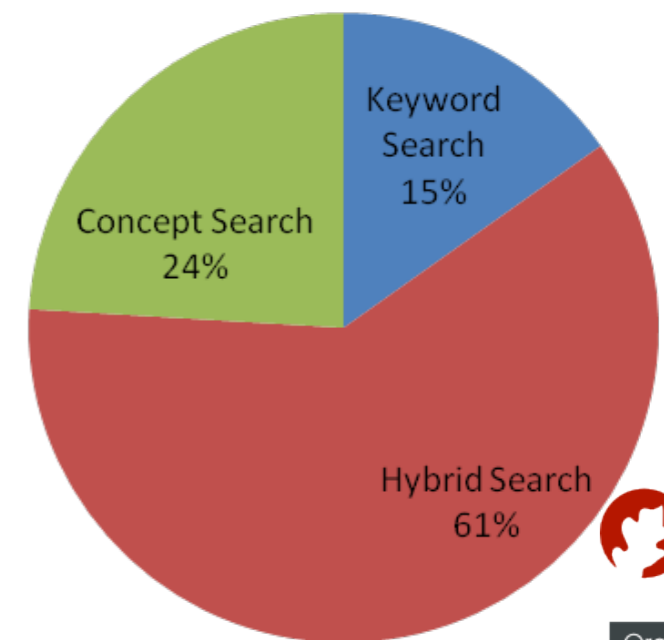
- When an ontology different from the original is used
 - ▶ the original query is mapped to the original ontology via the formal links.
 - ▶ For the parts that are not mapped the restrictions are turned into keywords



Search preferences: Service Engineers

- Service engineers showed a clear predilection for hybrid search:
 - 61% of the search were executed using the hybrid modality
 - 24% using semantic search
 - 15% using keyword search.

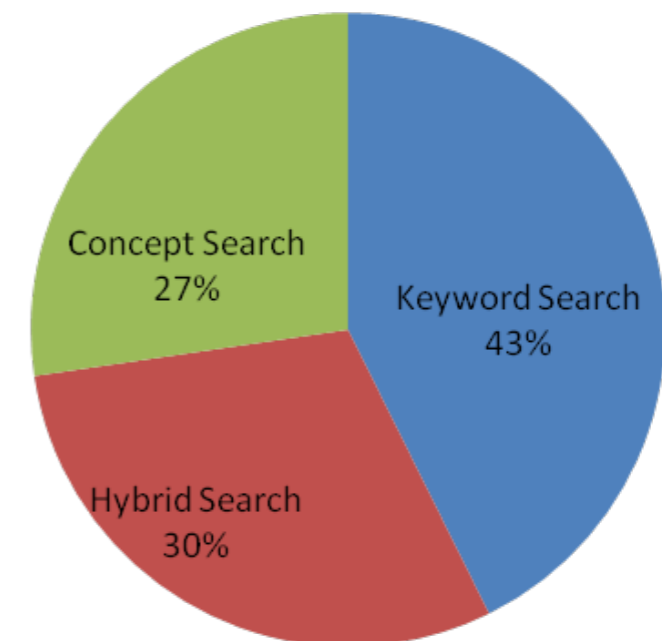
Reason: data they were looking for was not all covered by the metadata





Search preferences: designers

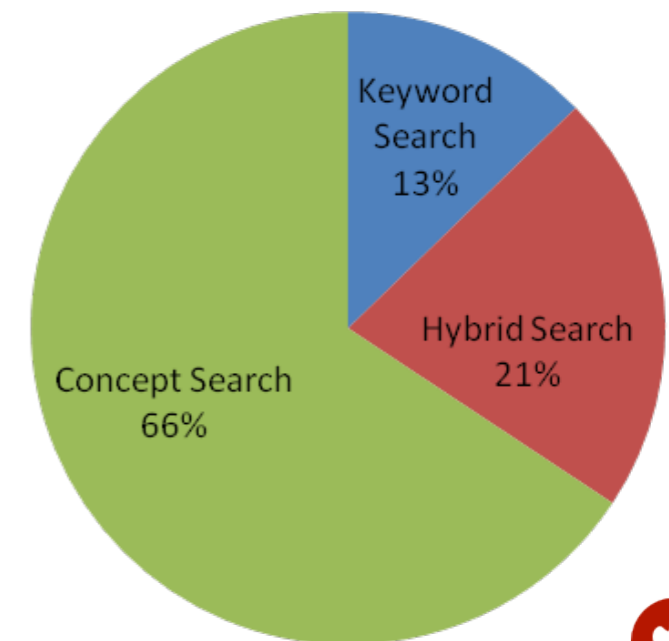
- Designers tended instead to favour keyword search:
 - 43% of the searches were executed using keyword search
 - 30% using hybrid
 - 27% using semantic search.





Search Strategies: Others

- The users belonging to other groups showed a predilection for concept search:
 - 66% of the searches were executed using semantic search
 - 24% using hybrid
 - 15% using keyword search.

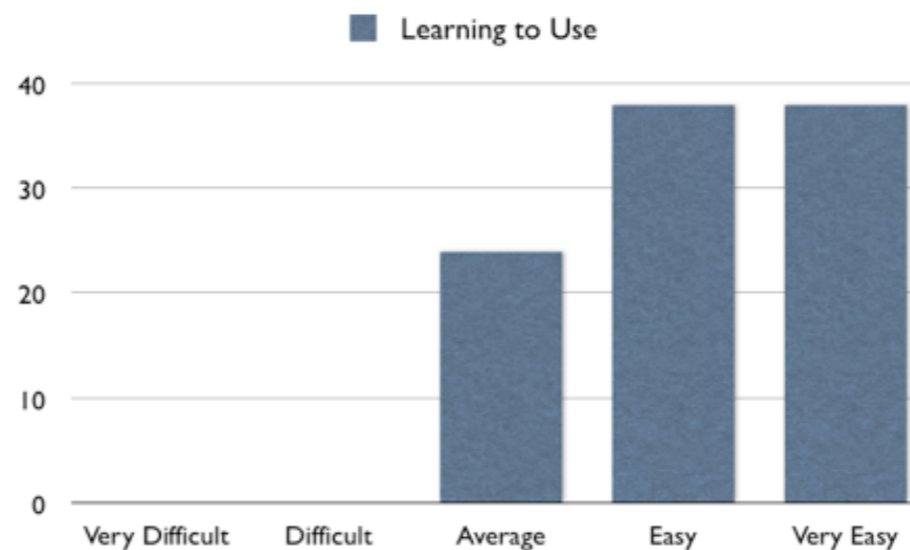
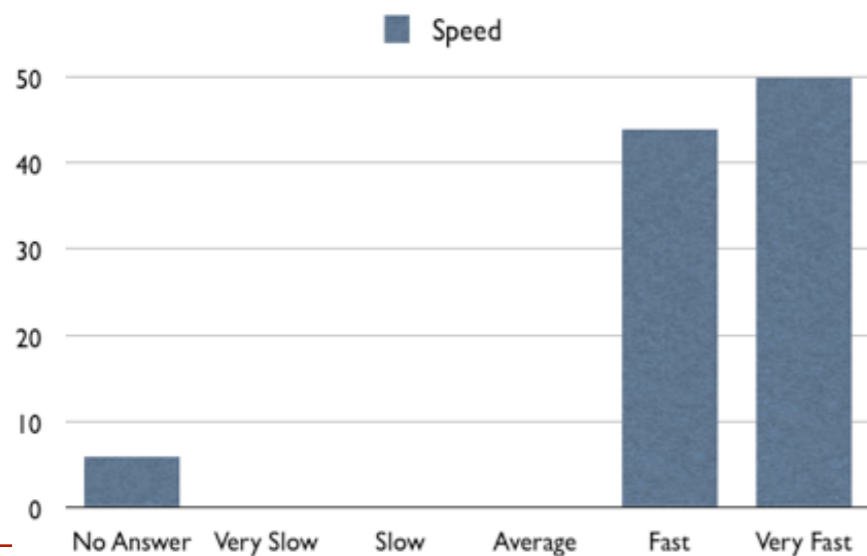
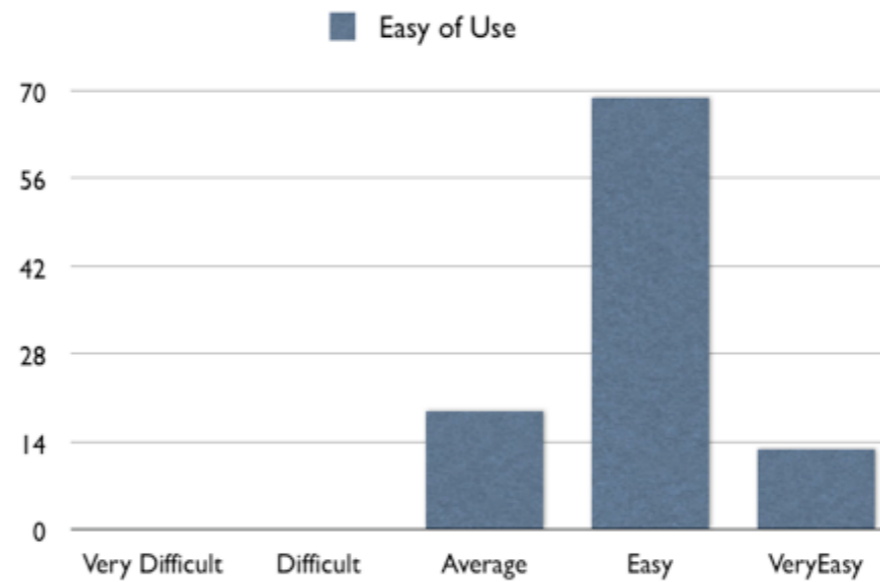
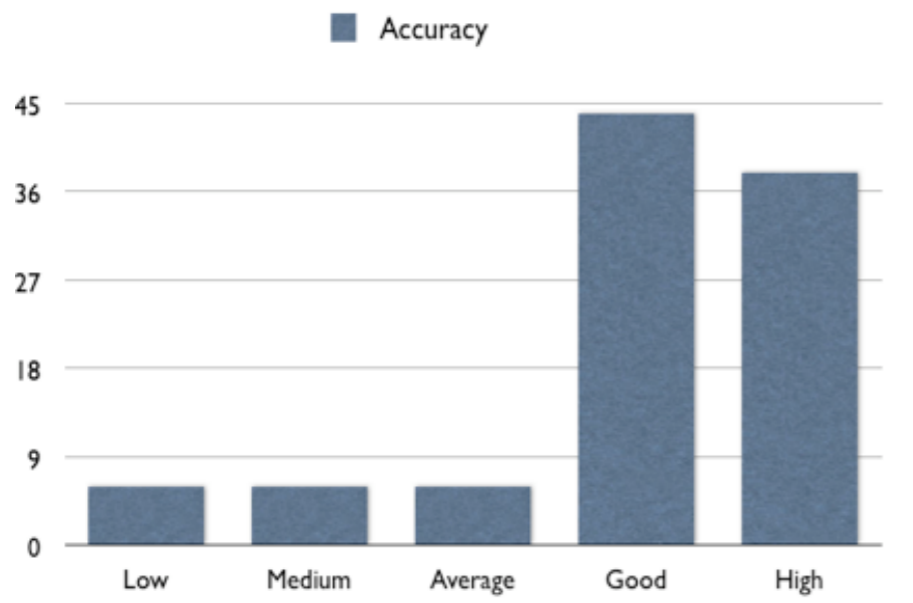
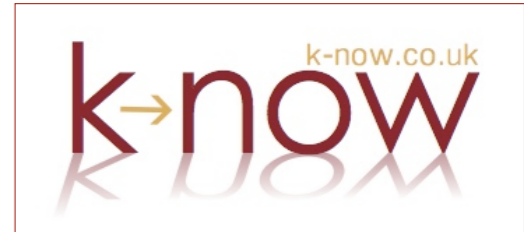




Liked by the users?

- K-Search + IE

- ▶ Finalist of Rolls-Royce Creativity Award 2007
 - Voted by employees for its innovation potential





Liked by Users?

- Developed as part of IPAS collaboration
 - ▶ 2005-2008
 - £240,000 (50% Rolls-Royce)
- Support to the design of new Trent XWB
 - ▶ Porting to 9 Information Sources
 - 2008-2009
 - around £100,000 (100% R-R)
- Funds from Rolls-Royce for use of K-Tools for use in manufacturing
 - ▶ around £340,000



Enriching the User Experience

- Adding knowledge to documents
 - Document enrichment: helping connecting the document to the rest of the knowledge
 - Associating Services
 - Magpie (Dzbor et al. 2004)
 - Connected to other documents
 - COHSE (Goble et al. 2001)

NASA GISS: A Stratospheric "Clock" to Measure Upper Atmosphere Circulation - Microsoft Internet Explorer provided by T

File Edit View Favorites Tools Help

Address http://www.giss.nasa.gov/research/intro/koch_01/ Go Links >>

Magpie Climatology Meteorology Physics Chemistry

collision of high-energy particles from space with nitrogen atoms in the atmosphere. Most tracer production occurs between about 30° 70° latitude in both hemispheres of the lower stratosphere, as indicated by the circled regions on the figure. These tracers, which are borne on aerosol particles, are removed from the stratosphere by radioactive decay. While beryllium-7 decays relatively quickly, with a half-life of 53 days, ¹⁰Be's decay rate is negligible. The only sink for ¹⁰Be occurs after it enters the troposphere, where the radionuclides are efficiently removed by precipitation. Therefore, if we look at the ratio of ¹⁰Be/⁷Be as air moves from the midlatitude production region to other parts of the stratosphere, the ratio will generally increase, as ⁷Be decays. Thus, the ¹⁰Be/⁷Be acts as a "clock" of air mass age.

The figure shows the ¹⁰Be/⁷Be ratio calculated in the GISS general circulation model (GCM) during January and March. In the tropical stratosphere, air rises from the troposphere and continues to ascend, but exchange with higher latitudes is inhibited. The ¹⁰Be/⁷Be ratio is very high (white region) since slow penetration of air from the mid-latitude production region allows much of the ⁷Be to decay. During the early northern hemisphere spring, air from the lower tropical stratosphere moves to higher latitudes relatively quickly. The result is the green blob of relatively high ¹⁰Be/⁷Be air at

March

Altitude (km)

Latitude

Ratio of ¹⁰Be/⁷Be

¹⁰Be/⁷Be ratio calculated in the GISS general circulation model during January and March. Circled areas indicate maximum

- Explain concept
- Relevant parts in S199
- Analysis of effects by IPCC
- CPDN results analysis
- Background reading
- Scientific articles

Done Internet

Sparks O₃ Browser

Augmenting the Web with Semantic Overlays

Grégoire Burel¹, Amparo E. Cano¹ and Vitaveska Lanfranchi¹

¹OAK Group, Department of Computer Science,
University of Sheffield

{G.Burel, A.Cano, V.Lanfranchi}@dcs.shef.ac.uk



ESWC09, 5th Workshop on Scripting and Development for the Semantic Web – 31st May 2009
Enlighten the Web.





My Work at W3C

I am [Semantic Web Activity Lead](#); that is my main work at [W3C](#). I am member of [IW3C2 \(International World Wide Web Conference Committee\)](#) (the committee coordinating the yearly WWW conference series), serving as a liaison for W3C, and of [SWSA \(Semantic Web Science Association\)](#), the committee responsible for the International Semantic Web Conferences series.

As part of my work, I also participate in lots of outreach activities, and I regularly make presentations, tutorials, etc. You can consult my [list of presentations](#) for further details.

Contact information

Email:

ivan@w3.org
(sha1sum: 5ac8032d5f6012aa1775ea2f3e1676bafd5e80b)

Postal address:

C/o Centre for Mathematics and Computer Sciences (CWI)
Kruislaan 413, P.O. Box 94079, 1090 CB Amsterdam, The Netherlands.

Phone numbers:

phone: +31-20-5924163
mobile phone: +31-64-044153
fax: +31-20-5924312

PGP/GPG:

My [GnuPGP key and signature](#) is available on-line.

FOAF:

You can either extract a short [FOAF information](#) from this page, or consult my [more complete, public FOAF file](#).

Misc:

I am often on freenode, (acc. name [IvanHerman](#); primarily on the [#swig](#) channel).
I am (of course...) present on a number of online accounts and services, like: [LinkedIn](#) (acc. number [2352277](#)), [Dopplr](#) (acc. name [IvanHerman](#)), [TripIt](#) (acc. name [ivan_herman](#)), [Twitter](#) (acc. name [ivan_herman](#)), [Flickr](#) (acc. [ivan_herman](#)).
My URI (as a real person): <http://www.ivan-herman.net/foaf#me>

Short CV

I graduated as mathematician at the [Eötvös Loránd University of Budapest](#), Hungary, in 1979. After a brief scholarship at the Université Paris VI I joined the Hungarian research institute in computer science ([SZTAKI](#)) where I worked for 6 years (and turned into a computer scientist...). I left Hungary in 1986 and, after a few years in industry in Munich, Germany, I joined the [Centre for Mathematics and Computer Sciences \(CWI\)](#) in Amsterdam where I have a tenure position since 1988. I received a PhD degree in [Computer Science](#) in 1990 at the [University of Leiden](#), in the Netherlands. I joined the [W3C Team](#) as Head of [W3C Offices](#) in [Amsterdam](#). I served as Head of Offices until June 2006, when I was asked to take the [position at CWI](#), which is now my principal work at W3C.

I have also worked in quite different areas (distributed and dataflow programming, language design, system research years in computer graphics and information visualization. I also participated in [outreach](#) activities and software developments. My [separate "professional" home page](#) contains





Upcoming Trips

- 7 to 22 May, 2009: **Visits around Australia, organized by the local W3C Office**
- 13 to 21 June, 2009: **Semantic Technology Conference, San Jose, CA, USA**

References to my public presentations

I have a number of [slide sets "in progress"](#), which I use for finalized presentations I have given or will give at various events.

The last 6 months:

15 January

Ivan Herman gives an invited talk on behalf of the [Benelux Office](#) ("Years Reception)" on Thursday, 15 January 2009, in Amsterdam,

16 March

Ivan Herman gives a talk entitled "[Some W3C SW technologies](#)" in Amsterdam, The Netherlands.

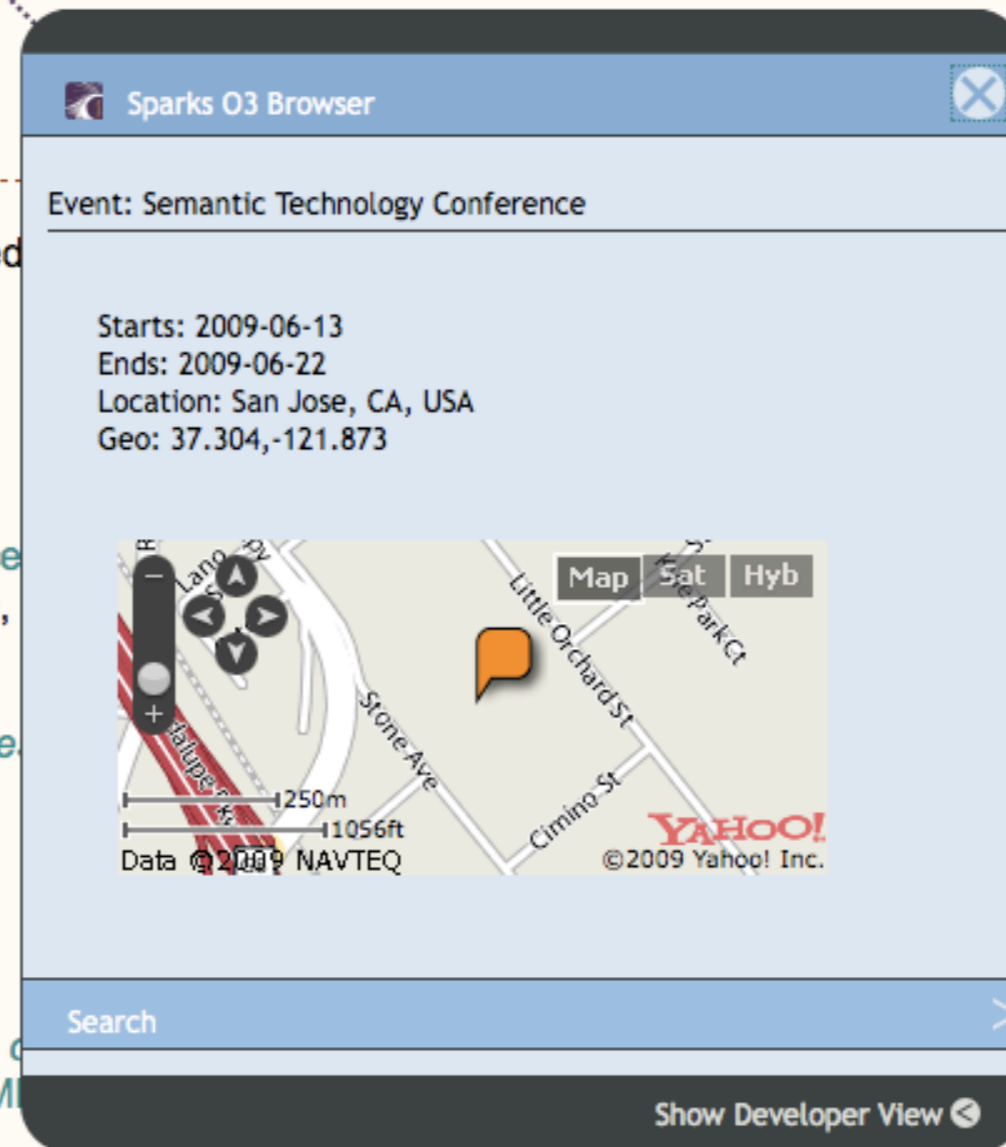
Upcoming:

12 May

Ivan Herman gives a talk entitled "[Introduction and Applications of Semantic Web Technologies](#)" in the [Presentation series "The future of web standards, HTML5, XHTML5"](#)

13 May

Ivan Herman gives a talk on behalf of the [Australia Office](#) entitled "[Introduction and Applications of Semantic Web](#)" at the "Presentation series "T



Knowledge Visualisation



Visualisation

- “the Semantic Web emphasises formal, machine readable [...] approaches. It focuses on the formal and even the meaning achieved through rigorously defined forms.
- Information visualization emphasizes the semantics and the meaning that can be conveyed by visual-spatial models to the users.” [6].



Requirements

- Going beyond the search/retrieval of documents and facts
 - Supporting exploration
 - Towards identification of trends
- Support flexible data exploration (search and browse) to make sense of a complex environment
 - data can be very dense (large amount, very similar)
 - ontology can be very large, several repositories
- Effective and efficient access to data



Semantic Data in Context

- Visualizations according to some main semantic dimensions familiar to users
 - to make interaction transparent
 - to engage them in data exploration
- Data exploration via dynamic query according to all semantic dimensions
 - supports personal investigation strategies
 - instantaneous change of investigation path



Example: Application to Jet Engines

- User studies found that most appropriate are:
 - Visualizations:
 - Time: show reoccurring events
 - Geography: show flying-root related phenomena
 - Topology: show part(s) of the engine involved
 - Data exploration:
 - According to dozens of filters
 - Defined by a company wide ontology

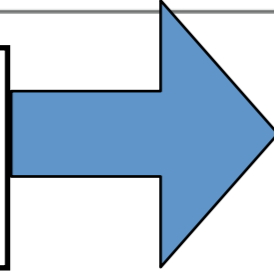
```
<rdf:Description rdf:about="http://kmi.open.ac.uk/projects/xmedia/
RR1.owl#Event_Report.BKK.Event_Report_237">
<rdf:type rdf:resource="http://kmi.open.ac.uk/projects/xmedia/RR1.owl#Event_Report"/>
<j.0:has_file_location>BKK/Event_Report_237</j.0:has_file_location>
<j.0:hasFormattedEventDate>26-Jul-1922</j.0:hasFormattedEventDate>
<j.0:hasEventDate>26-Jul-22</j.0:hasEventDate>
<j.0:hasAssociatedDate>28-Aug-22</j.0:hasAssociatedDate>
<j.0:hasTSN>14613</j.0:hasTSN>
<j.0:hasEngine_Serial_Number>2551.55</j.0:hasEngine_Serial_Number>
<j.0:hasLocation>BKK</j.0:hasLocation>
<j.0:hasRegime>GROUND</j.0:hasRegime>
<j.0:hasCSN>5362</j.0:hasCSN>
<j.0:hasComponent>Fuel</j.0:hasComponent>
</rdf:Description>
```


RDF triples



The
University
Of
Sheffield.

```
<rdf:Description rdf:about="http://kmi.open.ac.uk/projects/amedia/
RR.Low/Event_Report/BKK_Event_Report_237">
<rdf:type rdf:resource="http://kmi.open.ac.uk/projects/amedia/RR.Low/Event_Report"/>
<?has_file_location=BKK/Event_Report_237?>
<?hasFormattedEventDate=26-Jul-1922?>
<?hasEventDate=26-Jul-22?>
<?hasAssociatedDate=28-Aug-22?>
<?hasTSN=14613?>
<?hasEngine_Serial_Number=251.55?>
<?hasLocation=BKK?>
<?hasRegime=GROUND?>
<?hasCSN=5362?>
<?hasComponent=Fuel?>
<rdf:Description>
```



Unplotted Documents									
has_file_location	hasLocation	hasEventDate	hasEngine_Serial_Number	hasCategory	hasType	hasTSN	hasCSN	hasFormat...	ha
[BKK/Event_Report_174]	[UBP]	[18-May-22]	[2553.25, 2553.05]			[12136]	[4193]	[18-May-...	
[BKK/Event_Report_192]	[UBP]	[28-May-22]	[2553.05]			[11685]	[4482]	[28-May-...	[G
[BKK/Event_Report_29]	[TPE]	[26-Aug-21]	[2553.95, 2550.05]			[8365]	[2878]	[26-Aug-...	[T
[DXB/Event_Report_44]	[TIP]	[15/05/23]	[2058.65]	[Basic]	[Delay]			[15-May-...	
[DXB/Event_Report_78]	[THR]	[16/06/23]	[2053.65]	[Non-Basic]	[Delay]			[16-Jun-1...	
[DXB/Event_Report_149]	[SYD]	[07/09/23]	[2569.35]	[Non-Basic]	[Cancellation]			[07-Sep-...	
[DXB/Event_Report_6]	[SYD]	[18/01/24]	[3557.35/3554.00]	[Basic]	[Delay]			[18-Jan-1...	
[DXB/Event_Report_74]	[SIN]	[12/06/23]	[2552.55]	[Basic]	[Delay]			[12-Jun-1...	
[DXB/Event_Report_108]	[SIN]		[2060.80]	[Basic]	[Delay]			[23-Jul-1...	
[BKK/Event_Report_2]	[SIN]	[31-Oct-21]	[2551.55, 2551.80, 2...			[13699]	[5267]	[31-Oct-...	[C
[DXB/Event_Report_195]	[SAH]	[15/10/23]	[2052.90]	[Non-Basic]	[Air Turnback]			[15-Oct-...	
[DXB/Event_Report_122]	[NBO]	[5/08/23]	[2061.20]	[Basic]	[Delay]			[05-Aug-...	
[DXB/Event_Report_169]	[NBO]	[29/09/23]	[2061.85]	[Non-Basic]	[Delay]			[29-Sep-...	
[DXB/Event_Report_121]	[NBO]	[05/08/23]	[2058.00, 2]	[Basic]	[Delay]			[05-Aug-...	
[DXB/Event_Report_118]	[NBO]		[2060.30]	[Basic]	[Delay]			[03-Aug-...	
[DXB/Event_Report_143]	[NBO]		[2053.90]	[Basic]	[Delay]			[23-Aug-...	
[DXB/Event_Report_71]	[MXP]	[11/06/23]	[2560.35]	[Non-Basic]	[ABTO(high speed)/...			[11-Jun-1...	
[DXB/Event_Report_151]	[MUC]	[09/09/23]	[2054.95]	[Basic]	[Delay]			[09-Sep-...	
[DXB/Event_Report_125]	[MEL]	[31/7/23]	[2568.90]	[Open]	[Cancellation]			[31-Jul-1...	
[DXB/Event_Report_115]	[MEL]		[2055.95]	[Non-Basic]	[Delay]			[28-Jul-1...	
[DXB/Event_Report_194]	[MCT]		[2058.25]	[Basic]	[Delay]			[26-Oct-...	
[AIH/Event_Report_1]	[MAN]	[24/05/22]	[2057.40]			[0]		[24-May-...	[G
[DXB/Event_Report_65]	[MAN]	[08/06/23]	[2062.50]	[Non-Basic]	[Delay]			[08-Jun-1...	
[DXB/Event_Report_81]	[MAN]	[26/06/23]	[2056.75]	[Basic]	[Delay]			[26-Jun-1...	
[DXB/Event_Report_200]	[LHR]	[01/11/23]	[2571.45]	[Non-Basic]	[Delay]			[01-Nov-...	
[DXB/Event_Report_75]	[LHR]	[14/06/23]	[2059.45]	[Basic]	[Delay]			[14-Jun-1...	
[DXB/Event_Report_58]	[LGW]	[28/05/23]	[2058.00]	[Basic]	[Delay]			[28-May-...	
[DXB/Event_Report_49]	[KUL]	[19/05/23]	[2552.55]	[Basic]	[Delay]			[19-May-...	
[DXB/Event_Report_170]	[KHI]		[2061.85]	[Basic]	[Delay]			[01-Oct-...	
[DXB/Event_Report_161]	[IST]	[23/09/23]	[2062.80]	[Basic]	[Diversion/Delay]			[23-Sep-...	
[BKK/Event_Report_61]	[ICN]	[24-Nov-21]	[2555.70]			[10917]	[2805]	[24-Nov-...	[G
[BKK/Event_Report_157]	[ICN]	[18-Apr-22]	[2553.65]			[12477]	[4572]	[18-Apr-...	[G
[BKK/Event_Report_225]	[HKG]	[07-Jul-22]				[7611]	[1893]	[07-Jul-1...	[G
[BKK/Event_Report_250]	[HKG]	[17-Sep-22]	[2550.55]			[16608]	[6427]	[17-Sep-...	[G
[BKK/Event_Report_56]	[HKG]	[26-Nov-21]	[2550.90]			[11478]	[4180]	[26-Nov-...	[G

A row is a document, a column is a concept

A cell is a triple

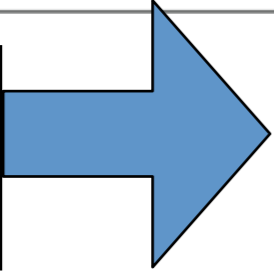


RDF triples



The University Of Sheffield.

```
<rdf:Description rdfs:about="http://kmi.open.ac.uk/projects/amedia/RR1/low/Event_Report_BKK/Event_Report_237">
<rdf:type rdf:resource="http://kmi.open.ac.uk/projects/amedia/RR1/low/Event_Report"/>
<?has_file_location=BKK/Event_Report_237?>
<?hasFormatedEventDate=26-Jul-1922?>
<?hasEventDate=26-Jul-22?>
<?hasAssociatedDate=28-Aug-22?>
<?hasTSN=14613?>
<?hasEngine_Serial_Number=2551.55?>
<?hasLocation=BKK?>
<?hasRegime=GROUND?>
<?hasCSN=5362?>
<?hasComponent=Fuel?>
<?hasDescription?>
```



RDF table

has_file_location	hasLocation	hasEventDate	hasEngine_Serial_Number	hasCategory	hasType	hasTSN	hasCSN	hasFormat...
[BKK/Event_Report_174]	[BKP]	[18-May-22]	[2553.25, 2553.05]			[12136]	[4193]	[18-May-...
[BKK/Event_Report_194]	[BKP]	[26-Aug-21]	[2553.95, 2550.05]			[11685]	[6482]	[26-Aug-...
[BKK/Event_Report_29]	[TPE]	[15/05/23]	[2058.65]	[Basic]	[Delay]	[8365]	[2878]	[26-Aug-...
[DNB/Event_Report_44]	[TPE]	[15/05/23]	[2058.65]	[Non-Basic]	[Delay]			[15-May-...
[DNB/Event_Report_78]	[THRE]	[16/06/23]	[2055.65]	[Non-Basic]	[Cancellation]			[16-Jun-...
[DNB/Event_Report_149]	[SYD]	[07/09/23]	[2569.35]	[Non-Basic]	[Cancellation]			[07-Sep-...
[DNB/Event_Report_8]	[SYD]	[18/01/24]	[3557.35/3554.00]	[Basic]	[Delay]			[18-Jan-...
[DNB/Event_Report_74]	[SNI]	[12/06/23]	[2055.55]	[Basic]	[Delay]			[12-Jun-...
[DNB/Event_Report_108]	[SNI]	[12/06/23]	[2060.80]	[Basic]	[Delay]			[23-Jul-...
[BKK/Event_Report_2]	[SNI]	[31-Oct-21]	[2551.55, 2551.80, 2...	[Non-Basic]	[Air Turnback]	[13699]	[5267]	[31-Oct-...
[DNB/Event_Report_195]	[SNI]	[15/10/23]	[2055.90]	[Basic]	[Delay]			[15-Oct-...
[DNB/Event_Report_122]	[NBO]	[5/08/23]	[2061.20]	[Basic]	[Delay]			[05-Aug-...
[DNB/Event_Report_169]	[NBO]	[29/09/23]	[2061.85]	[Non-Basic]	[Delay]			[29-Sep-...
[DNB/Event_Report_121]	[NBO]	[05/08/23]	[2058.00, 2]	[Basic]	[Delay]			[05-Aug-...
[DNB/Event_Report_118]	[NBO]	[05/08/23]	[2060.30]	[Basic]	[Delay]			[03-Aug-...
[DNB/Event_Report_143]	[NBO]	[05/08/23]	[2053.90]	[Basic]	[Delay]			[23-Aug-...
[DNB/Event_Report_71]	[NBP]	[11/06/23]	[2360.35]	[Non-Basic]	[AETC high speed]...			[11-Jun-...
[DNB/Event_Report_151]	[MUC]	[09/09/23]	[2054.95]	[Basic]	[Delay]			[09-Sep-...
[DNB/Event_Report_125]	[MEL]	[31/7/23]	[2366.90]	[Open]	[Cancellation]			[31-Jul-...
[DNB/Event_Report_135]	[MEL]	[31/7/23]	[2055.95]	[Non-Basic]	[Delay]			[28-Jul-...
[DNB/Event_Report_194]	[MCT]	[2058.25]		[Basic]	[Delay]			[28-Oct-...
[API/Event_Report_1]	[MAN]	[24/05/22]	[2057.40]			[0]		[24-May-...
[DNB/Event_Report_85]	[MAN]	[08/06/23]	[2062.50]	[Non-Basic]	[Delay]			[08-Jun-...
[DNB/Event_Report_81]	[MAN]	[26/06/23]	[2056.75]	[Basic]	[Delay]			[26-Jun-...
[DNB/Event_Report_200]	[LHR]	[01/11/23]	[2571.45]	[Non-Basic]	[Delay]			[01-Nov-...
[DNB/Event_Report_75]	[LHR]	[14/06/23]	[2059.45]	[Basic]	[Delay]			[14-Jun-...
[DNB/Event_Report_58]	[LGW]	[28/05/23]	[2058.00]	[Basic]	[Delay]			[28-May-...
[DNB/Event_Report_49]	[KLA]	[19/05/23]	[2552.55]	[Basic]	[Delay]			[19-May-...
[DNB/Event_Report_170]	[KHE]	[2061.85]		[Basic]	[Delay]			[01-Oct-...
[DNB/Event_Report_161]	[IST]	[23/09/23]	[2062.80]	[Basic]	[Diversion/Delay]			[23-Sep-...
[BKK/Event_Report_61]	[ICN]	[24-Nov-21]	[2555.70]			[10917]	[2805]	[24-Nov-...
[BKK/Event_Report_157]	[ICN]	[18-Apr-22]	[2553.65]			[12477]	[4572]	[18-Apr-...
[BKK/Event_Report_225]	[HKG]	[07-Jul-21]				[7611]	[1893]	[07-Jul-...
[BKK/Event_Report_250]	[HKG]	[17-Sep-22]	[2550.55]			[16608]	[6427]	[17-Sep-...
[BKK/Event_Report_56]	[HKG]	[26-Nov-21]	[2550.90]			[11478]	[6380]	[26-Nov-...



RDF triples



The University Of Sheffield.

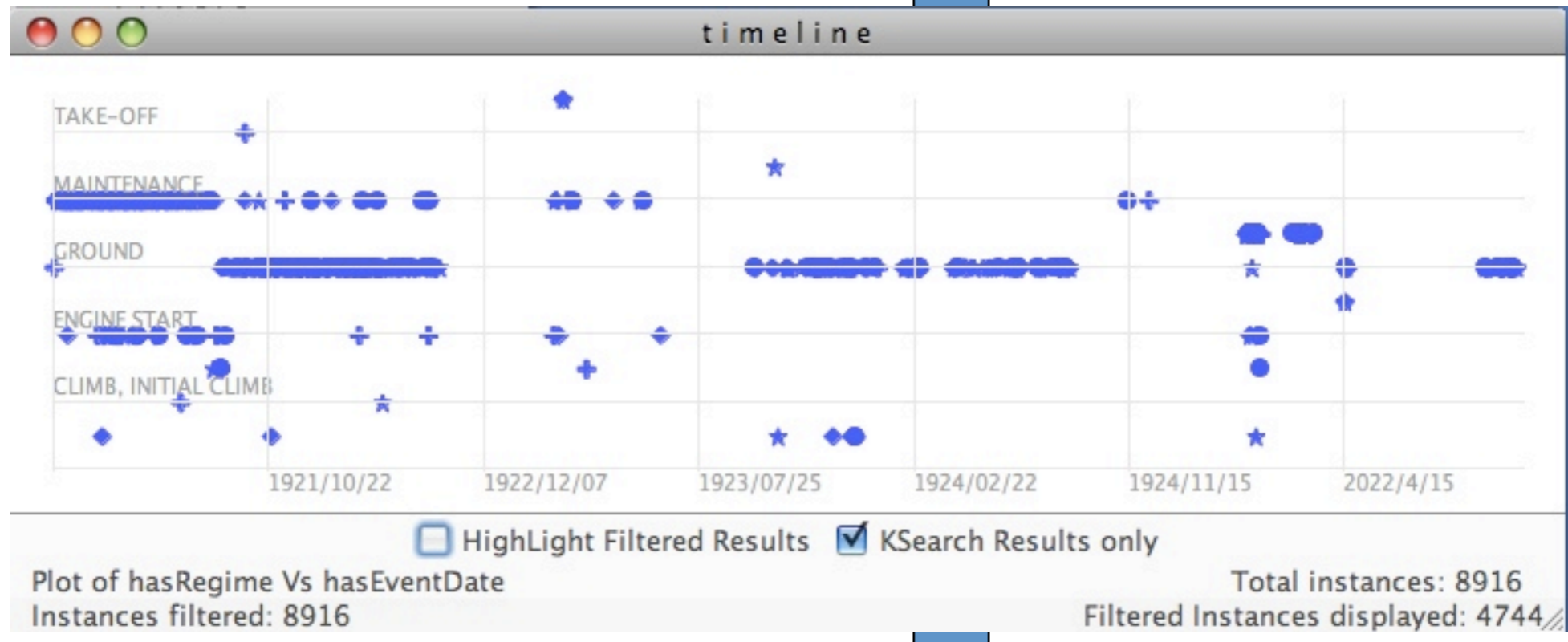
RDF table

hasFileLocation	hasLocation	hasEventDate	hasEngine_Serial_Number	hasCategory	hasType	hasTSN	hasCSN	hasFormat...	has...
[BKK:Event_Report_174]	[BPP]	[18-May-22]	[2553.25, 2553.05]			[12136]	[4193]	[18-May-...	[C]
[BKK:Event_Report_194]	[BPP]	[28-May-22]	[2553.05]			[11685]	[6482]	[28-May-...	[C]
[BKK:Event_Report_29]	[TPE]	[26-Aug-21]	[2553.95, 2550.05]			[8365]	[2878]	[26-Aug-...	[T]
[DNB:Event_Report_44]	[TPE]	[15/05/23]	[2058.65]	[Basic]	[Delay]			[15-May-...	
[DNB:Event_Report_78]	[THR]	[16/06/23]	[2058.65]	[Non-Basic]	[Delay]			[16-Jun-...	
[DNB:Event_Report_149]	[SYD]	[07/09/23]	[2569.35]	[Non-Basic]	[Cancellation]			[07-Sep-...	
[DNB:Event_Report_6]	[SYD]	[18/01/24]	[3557.35/3554.00]	[Basic]	[Delay]			[18-Jan-...	
[DNB:Event_Report_74]	[SN]	[12/06/23]	[2552.55]	[Basic]	[Delay]			[12-Jun-...	
[DNB:Event_Report_108]	[SN]		[2060.80]	[Basic]	[Delay]			[23-Jul-...	
[BKK:Event_Report_2]	[SN]	[31-Oct-21]	[2551.55, 2551.80, 2...			[13699]	[5267]	[31-Oct-...	[C]
[DNB:Event_Report_195]	[SAB]	[15/10/23]	[2052.90]					[15-Oct-...	
[DNB:Event_Report_122]	[NBO]	[5/08/23]	[2061.20]	[Non-Basic]	[Air Turnback]			[05-Aug-...	
[DNB:Event_Report_169]	[NBO]	[29/09/23]	[2061.85]	[Non-Basic]	[Delay]			[29-Sep-...	
[DNB:Event_Report_121]	[NBO]	[05/08/23]	[2058.00, 2]	[Basic]	[Delay]			[05-Aug-...	
[DNB:Event_Report_118]	[NBO]		[2060.30]	[Basic]	[Delay]			[03-Aug-...	
[DNB:Event_Report_143]	[NBO]		[2053.90]	[Basic]	[Delay]			[23-Aug-...	
[DNB:Event_Report_71]	[NBP]	[11/06/23]	[2360.35]	[Non-Basic]	[ATCO high speed]...			[11-Jun-...	
[DNB:Event_Report_151]	[MUC]	[09/09/23]	[2054.95]	[Basic]	[Delay]			[09-Sep-...	
[DNB:Event_Report_125]	[MEL]	[31/7/23]	[2368.90]	[Open]	[Cancellation]			[31-Jul-...	
[DNB:Event_Report_131]	[MEL]		[2055.95]	[Non-Basic]	[Delay]			[28-Jul-...	
[DNB:Event_Report_194]	[MCT]		[2058.25]	[Basic]	[Delay]			[28-Oct-...	
[BKK:Event_Report_1]	[MAN]	[24/05/22]	[2057.40]			[0]		[24-May-...	[C]
[DNB:Event_Report_65]	[MAN]	[08/06/23]	[2062.50]	[Non-Basic]	[Delay]			[08-Jun-...	
[DNB:Event_Report_81]	[MAN]	[26/06/23]	[2056.75]	[Basic]	[Delay]			[26-Jun-...	
[DNB:Event_Report_200]	[LHR]	[01/11/23]	[2571.45]	[Non-Basic]	[Delay]			[01-Nov-...	
[DNB:Event_Report_75]	[LHR]	[14/06/23]	[2059.45]	[Basic]	[Delay]			[14-Jun-...	
[DNB:Event_Report_58]	[LGW]	[28/05/23]	[2058.00]	[Basic]	[Delay]			[28-May-...	
[DNB:Event_Report_49]	[KJA]	[19/05/23]	[2552.55]	[Basic]	[Delay]			[19-May-...	
[DNB:Event_Report_170]	[KJE]		[2061.85]	[Basic]	[Delay]			[01-Oct-...	
[DNB:Event_Report_161]	[IST]	[23/09/23]	[2062.80]	[Basic]	[Diversion/Delay]			[23-Sep-...	
[BKK:Event_Report_61]	[ICN]	[24-Nov-21]	[2555.70]			[10917]	[2805]	[24-Nov-...	[C]
[BKK:Event_Report_157]	[ICN]	[18-Apr-22]	[2553.65]			[12477]	[4572]	[18-Apr-...	[C]
[BKK:Event_Report_225]	[HKG]	[07-Jul-21]				[7611]	[1893]	[07-Jul-...	[C]
[BKK:Event_Report_250]	[HKG]	[17-Sep-22]	[2550.55]			[16608]	[6427]	[17-Sep-...	[C]
[BKK:Event_Report_165]	[HKG]	[26-Nov-21]	[2550.90]			[11478]	[6300]	[26-Nov-...	[C]

```

<rdf:Description rdfs:about="http://kmi.open.ac.uk/projects/amedia/RR1.LowIfEvent_Report_BKK_Event_Report_237">
<rdf:type rdf:resource="http://kmi.open.ac.uk/projects/amedia/RR1.LowIfEvent_Report"/>
<rdf:has_file_location-BKK/Event_Report_237/>
<rdf:hasFormatofEventDate>26-Jul-1922/>
<rdf:hasEventDate>26-Jul-22/>
<rdf:hasAssociatedDate>28-Aug-22/>
<rdf:hasTSN>14613/>
<rdf:hasEngine_Serial_Number>2551.55/>
<rdf:hasLocation>BKK/>
<rdf:hasRegime>GROUND/>
<rdf:hasCSN>5362/>
<rdf:hasComponent-Fuel/>
<rdf:Description>

```



RDF triples



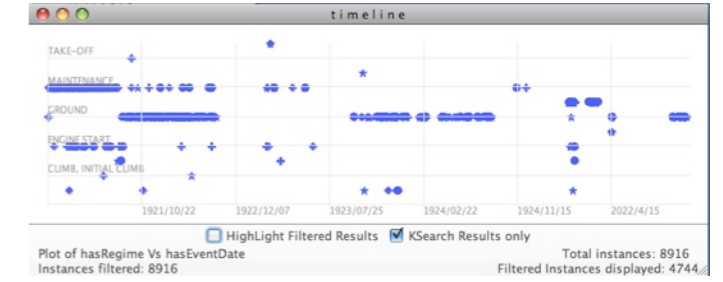
The University of Sheffield.

```
<rdf:Description rdfs:about="http://kmi.open.ac.uk/projects/amedia/RR1LowIfEvent_Report_BKK_Event_Report_237">
  <rdf:type rdf:resource="http://kmi.open.ac.uk/projects/amedia/RR1LowIfEvent_Report"/>
  <j.0.has_file_location-BKK/Event_Report_237.rj.0.has_file_location>
  <j.0.hasFormattedEventDate-26-Jul-1922.rj.0.hasFormattedEventDate>
  <j.0.hasEventDate-26-Jul-22.rj.0.hasEventDate>
  <j.0.hasAssociatedDate-28-Aug-22.rj.0.hasAssociatedDate>
  <j.0.hasTSN-14613.rj.0.hasTSN>
  <j.0.hasEngine_Serial_Number-2551.55.rj.0.hasEngine_Serial_Number>
  <j.0.hasLocation-BKK.rj.0.hasLocation>
  <j.0.hasRegime-GROUND.rj.0.hasRegime>
  <j.0.hasCSN-5362.rj.0.hasCSN>
  <j.0.hasComponent-Fuel.rj.0.hasComponent>
  <rdf:Description>
```

RDF table

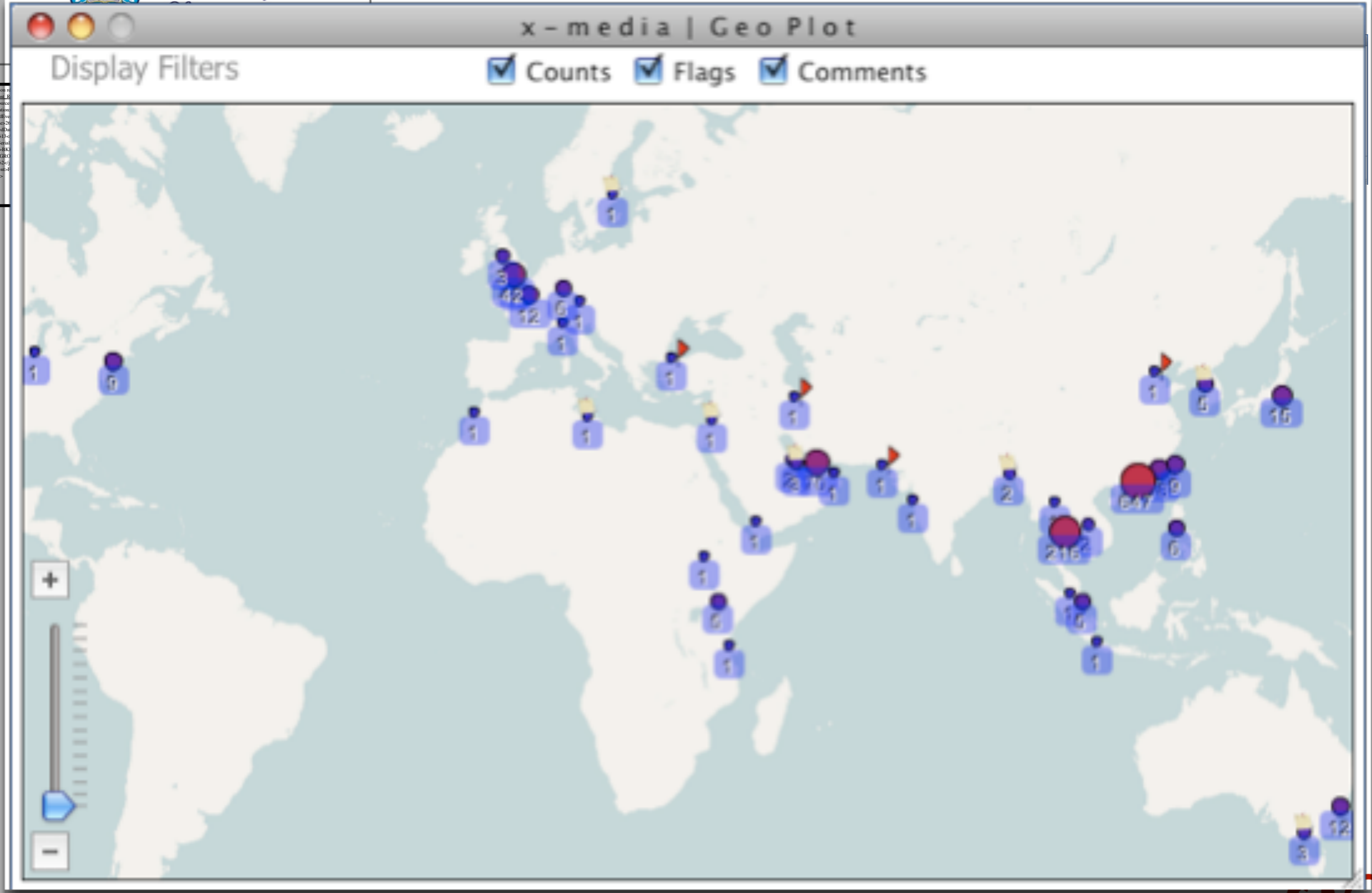
has_file_location	hasLocation	hasEventDate	hasEngine_Serial_Number	hasCategory:hasType	hasTSN	hasCSN	hasFormat...
[BKK/Event_Report_174]	[BKP]	[18-May-22]	[2553.25, 2553.05]		[12136]	[4193]	[18-May-...
[BKK/Event_Report_194]	[BKP]	[28-May-22]	[2553.05]		[11685]	[6482]	[28-May-...
[BKK/Event_Report_29]	[TPE]	[26-Aug-21]	[2553.95, 2550.05]		[8365]	[2878]	[26-Aug-...
[DNB/Event_Report_44]	[TPE]	[15/05/23]	[2058.65]	[Basic] [Delay]			[15-May-...
[DNB/Event_Report_78]	[THRE]	[16/06/23]	[2058.65]	[Non-Basic] [Delay]			[16-Jun-...
[DNB/Event_Report_149]	[SYD]	[07/09/23]	[2569.35]	[Non-Basic] [Cancellation]			[07-Sep-...
[DNB/Event_Report_8]	[SYD]	[18/01/24]	[3557.35/3554.00]	[Basic] [Delay]			[18-Jan-...
[DNB/Event_Report_74]	[SNO]	[12/06/23]	[2052.55]	[Basic] [Delay]			[12-Jun-...
[DNB/Event_Report_108]	[SNO]		[2060.80]	[Basic] [Delay]			[23-Jul-...
[BKK/Event_Report_2]	[SNO]	[31-Oct-21]	[2551.55, 2551.80, 2...	[Non-Basic] [Air Turnback]	[13699]	[5267]	[31-Oct-...
[DNB/Event_Report_195]	[SAB]	[15/10/23]	[2055.90]	[Basic] [Delay]			[15-Oct-...
[DNB/Event_Report_122]	[NBO]	[5/08/23]	[2061.20]	[Basic] [Delay]			[05-Aug-...
[DNB/Event_Report_169]	[NBO]	[29/09/23]	[2061.85]	[Non-Basic] [Delay]			[29-Sep-...
[DNB/Event_Report_121]	[NBO]	[05/08/23]	[2058.00, 2]	[Basic] [Delay]			[05-Aug-...
[DNB/Event_Report_118]	[NBO]		[2060.30]	[Basic] [Delay]			[03-Aug-...
[DNB/Event_Report_143]	[NBO]		[2053.90]	[Basic] [Delay]			[23-Aug-...
[DNB/Event_Report_71]	[MOP]	[11/06/23]	[2360.35]	[Non-Basic] [ATCR high speed]...			[11-Jun-...
[DNB/Event_Report_151]	[MUC]	[09/09/23]	[2054.95]	[Basic] [Delay]			[09-Sep-...
[DNB/Event_Report_125]	[MEL]	[31/7/23]	[2368.90]	[Open] [Cancellation]			[31-Jul-...
[DNB/Event_Report_135]	[MEL]		[2055.95]	[Non-Basic] [Delay]			[28-Jul-...
[DNB/Event_Report_194]	[MCT]		[2058.25]	[Basic] [Delay]			[28-Oct-...
[AHI/Event_Report_1]	[MAN]	[24/05/22]	[2057.40]	[Non-Basic] [Delay]		[0]	[24-May-...
[DNB/Event_Report_85]	[MAN]	[08/06/23]	[2062.50]	[Basic] [Delay]			[08-Jun-...
[DNB/Event_Report_81]	[MAN]	[26/06/23]	[2056.75]	[Basic] [Delay]			[26-Jun-...
[DNB/Event_Report_200]	[LHR]	[01/11/23]	[2571.45]	[Non-Basic] [Delay]			[01-Nov-...
[DNB/Event_Report_75]	[LHR]	[14/06/23]	[2058.45]	[Basic] [Delay]			[14-Jun-...
[DNB/Event_Report_58]	[LGW]	[28/05/23]	[2058.00]	[Basic] [Delay]			[28-May-...
[DNB/Event_Report_49]	[KLA]	[19/05/23]	[2552.55]	[Basic] [Delay]			[19-May-...
[DNB/Event_Report_170]	[KHE]		[2061.85]	[Basic] [Delay]			[01-Oct-...
[DNB/Event_Report_161]	[IST]	[23/09/23]	[2062.80]	[Basic] [Diversion/Delay]			[23-Sep-...
[BKK/Event_Report_61]	[ICN]	[24-Nov-21]	[2555.70]		[10917]	[2805]	[24-Nov-...
[BKK/Event_Report_157]	[ICN]	[18-Apr-22]	[2553.65]		[12477]	[4572]	[18-Apr-...
[BKK/Event_Report_225]	[HKG]	[07-Jul-22]			[7611]	[1893]	[07-Jul-...
[BKK/Event_Report_250]	[HKG]	[17-Sep-22]	[2550.55]		[16608]	[6427]	[17-Sep-...
[BKK/Event_Report_165]	[HKG]	[26-Nov-21]	[2550.90]		[11478]	[6380]	[26-Nov-...

Timeline





```
<rdf:Description r
RR:Loc#Event_R
<rdf:type rdf:resourc
<rdf:has_#loc_location
<rdf:has_#form_activit
<rdf:has_#event_date_2
<rdf:has_#assoc_data
<rdf:has_#tsn_14613-
<rdf:has_#engin_sch
<rdf:has_#loc_instr_3IK
<rdf:has_#regime_#GR
<rdf:has_#cnn_3362-
<rdf:has_#compone-
<rdf:Description
```



RDF triples



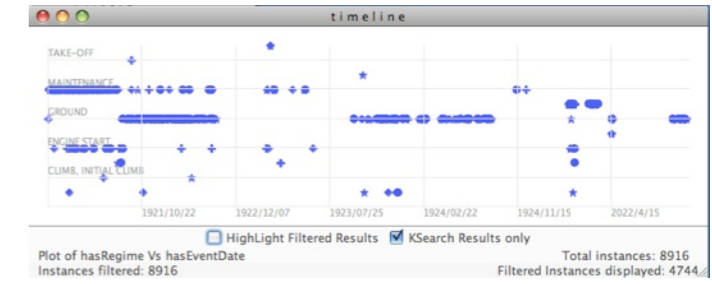
The University Of Sheffield.

```
<rdf:Description rdfs:about="http://kmi.open.ac.uk/projects/amedia/RR1/low/Event_Report_BKK_Event_Report_237">
<rdf:type rdf:resource="http://kmi.open.ac.uk/projects/amedia/RR1/low/Event_Report"/>
<ij:has_file_location-BKK/Event_Report_237/ij:has_file_location>
<ij:hasFormattedEventDate>26-Jul-1922<ij:hasFormattedEventDate>
<ij:hasEventDate>26-Jul-22<ij:hasEventDate>
<ij:hasAssociatedDate>28-Aug-22<ij:hasAssociatedDate>
<ij:hasTSN>14613<ij:hasTSN>
<ij:hasEngine_Serial_Number>2551.55<ij:hasEngine_Serial_Number>
<ij:hasLocation-BKK<ij:hasLocation>
<ij:hasRegime-GROUND<ij:hasRegime>
<ij:hasCSN>5362<ij:hasCSN>
<ij:hasComponent-Fuel<ij:hasComponent>
<ij:hasDescription>
```

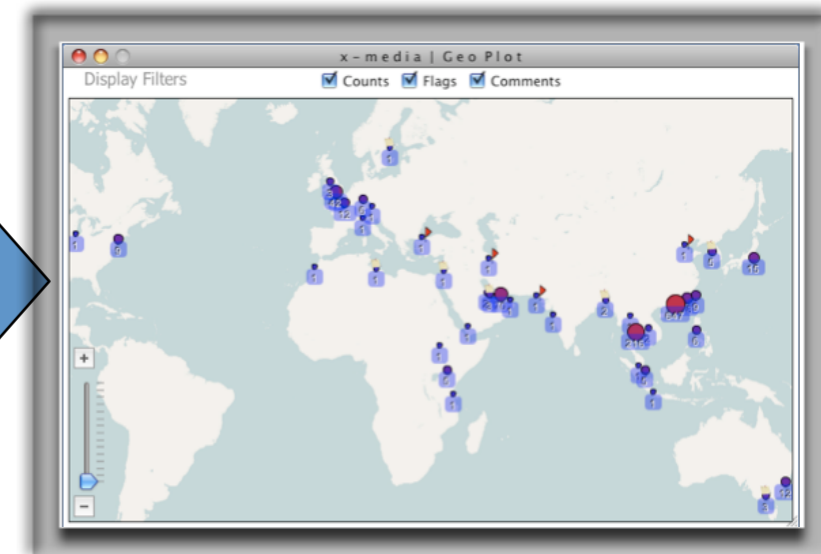
RDF table

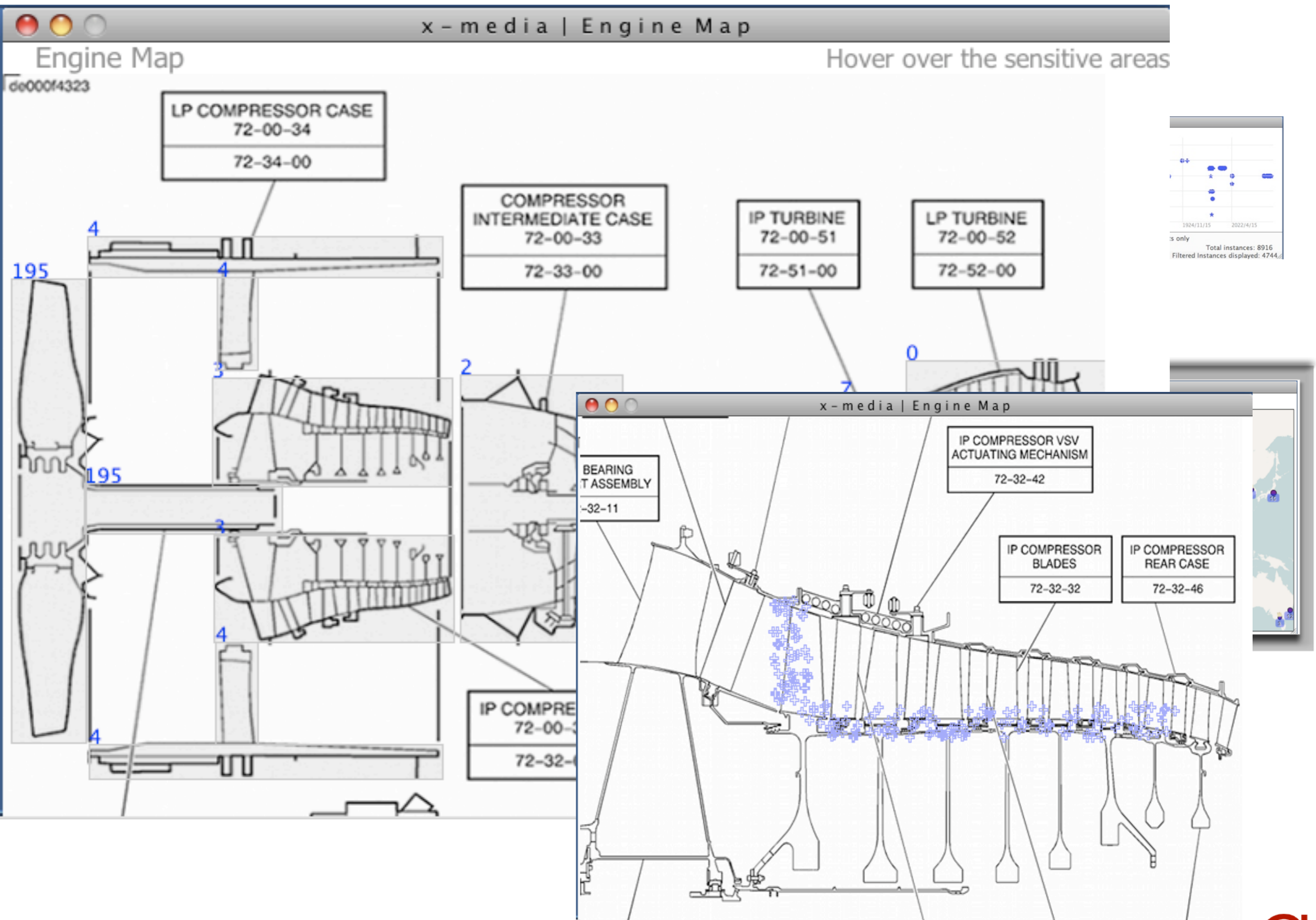
has_file_location	hasLocation	hasEventDate	hasEngine_Serial_Number	hasCategory:hasType	hasTSN	hasCSN	hasFormat...
[BKK/Event_Report_174]	[BKP]	[18-May-22]	[2553.25, 2553.05]	[Basic]	[12136]	[4193]	[18-May-...
[BKK/Event_Report_194]	[BKP]	[28-May-22]	[2553.05]	[Basic]	[11685]	[4482]	[28-May-...
[BKK/Event_Report_29]	[TPE]	[26-Aug-21]	[2553.95, 2550.05]	[Basic]	[8365]	[2878]	[26-Aug-...
[BKK/Event_Report_44]	[TPE]	[15/05/21]	[2058.65]	[Non-Basic]	[Delay]		[15-May-...
[DNB/Event_Report_78]	[THR]	[18/06/21]	[2053.65]	[Non-Basic]	[Delay]		[18-Jun-...
[DNB/Event_Report_149]	[SYD]	[07/09/21]	[2569.35]	[Non-Basic]	[Cancellation]		[07-Sep-...
[DNB/Event_Report_8]	[SYD]	[18/01/24]	[3557.35/3554.00]	[Basic]	[Delay]		[18-Jan-...
[DNB/Event_Report_74]	[SNG]	[12/06/21]	[2552.55]	[Basic]	[Delay]		[12-Jun-...
[DNB/Event_Report_108]	[SNG]		[2060.80]	[Basic]	[Delay]		[23-Jul-...
[BKK/Event_Report_2]	[SNG]	[31-Oct-21]	[2551.55, 2551.80, 2...	[Non-Basic]	[Air Turnback]	[13699]	[5267]
[DNB/Event_Report_195]	[SAB]	[15/10/21]	[2055.90]	[Basic]	[Delay]		[15-Oct-...
[DNB/Event_Report_122]	[NBO]	[5/08/21]	[2061.20]	[Basic]	[Delay]		[05-Aug-...
[DNB/Event_Report_169]	[NBO]	[29/09/21]	[2061.85]	[Non-Basic]	[Delay]		[29-Sep-...
[DNB/Event_Report_121]	[NBO]	[05/08/21]	[2058.00, 2]	[Basic]	[Delay]		[05-Aug-...
[DNB/Event_Report_118]	[NBO]		[2060.30]	[Basic]	[Delay]		[03-Aug-...
[DNB/Event_Report_143]	[NBO]		[2053.90]	[Basic]	[Delay]		[23-Aug-...
[DNB/Event_Report_71]	[NBP]	[11/06/21]	[2360.35]	[Non-Basic]	[AETC high speed 3/...		[11-Jun-...
[DNB/Event_Report_151]	[MUC]	[09/09/21]	[2054.95]	[Basic]	[Delay]		[09-Sep-...
[DNB/Event_Report_125]	[MEL]	[31/7/23]	[2368.90]	[Open]	[Cancellation]		[31-Jul-...
[DNB/Event_Report_135]	[MEL]		[2055.95]	[Non-Basic]	[Delay]		[28-Jul-...
[DNB/Event_Report_194]	[MCT]		[2058.25]	[Basic]	[Delay]		[28-Oct-...
[AFL/Event_Report_1]	[MAN]	[24/05/22]	[2057.40]	[Non-Basic]	[Delay]	[0]	[24-May-...
[DNB/Event_Report_85]	[MAN]	[08/06/21]	[2062.50]	[Basic]	[Delay]		[08-Jun-...
[DNB/Event_Report_81]	[MAN]	[26/06/21]	[2056.75]	[Basic]	[Delay]		[26-Jun-...
[DNB/Event_Report_200]	[LHR]	[01/11/21]	[2571.45]	[Non-Basic]	[Delay]		[01-Nov-...
[DNB/Event_Report_75]	[LHR]	[14/06/21]	[2059.45]	[Basic]	[Delay]		[14-Jun-...
[DNB/Event_Report_58]	[LGW]	[28/05/21]	[2058.00]	[Basic]	[Delay]		[28-May-...
[DNB/Event_Report_49]	[KLA]	[19/05/21]	[2552.55]	[Basic]	[Delay]		[19-May-...
[DNB/Event_Report_170]	[KHE]		[2061.85]	[Basic]	[Delay]		[01-Oct-...
[DNB/Event_Report_161]	[IST]	[23/09/21]	[2062.80]	[Basic]	[Diversion/Delay]		[23-Sep-...
[BKK/Event_Report_61]	[ICN]	[24-Nov-21]	[2555.70]		[10917]	[2805]	[24-Nov-...
[BKK/Event_Report_157]	[ICN]	[18-Apr-22]	[2553.65]		[12477]	[4572]	[18-Apr-...
[BKK/Event_Report_225]	[HKG]	[07-Jul-22]	[2550.55]		[7611]	[1893]	[07-Jul-...
[BKK/Event_Report_250]	[HKG]	[17-Sep-22]	[2550.55]		[16608]	[6427]	[17-Sep-...
[BKK/Event_Report_165]	[HKG]	[26-Nov-21]	[2550.90]		[13478]	[6380]	[26-Nov-...

Timeline



GeoPlot





Number of levels depends on ontology granularity

RDF triples



The University of Sheffield.

```
<rdf:Description rdfs:about="http://kmi.open.ac.uk/projects/amedia/RR1LowIfEvent_Report_BKK_Event_Report_237">
  <rdf:type rdf:resource="http://kmi.open.ac.uk/projects/amedia/RR1LowIfEvent_Report"/>
  <?has_file_location=BKK/Event_Report_237?>
  <?hasFormattedEventDate=26-Jul-1922?>
  <?hasEventDate=26-Jul-22?>
  <?hasAssociatedDate=28-Aug-22?>
  <?hasTSN=14613?>
  <?hasEngine_Serial_Number=2551.55?>
  <?hasLocation=BKK?>
  <?hasRegime=GROUND?>
  <?hasCSN=5362?>
  <?hasComponent=Fuel?>
  <?hasDescription?>
```

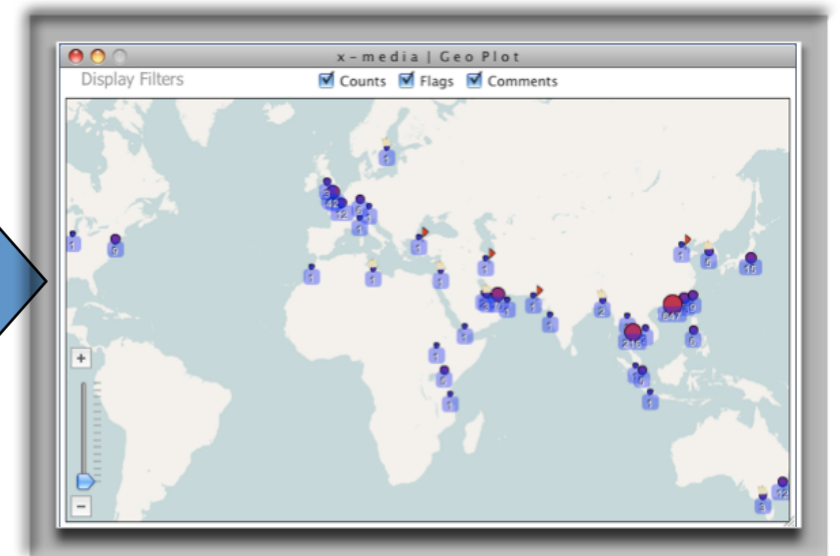
RDF table

has_file_location	hasLocation	hasEventDate	hasEngine_Serial_Number	hasCategory:hasType	hasTSN	hasCSN	hasFormat...
[BKK/Event_Report_174]	[BKP]	[18-May-22]	[2553.25, 2553.05]	[Basic] [Delay]	[12136]	[4193]	[18-May-...
[BKK/Event_Report_194]	[BKP]	[28-May-22]	[2553.05]	[Basic] [Delay]	[11685]	[6482]	[28-May-...
[BKK/Event_Report_29]	[TPE]	[26-Aug-21]	[2553.95, 2550.05]	[Basic] [Delay]	[8365]	[2878]	[26-Aug-...
[DNB/Event_Report_44]	[TPE]	[15/05/23]	[2058.65]	[Non-Basic] [Delay]			[15-May-...
[DNB/Event_Report_78]	[TPE]	[16/06/23]	[2058.65]	[Non-Basic] [Cancellation]			[16-Jun-...
[DNB/Event_Report_149]	[SVD]	[07/09/23]	[2569.35]	[Basic] [Delay]			[07-Sep-...
[DNB/Event_Report_8]	[SVD]	[18/01/24]	[3557.35/3554.00]	[Basic] [Delay]			[18-Jan-...
[DNB/Event_Report_74]	[SN]	[12/06/23]	[2055.55]	[Basic] [Delay]			[12-Jun-...
[DNB/Event_Report_108]	[SN]		[2060.80]	[Basic] [Delay]			[23-Jul-...
[BKK/Event_Report_2]	[SN]	[31-Oct-21]	[2551.55, 2551.80, 2...	[Non-Basic] [Air Turbback]	[13699]	[5267]	[31-Oct-...
[DNB/Event_Report_195]	[SAB]	[15/10/23]	[2055.90]	[Basic] [Delay]			[15-Oct-...
[DNB/Event_Report_122]	[NBO]	[5/08/23]	[2061.20]	[Basic] [Delay]			[05-Aug-...
[DNB/Event_Report_169]	[NBO]	[29/09/23]	[2061.85]	[Basic] [Delay]			[29-Sep-...
[DNB/Event_Report_121]	[NBO]	[05/08/23]	[2058.00, 2]	[Basic] [Delay]			[05-Aug-...
[DNB/Event_Report_118]	[NBO]		[2060.30]	[Basic] [Delay]			[03-Aug-...
[DNB/Event_Report_143]	[NBO]		[2053.90]	[Basic] [Delay]			[23-Aug-...
[DNB/Event_Report_71]	[NBP]	[11/06/23]	[2360.35]	[Non-Basic] [AETCO high speed]...			[11-Jun-...
[DNB/Event_Report_151]	[MUC]	[09/09/23]	[2054.95]	[Basic] [Delay]			[09-Sep-...
[DNB/Event_Report_125]	[MEL]	[31/7/23]	[2368.90]	[Open] [Cancellation]			[31-Jul-...
[DNB/Event_Report_135]	[MEL]		[2055.95]	[Non-Basic] [Delay]			[28-Jul-...
[DNB/Event_Report_194]	[MCT]		[2058.25]	[Basic] [Delay]			[28-Oct-...
[APL/Event_Report_1]	[MAN]	[24/05/22]	[2057.40]	[Non-Basic] [Delay]			[24-May-...
[DNB/Event_Report_85]	[MAN]	[08/06/23]	[2062.50]	[Basic] [Delay]			[08-Jun-...
[DNB/Event_Report_81]	[MAN]	[26/06/23]	[2056.75]	[Basic] [Delay]			[26-Jun-...
[DNB/Event_Report_200]	[LHR]	[01/11/23]	[2571.45]	[Non-Basic] [Delay]			[01-Nov-...
[DNB/Event_Report_75]	[LHR]	[14/06/23]	[2059.45]	[Basic] [Delay]			[14-Jun-...
[DNB/Event_Report_58]	[LGW]	[28/05/23]	[2058.00]	[Basic] [Delay]			[28-May-...
[DNB/Event_Report_49]	[KLA]	[19/05/23]	[2552.55]	[Basic] [Delay]			[19-May-...
[DNB/Event_Report_170]	[KHE]		[2061.85]	[Basic] [Delay]			[01-Oct-...
[DNB/Event_Report_161]	[IST]	[23/09/23]	[2062.80]	[Basic] [Diversion/Delay]			[23-Sep-...
[DNB/Event_Report_61]	[ICN]	[24-Nov-21]	[2555.70]		[10917]	[2805]	[24-Nov-...
[BKK/Event_Report_157]	[ICN]	[18-Apr-22]	[2553.65]		[12477]	[4572]	[18-Apr-...
[BKK/Event_Report_225]	[HKG]	[07-Jul-22]			[7613]	[1893]	[07-Jul-...
[BKK/Event_Report_250]	[HKG]	[17-Sep-22]	[2550.55]		[16608]	[6427]	[17-Sep-...
[BKK/Event_Report_165]	[HKG]	[26-Nov-21]	[2550.90]		[11478]	[6380]	[26-Nov-...

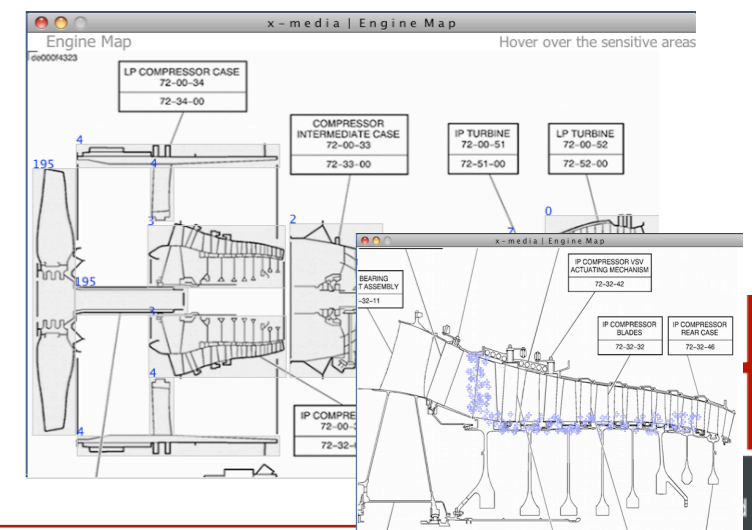
Timeline



GeoPlot



TopologicalMap



RDF triples



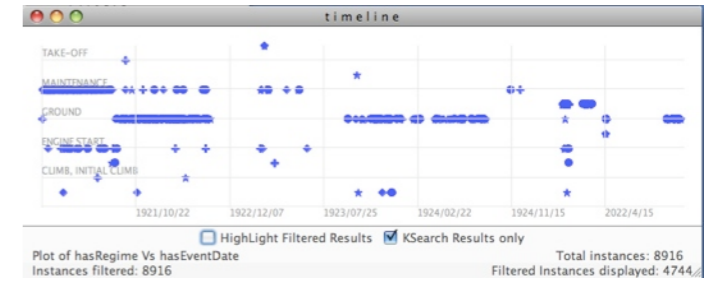
The University Of Sheffield.

```
<rdf:Description rdf:about="http://kmi.open.ac.uk/projects/amedia/RR1Low/Event_Report_BKK_Event_Report_237">
  <rdf:type rdf:resource="http://kmi.open.ac.uk/projects/amedia/RR1Low/Event_Report"/>
  <j0has_file_location-BKK/Event_Report_237/>
  <j0hasFormatofEventDate=26-Jul-1922/>
  <j0hasEventDate=26-Jul-22/>
  <j0hasAssociatedDate=28-Aug-22/>
  <j0hasTSN=14613/>
  <j0hasEngine_Serial_Number=255155/>
  <j0hasLocation-BKK/>
  <j0hasRegime-GROUND/>
  <j0hasCSN=5362/>
  <rdf:Description>
```

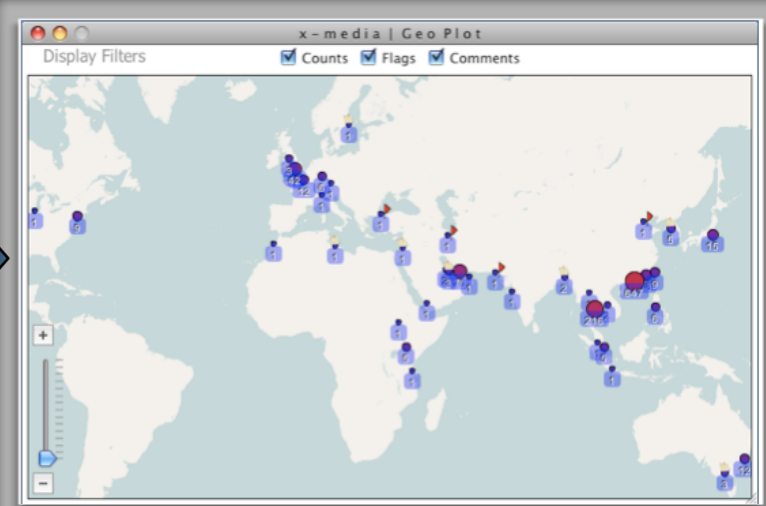
RDF table

hasFileLocation	hasLocation	hasEventDate	hasEngine_Serial_Number	hasCategory:hasType	hasTSN	hasCSN	hasFormatofEventDate
[BKK/Event_Report_174]	[BKP]	[18-May-22]	[2553.25, 2553.05]	[Basic] [Delay]	[12136]	[4193]	[18-May-22]
[BKK/Event_Report_194]	[BKP]	[28-May-22]	[2553.05]	[Basic] [Delay]	[11685]	[4482]	[28-May-22]
[BKK/Event_Report_29]	[TPE]	[26-Aug-21]	[2553.95, 2550.05]	[Basic] [Delay]	[8365]	[2878]	[26-Aug-21]
[DNB/Event_Report_44]	[TPE]	[15/05/23]	[2058.65]	[Non-Basic] [Delay]			[15-May-23]
[DNB/Event_Report_78]	[TPE]	[16/06/23]	[2058.65]	[Non-Basic] [Cancellation]			[16-Jun-23]
[DNB/Event_Report_149]	[SYD]	[07/09/23]	[2569.35]	[Basic] [Delay]			[07-Sep-23]
[DNB/Event_Report_6]	[SYD]	[18/01/24]	[3557.35,3554.00]	[Basic] [Delay]			[18-Jan-24]
[DNB/Event_Report_74]	[SNO]	[12/06/23]	[2552.55]	[Basic] [Delay]			[12-Jun-23]
[DNB/Event_Report_108]	[SNO]	[31-Oct-21]	[2551.55, 2551.80, 2...	[Non-Basic] [Air Turback]	[13699]	[5267]	[31-Oct-21]
[DNB/Event_Report_2]	[SNO]	[15/10/23]	[2055.90]	[Basic] [Delay]			[15-Oct-23]
[DNB/Event_Report_195]	[SNO]	[05/08/23]	[2061.20]	[Basic] [Delay]			[05-Aug-23]
[DNB/Event_Report_122]	[SNO]	[05/08/23]	[2061.20]	[Basic] [Delay]			[05-Aug-23]
[DNB/Event_Report_169]	[SNO]	[29/09/23]	[2061.85]	[Basic] [Delay]			[29-Sep-23]
[DNB/Event_Report_121]	[SNO]	[05/08/23]	[2058.00, 2]	[Basic] [Delay]			[05-Aug-23]
[DNB/Event_Report_118]	[SNO]	[05/08/23]	[2060.30]	[Basic] [Delay]			[05-Aug-23]
[DNB/Event_Report_143]	[SNO]	[05/08/23]	[2053.90]	[Basic] [Delay]			[05-Aug-23]
[DNB/Event_Report_71]	[MOP]	[11/06/23]	[2360.35]	[Non-Basic] [AETCO high speed]...			[11-Jun-23]
[DNB/Event_Report_151]	[MUC]	[09/09/23]	[2054.95]	[Basic] [Delay]			[09-Sep-23]
[DNB/Event_Report_125]	[MEL]	[31/7/23]	[2368.90]	[Open] [Cancellation]			[31-Jul-23]
[DNB/Event_Report_135]	[MEL]	[31/7/23]	[2055.95]	[Non-Basic] [Delay]			[28-Jul-23]
[DNB/Event_Report_194]	[MCT]	[24/05/22]	[2057.40]	[Basic] [Delay]			[24-May-22]
[DNB/Event_Report_1]	[MAN]	[08/06/23]	[2062.50]	[Non-Basic] [Delay]			[08-Jun-23]
[DNB/Event_Report_81]	[MAN]	[26/06/23]	[2056.75]	[Basic] [Delay]			[26-Jun-23]
[DNB/Event_Report_209]	[LHR]	[01/11/23]	[2571.45]	[Non-Basic] [Delay]			[01-Nov-23]
[DNB/Event_Report_75]	[LHR]	[14/06/23]	[2059.45]	[Basic] [Delay]			[14-Jun-23]
[DNB/Event_Report_58]	[LGC]	[28/05/23]	[2058.00]	[Basic] [Delay]			[28-May-23]
[DNB/Event_Report_49]	[KLA]	[19/05/23]	[2552.55]	[Basic] [Delay]			[19-May-23]
[DNB/Event_Report_170]	[KHE]	[09/08/23]	[2061.85]	[Basic] [Delay]			[09-Aug-23]
[DNB/Event_Report_161]	[IST]	[23/09/23]	[2062.80]	[Basic] [Diversion/Delay]			[23-Sep-23]
[DNB/Event_Report_61]	[ICN]	[24-Nov-21]	[2555.70]	[Basic] [Delay]	[10917]	[2805]	[24-Nov-21]
[BKK/Event_Report_153]	[ICN]	[18-Apr-22]	[2553.65]	[Basic] [Delay]	[12477]	[4572]	[18-Apr-22]
[BKK/Event_Report_225]	[HKG]	[07-Jul-22]		[Basic] [Delay]	[7613]	[1893]	[07-Jul-22]
[BKK/Event_Report_250]	[HKG]	[17-Sep-22]	[2550.55]	[Basic] [Delay]	[16608]	[6427]	[17-Sep-22]
[BKK/Event_Report_165]	[HKG]	[26-Nov-21]	[2550.90]	[Basic] [Delay]	[11478]	[6380]	[26-Nov-21]

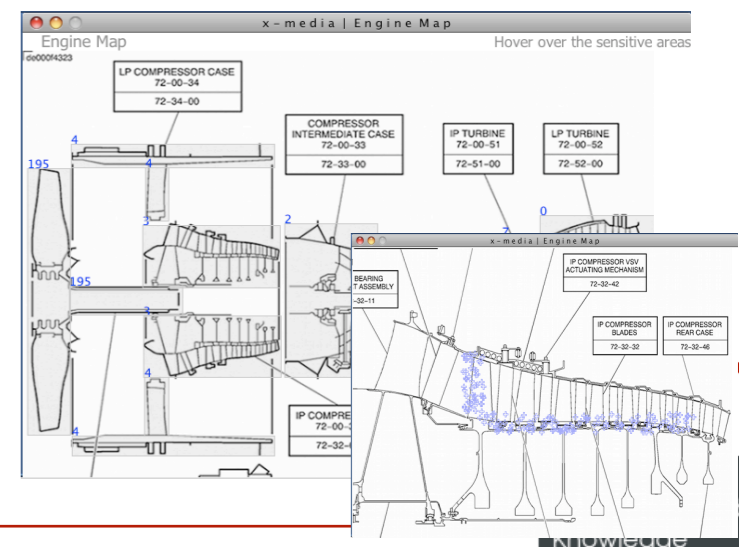
Timeline



GeoPlot



TopologicalMap



- slide: numeric range
- list: small sets (<25)
- text: large sets

Filters

Reset
K-Search
Results

Slider Filters

CSN ✕

5725

7647

TSN ✕

List Filters

Category ✕

BASIC
 OPEN
 NON-BASIC

Type ✕

Regime ✕

Text Filters

AssociatedDate

Probable Cause

Location

RDF triples



The University of Sheffield.

```
<rdf:Description rdfs:about="http://kmi.open.ac.uk/projects/amedia/RR1/low/Event_Report_BKK/Event_Report_237">
  <rdf:type rdf:resource="http://kmi.open.ac.uk/projects/amedia/RR1/low/Event_Report"/>
  <rdf:has_file_location-BKK/Event_Report_237/>
  <rdf:hasFormatofEventDate=26-Jul-1922/>
  <rdf:hasEventDate=26-Jul-22/>
  <rdf:hasAssociatedDate=28-Aug-22/>
  <rdf:hasTSN=14613/>
  <rdf:hasEngine_Serial_Number=2551.55/>
  <rdf:hasLocation-BKK/>
  <rdf:hasRegime-GROUND/>
  <rdf:hasCSN=5362/>
  <rdf:hasComponent-Fuel/>
  <rdf:hasDescription/>
```

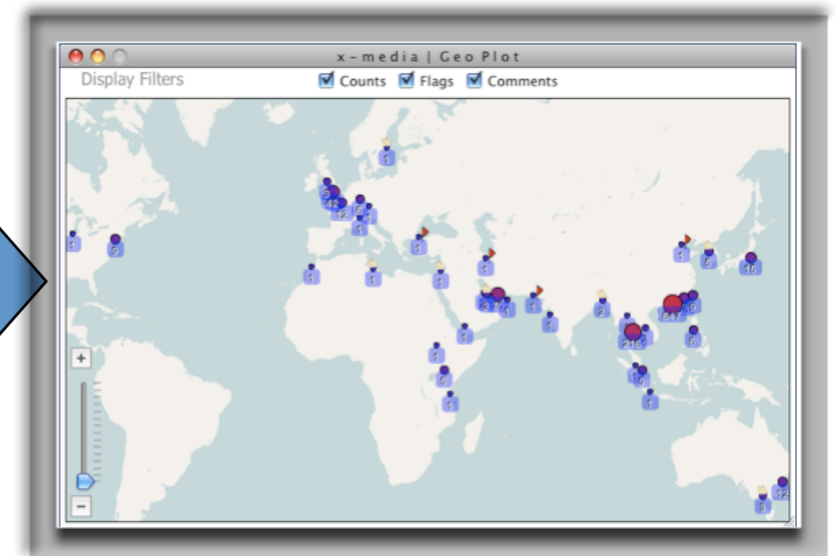
RDF table

has_file_location	hasLocation	hasEventDate	hasEngine_Serial_Number	hasCategory:hasType	hasTSN	hasCSN	hasFormat...
[BKK/Event_Report_174]	[BKP]	[18-May-22]	[2553.25, 2553.05]	[Basic] [Delay]	[12136]	[4193]	[18-May-...
[BKK/Event_Report_194]	[BKP]	[28-May-22]	[2553.05]	[Basic] [Delay]	[11685]	[6482]	[28-May-...
[BKK/Event_Report_29]	[TPE]	[26-Aug-21]	[2553.95, 2550.05]	[Basic] [Delay]	[8365]	[2878]	[26-Aug-...
[BKK/Event_Report_44]	[TPE]	[15/05/21]	[2058.65]	[Non-Basic] [Delay]			[15-May-...
[BKK/Event_Report_78]	[TWE]	[16/06/21]	[2058.65]	[Non-Basic] [Delay]			[16-Jun-...
[BKK/Event_Report_149]	[SVD]	[07/09/21]	[2569.35]	[Non-Basic] [Cancellation]			[07-Sep-...
[BKK/Event_Report_8]	[SVD]	[18/01/24]	[3557.35/3554.00]	[Basic] [Delay]			[18-Jan-...
[BKK/Event_Report_74]	[SN]	[12/06/21]	[2552.55]	[Basic] [Delay]			[12-Jun-...
[BKK/Event_Report_108]	[SN]		[2060.80]	[Basic] [Delay]			[23-Jul-...
[BKK/Event_Report_2]	[SN]	[31-Oct-21]	[2551.55, 2551.80, 2...	[Non-Basic] [Air Turback]	[13699]	[5267]	[31-Oct-...
[BKK/Event_Report_195]	[SAB]	[15/10/21]	[2055.90]	[Non-Basic] [Delay]			[15-Oct-...
[BKK/Event_Report_122]	[NBO]	[5/08/21]	[2061.20]	[Basic] [Delay]			[05-Aug-...
[BKK/Event_Report_169]	[NBO]	[29/09/21]	[2061.85]	[Non-Basic] [Delay]			[29-Sep-...
[BKK/Event_Report_121]	[NBO]	[05/08/21]	[2058.00, 2]	[Basic] [Delay]			[05-Aug-...
[BKK/Event_Report_118]	[NBO]		[2060.30]	[Basic] [Delay]			[03-Aug-...
[BKK/Event_Report_143]	[NBO]		[2053.90]	[Basic] [Delay]			[23-Aug-...
[BKK/Event_Report_71]	[NBP]	[11/06/21]	[2360.35]	[Non-Basic] [AETCO high speed]...			[11-Jun-...
[BKK/Event_Report_151]	[MUC]	[09/09/21]	[2054.95]	[Basic] [Delay]			[09-Sep-...
[BKK/Event_Report_125]	[MEL]	[31/7/23]	[2368.90]	[Open] [Cancellation]			[31-Jul-...
[BKK/Event_Report_135]	[MEL]		[2055.95]	[Non-Basic] [Delay]			[28-Jun-...
[BKK/Event_Report_194]	[MCT]		[2058.25]	[Basic] [Delay]			[26-Oct-...
[BKK/Event_Report_1]	[MAN]	[24/05/22]	[2057.40]	[Non-Basic] [Delay]			[24-May-...
[BKK/Event_Report_85]	[MAN]	[08/06/21]	[2062.50]	[Non-Basic] [Delay]			[08-Jun-...
[BKK/Event_Report_81]	[MAN]	[26/06/21]	[2056.75]	[Basic] [Delay]			[26-Jun-...
[BKK/Event_Report_200]	[LHR]	[01/11/21]	[2571.45]	[Non-Basic] [Delay]			[01-Nov-...
[BKK/Event_Report_75]	[LHR]	[14/06/21]	[2059.45]	[Basic] [Delay]			[14-Jun-...
[BKK/Event_Report_58]	[LGW]	[28/05/21]	[2058.00]	[Basic] [Delay]			[28-May-...
[BKK/Event_Report_49]	[KLA]	[19/05/21]	[2552.55]	[Basic] [Delay]			[19-May-...
[BKK/Event_Report_170]	[KHE]		[2061.85]	[Basic] [Delay]			[01-Oct-...
[BKK/Event_Report_161]	[IST]	[23/09/21]	[2062.80]	[Basic] [Diversion/Delay]			[23-Sep-...
[BKK/Event_Report_61]	[ICN]	[24-Nov-21]	[2555.70]		[10917]	[2805]	[24-Nov-...
[BKK/Event_Report_157]	[ICN]	[18-Apr-22]	[2553.65]		[12477]	[4572]	[18-Apr-...
[BKK/Event_Report_225]	[HKG]	[07-Jul-22]			[7611]	[1893]	[07-Jul-...
[BKK/Event_Report_250]	[HKG]	[17-Sep-22]	[2550.55]		[16608]	[6427]	[17-Sep-...
[BKK/Event_Report_165]	[HKG]	[26-Nov-21]	[2550.90]		[11478]	[6380]	[26-Nov-...

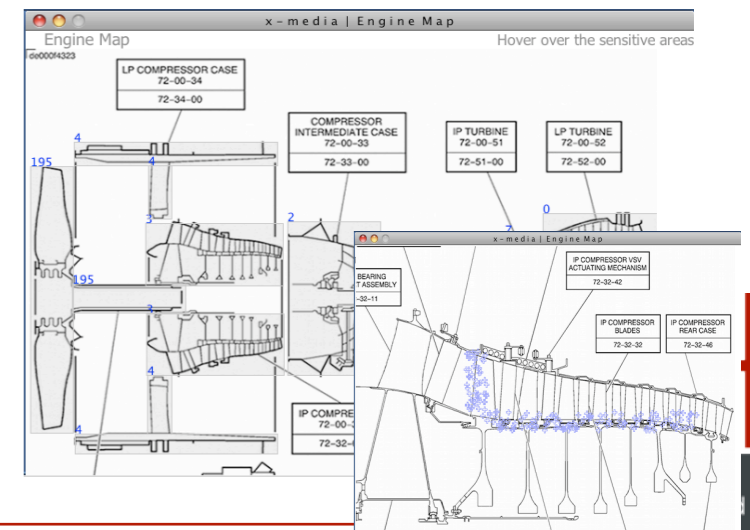
Timeline



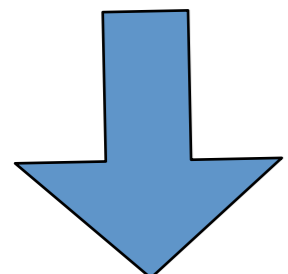
GeoPlot



TopologicalMap



Ontology



Filters

Reset K-Search Results

Slider Filters

CSN: 5725 - 7647

TSN

List Filters

Category

- BASIC
- OPEN
- NON-BASIC

Type

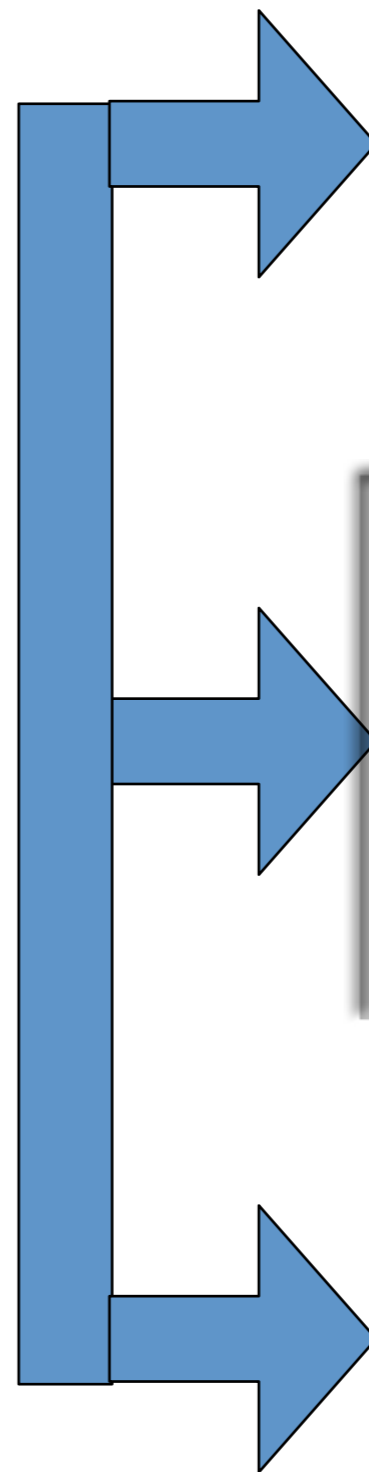
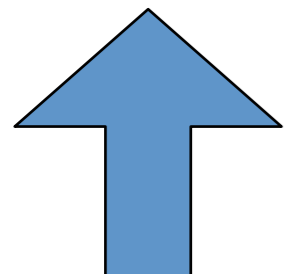
Regime

Text Filters

AssociatedDate

Probable Cause

Location



RDF triples



The University of Sheffield.

```
<rdf:Description rdfs:about="http://kmi.open.ac.uk/projects/amedia/RR1LowIfEvent_Report_BKK_Event_Report_237">
  <rdf:type rdf:resource="http://kmi.open.ac.uk/projects/amedia/RR1LowIfEvent_Report"/>
  <rdf:has_file_location-BKK/Event_Report_237/>
  <rdf:hasFormatofEventDate=26-Jul-1922/>
  <rdf:hasEventDate=26-Jul-22/>
  <rdf:hasAssociatedDate=28-Aug-22/>
  <rdf:hasTSN=14613/>
  <rdf:hasEngine_Serial_Number=255155/>
  <rdf:hasLocation-BKK/>
  <rdf:hasRegime-GROUND/>
  <rdf:hasCSN=5362/>
  <rdf:hasComponent-Fuel/>
  <rdf:hasDescription/>
```

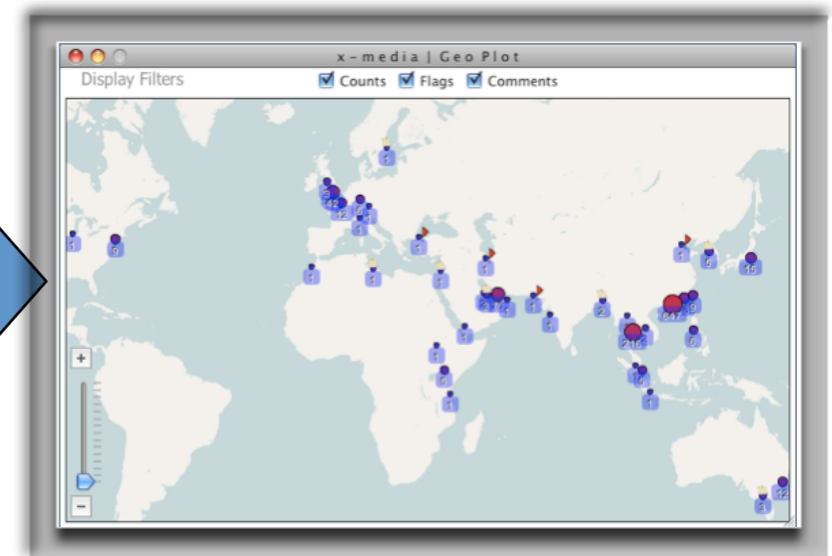
RDF table

has_file_location	hasLocation	hasEventDate	hasEngine_Serial_Number	hasCategory:hasType	hasTSN	hasCSN	hasFormat...
[BKK/Event_Report_174]	[BKP]	[18-May-22]	[2553.25, 2553.05]	[Basic] [Delay]	[12136]	[4193]	[18-May-...
[BKK/Event_Report_194]	[BKP]	[28-May-22]	[2553.05]	[Basic] [Delay]	[14685]	[6482]	[28-May-...
[BKK/Event_Report_29]	[TPE]	[26-Aug-21]	[2553.95, 2550.05]	[Basic] [Delay]	[8365]	[2878]	[26-Aug-...
[BKK/Event_Report_44]	[TPE]	[15/05/21]	[2058.65]	[Non-Basic] [Delay]			[15-May-...
[BKK/Event_Report_78]	[THRE]	[18/08/21]	[2058.65]	[Non-Basic] [Cancellation]			[18-Jan-...
[BKK/Event_Report_149]	[SYD]	[07/09/21]	[2569.35]	[Basic] [Delay]			[07-Sep-...
[BKK/Event_Report_8]	[SYD]	[18/01/24]	[3557.35/3554.00]	[Basic] [Delay]			[18-Jan-...
[BKK/Event_Report_74]	[SNG]	[12/08/21]	[2552.55]	[Basic] [Delay]			[12-Jan-...
[BKK/Event_Report_108]	[SNG]	[31-Oct-21]	[2551.55, 2551.80, 2...	[Basic] [Delay]	[13699]	[5267]	[23-Jul-...
[BKK/Event_Report_2]	[SNG]	[15/10/21]	[2055.90]	[Non-Basic] [Air Turnback]			[15-Oct-...
[BKK/Event_Report_195]	[SNG]	[15/10/21]	[2055.90]	[Basic] [Delay]			[15-Oct-...
[BKK/Event_Report_122]	[NBO]	[5/08/21]	[2061.20]	[Basic] [Delay]			[05-Aug-...
[BKK/Event_Report_169]	[NBO]	[29/09/21]	[2061.85]	[Basic] [Delay]			[29-Sep-...
[BKK/Event_Report_121]	[NBO]	[05/08/21]	[2058.00, 2]	[Basic] [Delay]			[05-Aug-...
[BKK/Event_Report_118]	[NBO]	[05/08/21]	[2060.30]	[Basic] [Delay]			[03-Aug-...
[BKK/Event_Report_143]	[NBO]	[05/08/21]	[2053.90]	[Basic] [Delay]			[23-Aug-...
[BKK/Event_Report_71]	[NBP]	[11/06/21]	[2360.35]	[Non-Basic] [AETCO high speed]...			[11-Jan-...
[BKK/Event_Report_151]	[MUC]	[09/09/21]	[2054.95]	[Basic] [Delay]			[09-Sep-...
[BKK/Event_Report_125]	[MEL]	[31/7/23]	[2368.90]	[Open] [Cancellation]			[31-Jul-...
[BKK/Event_Report_135]	[MEL]	[09/09/21]	[2055.95]	[Non-Basic] [Delay]			[28-Jul-...
[BKK/Event_Report_194]	[MCT]	[11/07/23]	[2058.25]	[Basic] [Delay]			[28-Oct-...
[BKK/Event_Report_1]	[MAN]	[24/05/22]	[2057.40]	[Non-Basic] [Delay]			[24-May-...
[BKK/Event_Report_85]	[MAN]	[08/08/21]	[2062.50]	[Basic] [Delay]			[08-Jun-...
[BKK/Event_Report_81]	[MAN]	[26/06/21]	[2056.75]	[Basic] [Delay]			[26-Jun-...
[BKK/Event_Report_200]	[LHR]	[01/11/21]	[2571.45]	[Non-Basic] [Delay]			[01-Nov-...
[BKK/Event_Report_75]	[LHR]	[14/08/21]	[2059.45]	[Basic] [Delay]			[14-Jan-...
[BKK/Event_Report_58]	[LGW]	[28/05/21]	[2058.00]	[Basic] [Delay]			[28-May-...
[BKK/Event_Report_49]	[KLA]	[19/05/21]	[2352.55]	[Basic] [Delay]			[19-May-...
[BKK/Event_Report_170]	[KHE]	[2061.85]		[Basic] [Delay]			[01-Oct-...
[BKK/Event_Report_161]	[IST]	[23/09/21]	[2062.80]	[Basic] [Diversion/Delay]			[23-Sep-...
[BKK/Event_Report_61]	[ICN]	[24-Nov-21]	[2555.70]		[10917]	[2805]	[24-Nov-...
[BKK/Event_Report_157]	[ICN]	[18-Apr-22]	[2553.65]		[12477]	[4572]	[18-Apr-...
[BKK/Event_Report_225]	[HKG]	[07-Jul-22]			[7611]	[1893]	[07-Jul-...
[BKK/Event_Report_250]	[HKG]	[17-Sep-22]	[2550.55]		[16608]	[6427]	[17-Sep-...
[BKK/Event_Report_165]	[HKG]	[26-Nov-21]	[2550.90]		[13478]	[6380]	[26-Nov-...

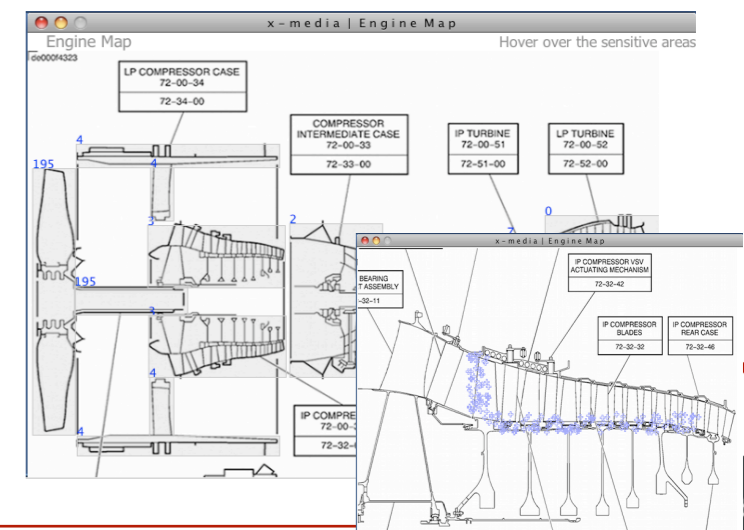
Timeline



GeoPlot

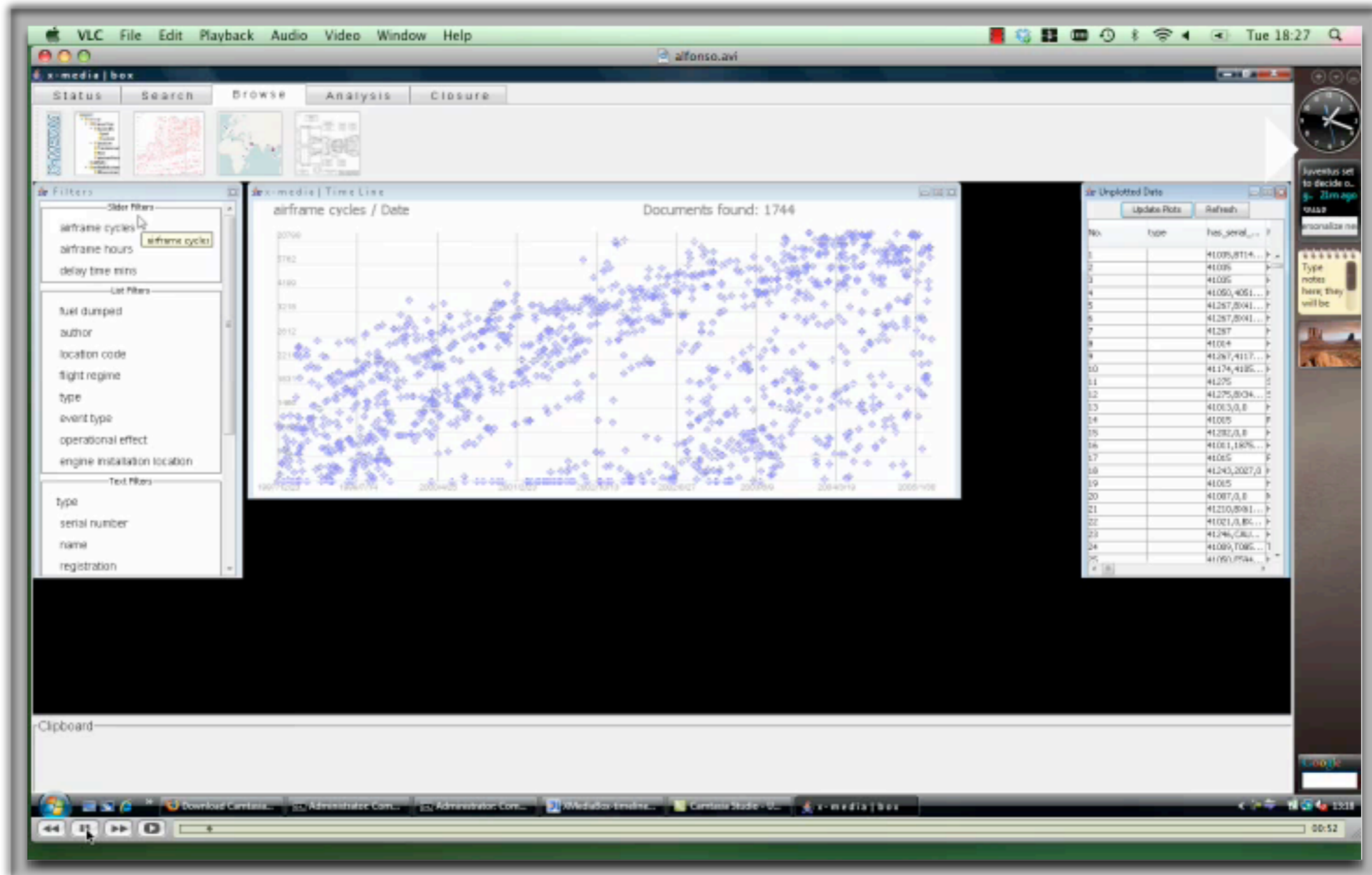


TopologicalMap



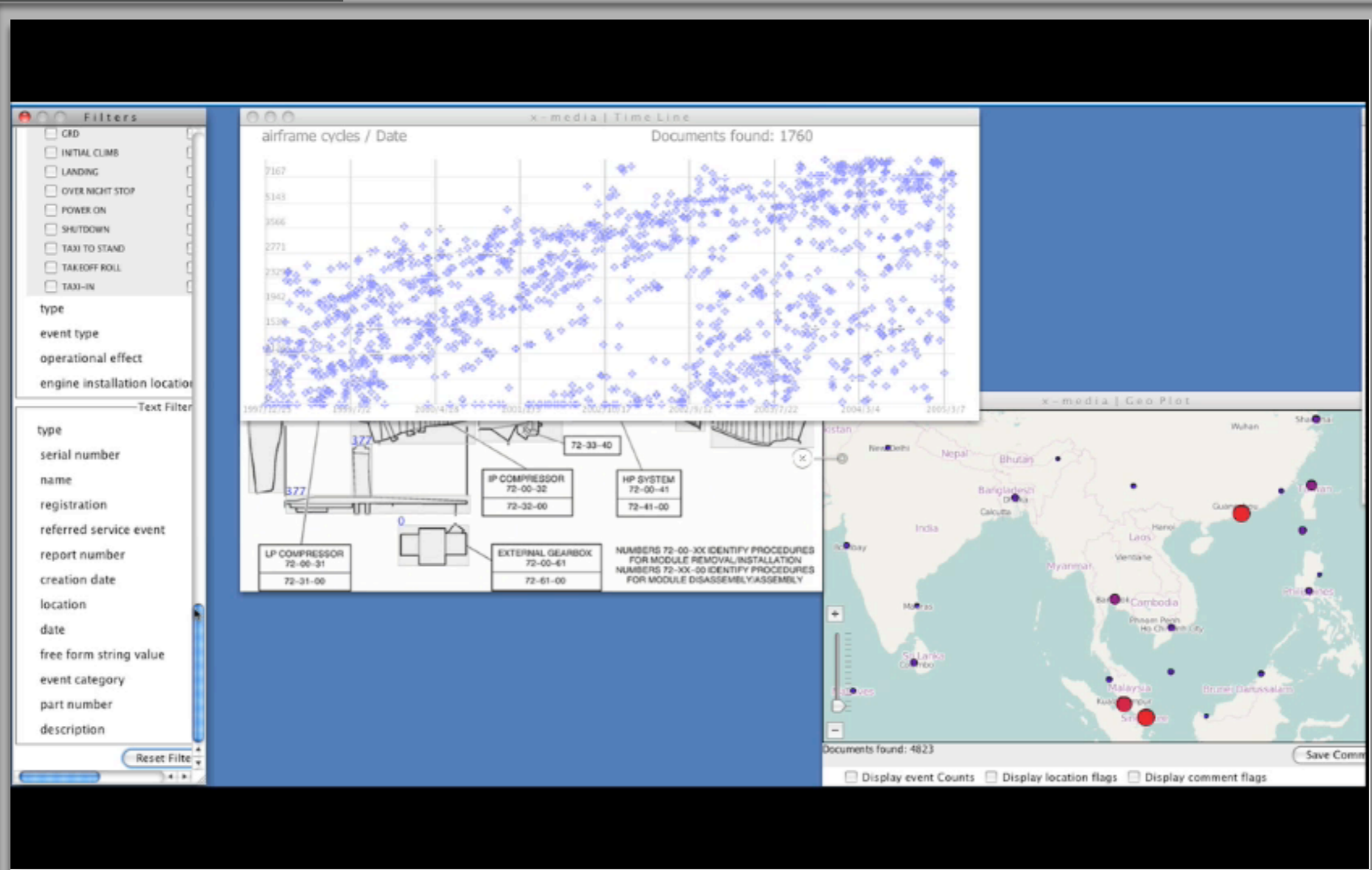
User

Ontology



Interaction with TimeLine





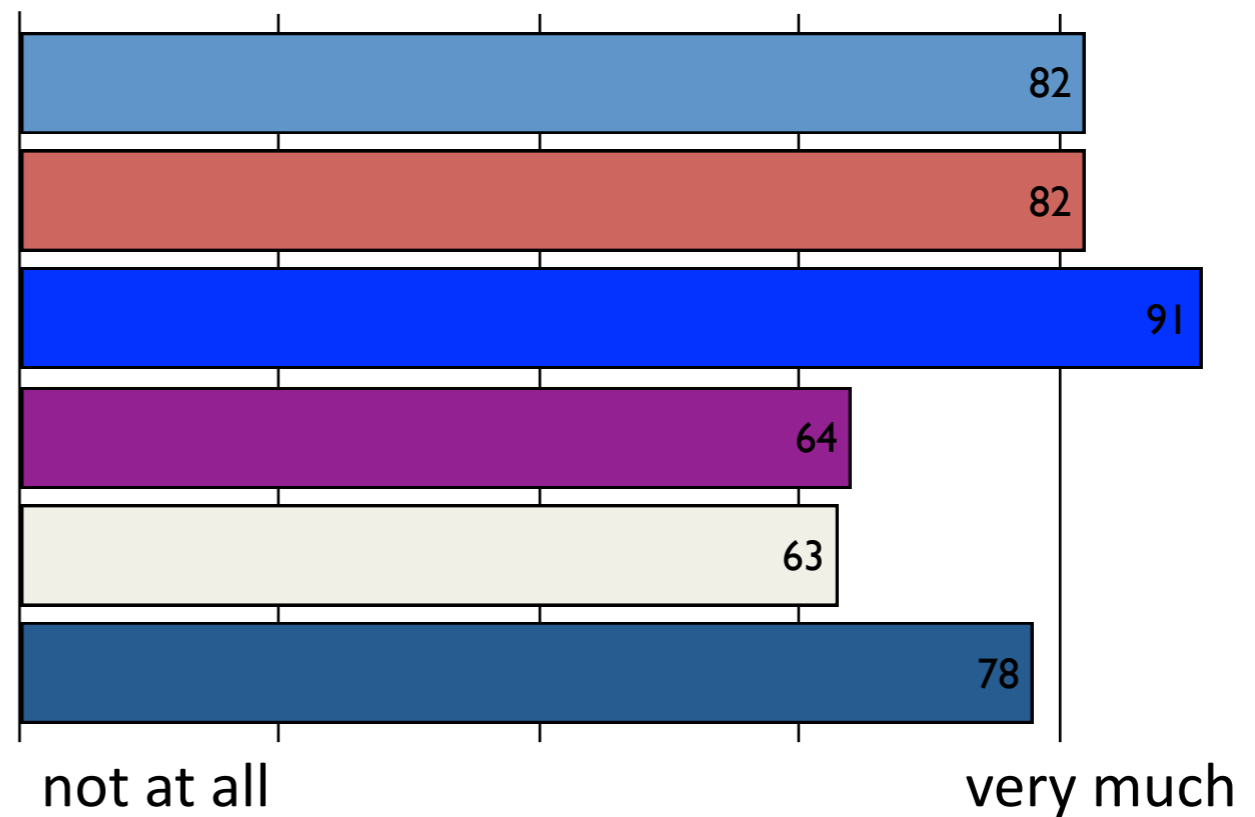
Multiple Visualisations





User Satisfaction

- Stimulating
- Fast
- Reliable
- Straightforward
- Easy to manipulate
- Easy to explore



- Divide: value of the tool vs. manipulation difficulties

Image cluster search

Search Criteria
No. of Matches: 100
Image Size: small

ontology perspectives

- Document
 - has distribution
 - Determination
 - has symptom
 - has mechanism
 - has impact
 - has customer impact
 - has action taken
 - has content
 - Event Context
 - has environmental condition
 - has operating regime
 - has location
 - has event type
 - has customer
 - Part
 - has type
 - has object type
 - has part number
 - has product
 - has feature
 - has position
 - has system

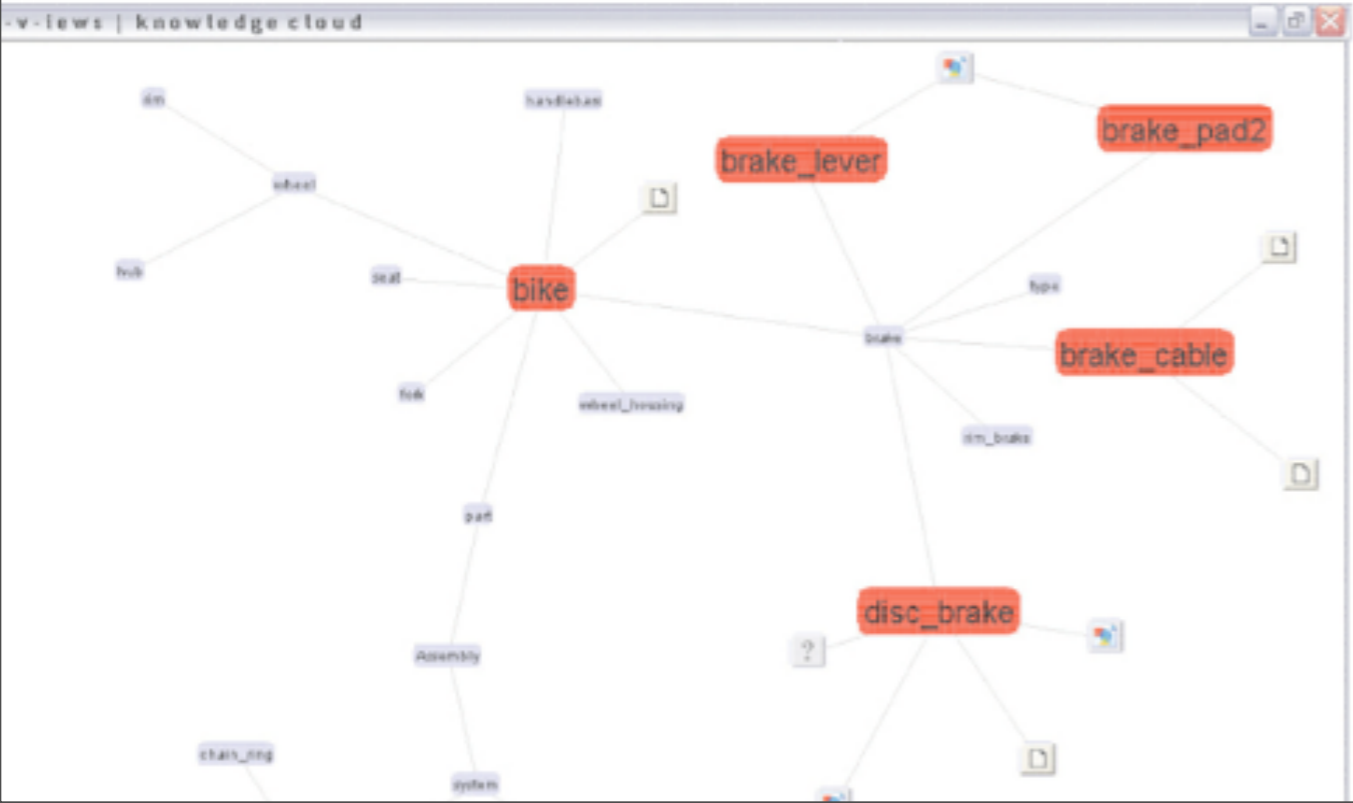
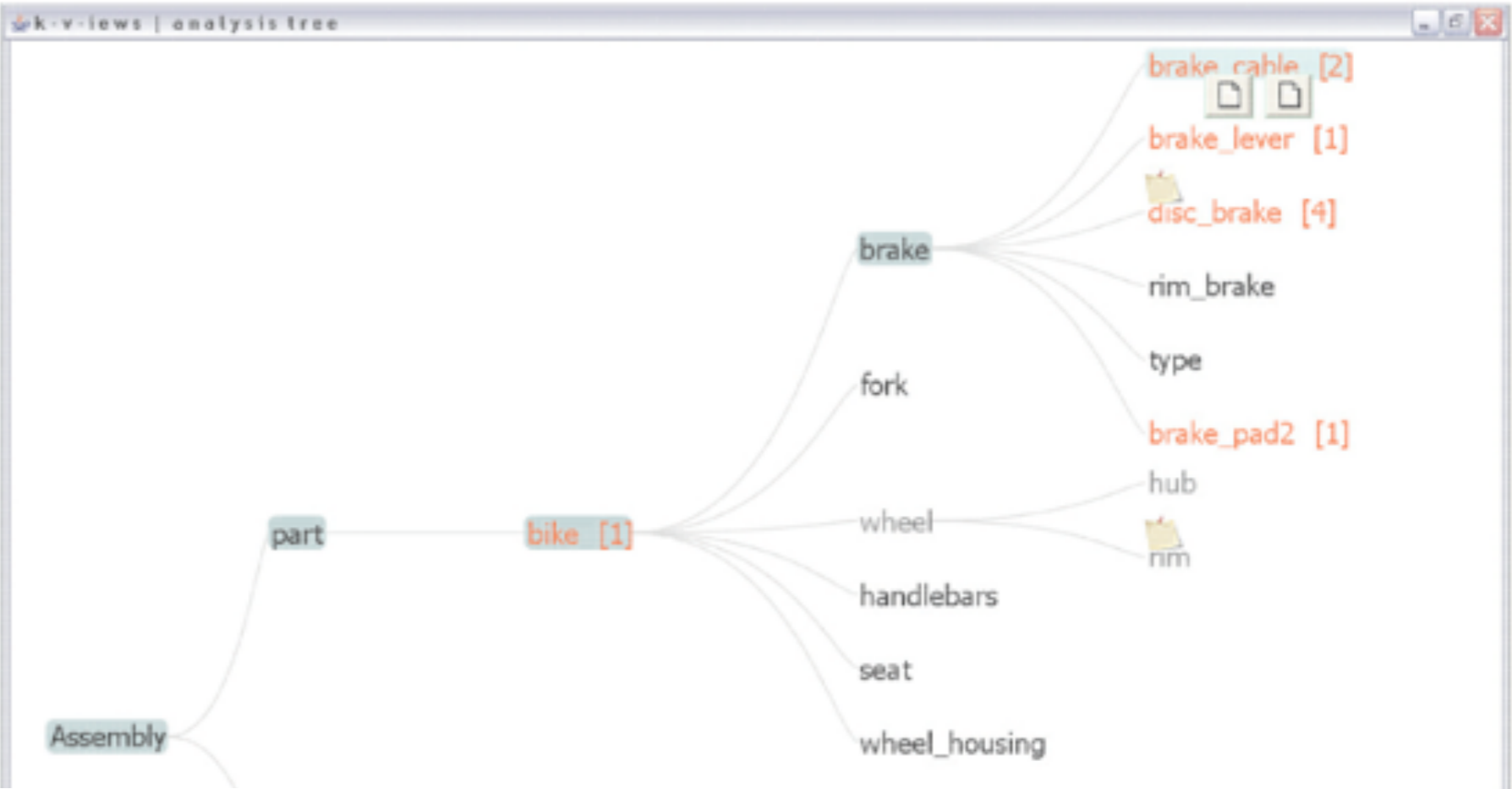
Image
Label: calliper brakes

part type = calliper

Results grid showing various bicycle images with labels like 'assorted_flick', 'bicycle_use_c', 'gears/129744', and 'bicycle_u'.

content-based image retrieval

Results grid showing similar images to the selected calliper brake image, with labels like 'assorted flick imgs/277', 'assorted flick imgs/141', 'assorted flick imgs/201', 'assorted flick imgs/221', 'assorted flick imgs/274', 'assorted flick imgs/251', and 'Similarity' values.



Reuse of Information and Knowledge



Steps in Reuse

- Set up:

- the sense-making activity is triggered by the need for new knowledge.
- Evidence in different forms and from all sources available must be collected

Context Definition

The screenshot shows a software window titled "Definition Stage" with a close button in the top right corner. The form contains the following fields and values:

- *Issue Definition: Bicycle Disk Brake Issue
- *Owner: Miriam
- Attributes**
- *Family: Mountain
- *Serial No.: 3457tr
- Module: Brake System
- Component: Disc Brake
- *Initial Observation: Crack
- Symptoms: cracking of brake cables, wear on housing
- Forensic Evidence: (Two small image thumbnails)
- *Consequences: Mid level
- Risk Level: Medium
- Shipping Location: ESP
- Exit Date: (Empty field)

At the bottom of the window are two buttons: "Suggest Experts" and "Update & Search KB".



Steps in Reuse

- Knowledge gathering:

- knowledge retrieval from the company repositories, supported by exploratory actions

Semantic Search

- The knowledge extracted:

Knowledge Browsing

- document content and metadata describing provenance and uncertainty, is used in subsequent phases.

The screenshot displays the IPAS (Intelligent Product Analysis System) interface. On the left, there's a 'content-based image retrieval' section with search criteria and image results. The main area shows search results for 'leak' with a list of documents. A detailed 'Event Report' is highlighted, showing a tree view of metadata including Report Number, Report Creation Date, Report Author, Operational Effect, Flight Regime, Event Location, Airframe Cycles, Airframe Hours, Engine Installed Location, Fuel Dumped, Delayed Time, Affected Engine, Engine, Engine Serial Number, Engine Type, Installed Part, Component, Part Number, Description, Removed Part, Event Description, and File Location. The event report data includes details for a Boeing 777-200 IGW, WB10, L/N 140, with an event date of 29-Apr-01, engine S/N 51899, and flight regime 'Engine Start'. The event type is 'Operational' and the category is 'Basic'. The reaction to the event shows 'Operational Effect: Delay' and 'Delay Time (mins): 108'. The interface also shows a 'Search' section with keywords and criteria, and a 'Document' list with search results for 'leak'.



Steps (ctd)

- Knowledge analysis and manipulation:
 - support is required for interacting with knowledge, and for capturing the content and context of implicit information
 - e.g., comments, open requests and action lists
 - ontology-based annotations, in order to create new, explicit knowledge.

Knowledge Capture

The screenshot shows the 'x-media | box' application window. The 'Analysis' tab is active, displaying an 'analysis tree' for the ontology 'resources/ontology/Bike_Assembly.owl'. The tree is a hierarchical diagram starting with 'Root_Cause' leading to 'Assembly', which branches into 'part' and 'system'. 'part' further branches into 'bike', 'fork', 'wheel', 'handlebars', and 'saddle'. 'system' branches into 'brake_system' and 'gear_system'. 'brake_system' branches into 'brake', 'brake_cable', 'brake_lever', 'disc_brake', 'rim_brake', and 'brake_pad'. 'brake' branches into 'brake_squeal' and 'brake_chatter'. 'disc_brake' branches into 'worn_brake_pad'. 'rim_brake' branches into 'brake_bearings'. 'wheel' branches into 'hub' and 'rim'. 'rim' branches into 'worn_rim', 'heated_rim', and 'rim_warpage'. 'gear_system' branches into 'cassette' and 'chain_ring'. On the right side of the window, a 'Hypothesis: brake_cable' window is open, showing a 'Set Status' dropdown set to 'Implicated', a text area for 'Enter summary' containing 'appears to be a significant contributor to damage', and three radio buttons for 'Documents', 'Conversations', and 'Q & A Sessions', all currently set to 'None'. There are 'View' and 'Add New' buttons at the bottom of this window.

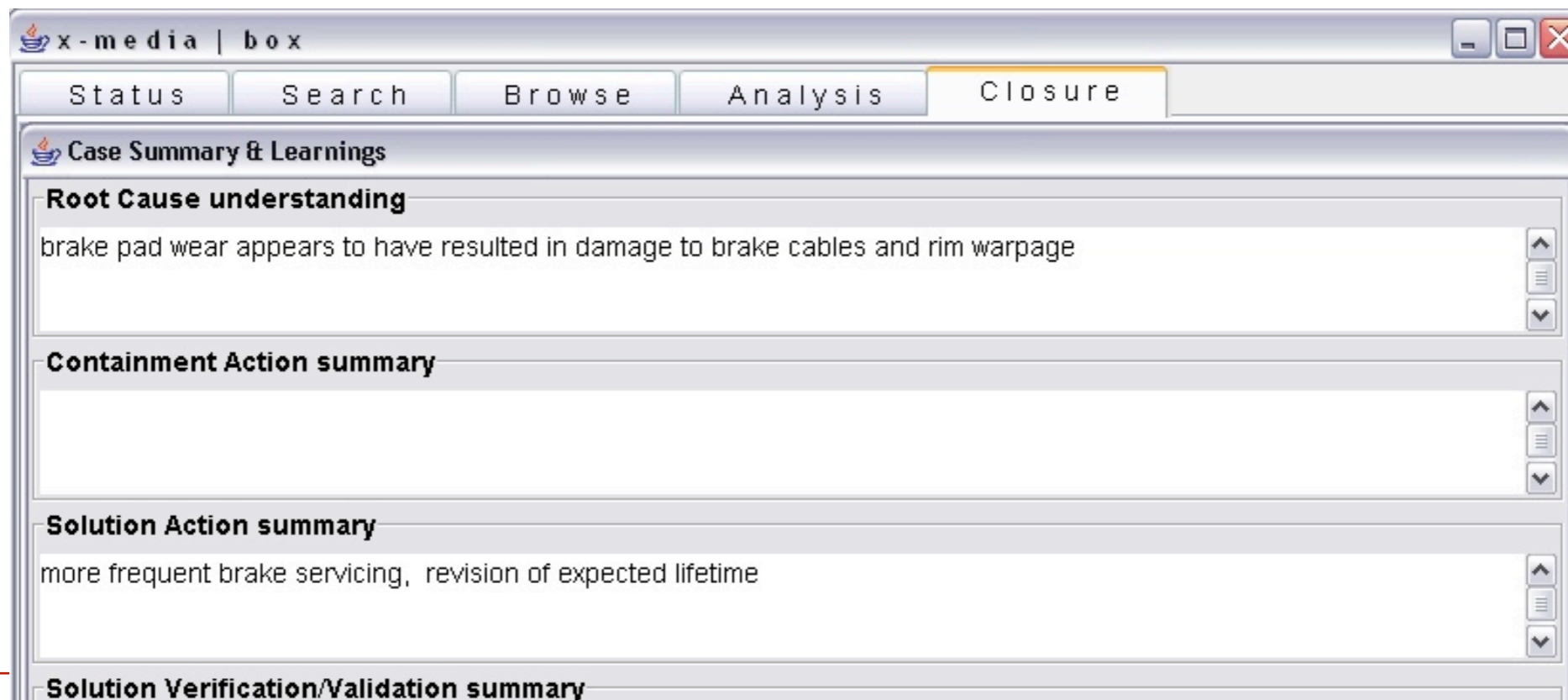


Steps (ctd)

- Recording of the current status:
 - Sense-making may span an extended period.
 - Levels of abstraction that record its status from a high level using summaries, to the level of knowledge analysis
 - the visualisation of islands of knowledge and their ontological relations,
 - to the detail and context of the knowledge – information provenance and co-occurring facts.

Provenance

Knowledge Capture



The screenshot shows the XMediaBox software interface. At the top, there are tabs for 'Status', 'Search', 'Browse', 'Analysis', and 'Closure'. Below these are several panels: 'Definition Stage' with fields for 'Issue Definition' (Disk Brake Failure), 'Owner' (Mitsun), and 'Attributes' (Family: Mountain, Serial No.: MM12345, Module: Brake System, Component: Disc Brake); 'Consequences' (Mid level, Risk Level: Low); 'Shipping Location' (ESP); and 'Exit Date'. A 'Suggest Experts' button is visible. An 'Issue Investigation Tracking' window is open, showing a table with columns for 'When', 'What', 'Where', and 'Who'. The 'When' row contains 'BMX brake pad recall - Apr 2008'. The 'Where' row contains 'disc brake failure on children's BMX Series M34'. A small window titled 'x-media : disc_brak...' displays a photograph of a bicycle wheel. At the bottom, there is a 'Clipboard' section with several icons.

- Set up
- Knowledge Retrieval
- Knowledge Browsing
- Analysis
- Closure



The
University
Of
Sheffield.

Conclusions

About Semantic KM



On KM and SW

- The Semantic WEB offers potentially key technologies to the development of future knowledge Management and the Web
 - More Web than Semantics, but:
 - A little semantics goes a long way (J. Hendler)
- The potential must be exploited addressing real world requirements
 - Rather than in principle AI-oriented requirements (e.g. closed world, small scale, etc.)
 - Scalability at no cost is mandatory (Google quality)
- Strong application pull can be obtained
 - Do not sell slogans, sell ideas and applications!
- Recent industrial funders of my activity:



Rolls-Royce



Environment Agency



- Our spin-out company





Challenges for Knowledge Management

- Complexity of KM Task requires large coordinated effort
 - Knowledge Acquisition is just one of them
 - Text Extraction is not enough
 - Cross-media
 - Integration
- New technologies are needed
 - Largely unsupervised technologies
 - Reasoning over imprecise input
 - Dynamicity of knowledge
- Innovation is across our special

■ User does not need to know it is SW technology



A final thought

- These technologies allow easy collection of and access to a *very* large amount of information/ knowledge
- Are we:
 - Preparing for a better Web/better world?
 - Preparing for a world with no privacy?
 - Big brother
 - Spam
 - Identity theft (e.g. Garlik)
 - Just adding hay to the haystack while searching for a needle?
 - Drowning in triples while trying to avoid drowning in texts?



The
University
Of
Sheffield.

Additional Bibliography



Semantic Search: biblio

- Uren, V., Lei, Y., Lopez, V., Liu, H., Motta, E. and Giordanino, M.: The usability of semantic search tools: a review, Knowledge Engineering Review, in press.
- Kaufmann, E. and Bernstein, A.: How Useful are Natural Language Interfaces to the Semantic Web for Casual End-users? Proceedings of the 6th International Semantic Web Conference and the 2nd Asian Semantic Web Conference, Busan, Korea, November 2007
- Lei, Y., Uren, V. and Motta, E. SemSearch: A Search Engine for the Semantic Web. in 15th International Conference on Knowledge Engineering and Knowledge Management Managing Knowledge in a World of Networks (EKAW 2006). 2006. Pödebrady.
- Guha, R., McCool, R. Miller, E. Semantic Search. in 12th International Conference on World Wide Web. 2003
- Gilardoni, L., Biasuzzi, C., Ferraro, M., Fonti, R., Slavazza, P.: LKMS – A Legal Knowledge Management System exploiting Semantic Web technologies, Proceedings of the 4th International Conference on the Semantic Web (ISWC), Galway, November 2005.
- Rocha, R., Schwabe, D. and Poggi de Aragão, M.: A Hybrid Approach for Searching in the Semantic Web, in the 2004 International World Wide Web Conference, May 17-22, 2004, New York, New York.
- Ravish Bhagdev, Sam Chapman, Fabio Ciravegna, Vitaveska Lanfranchi and Daniela Petrelli: Hybrid Search: Effectively Combining Keywords and Semantic Searches in Proceedings of the 5th European Semantic Web Conference, ESWC 08, Tenerife, June 2008
- Tran, T., Cimiano, P., Rudolph, R. and Studer, R.: Ontology-based Interpretation of Keywords for Semantic Search. Proceedings of the 6th International Semantic Web Conference and the 2nd Asian Semantic Web Conference, Busan, Korea, November 2007
- Catarci, T., Di Mascio, T., Franconi, E., Santucci, G., Tessaris, S. An Ontology Based Visual Tool for Query Formulation Support. in 16th European Conference on Artificial Intelligence (ECAI-04). 2004. Valencia, Spain.
- Kaufmann, E., Bernstein, A. and Zumstein, R. Querix: A natural language interface to query ontologies based on clarification dialogs. In 5th ISWC, pages 980–981, Athens, GA, 2006.
- Corby, O., Dieng-Kuntz, R., Faron-Zucker, C., and Gandon, F., Searching the Semantic Web: Approximate Query Processing Based on Ontologies. IEEE Intelligent Systems, 2006. 21(1)



Automatic Document Annotation: bib

- Fabio Ciravegna. Designing adaptive information extraction for the Semantic Web in Amilcare. In S. Handschuh and S. Staab, editors, Annotation for the Semantic Web, Frontiers in Artificial Intelligence and Applications. IOS Press, 2003.
- Fabio Ciravegna, Sam Chapman, Alexiei Dingli, and Yorick Wilks: Learning to Harvest Information for the Semantic Web, Proceedings of the First European Semantic Web Conference, Crete, May 2004
- A. Kiryakov, B. Popov, et al. Semantic Annotation, Indexing, and Retrieval. 2nd International Semantic Web Conference (ISWC2003), <http://www.ontotext.com/publications/index.html#KiryakovEtAl2003>
- S. Dill, N. Eiron, et al: <http://www.tomkinshome.com/papers/2Web/semtag.pdf> . SemTag and Seeker: Bootstrapping the semantic web via automated semantic annotation. WWW'03.
- Thomas Leonard and Hugh Glaser. Large scale acquisition and maintenance from the web without source access. In Siegfried Handschuh, Rose Dieng-Kuntz, and Steffen Staab, editors, Proceedings Workshop 4, Knowledge Markup and Semantic Annotation, K-CAP 2001, 2001
- Ireson, N., Ciravegna, F., Califf, M.E., Freitag, D., Kushmerick, N., Lavelli, A.: Evaluating Machine Learning for Information Extraction, Proceedings of the 22nd International Conference on Machine Learning (ICML 2005), Bonn, Germany, 2005
- Iria, J. and Ciravegna, F A Methodology and Tool for Representing Language Resources for Information Extraction. In Proc. of LREC 2006, Genoa, Italy, May 2006.
- F. Ciravegna: Challenges in Information Extraction from Text for Knowledge Management, in S. Staab, (ed), "Human Language Technologies for Knowledge Management", IEEE Intelligent Systems and Their Applications (Trends and Controversies), Vol. 16, No. 6, pp 88-90, 2001.
- Fabio Ciravegna. Adaptive information extraction from text by rule induction and generalisation. In Proceedings of 17th International Joint Conference on Artificial Intelligence (IJCAI), 2001. Seattle.
- H. Cunningham, D. Maynard, K. Bontcheva, V. Tablan. GATE: A Framework and Graphical Development Environment for Robust NLP Tools and Applications. 40th Anniversary Meeting of the Association for Computational Linguistics (ACL'02). 2002.
- I. Muslea, S. Minton, and C. Knoblock. 1998. Wrapper induction for semistructured webbased information sources. In Proceedings of the Conference on Automated Learning and Discovery (CONALD), 1998.
- Chakravarthy, A., Lanfranchi, V., Ciravegna, F.: Cross-media Document Annotation and Enrichment, Proceedings of the 1st Semantic Authoring and Annotation Workshop, 5th International Semantic Web Conference (ISWC2006), Athens, GA, USA, 2006
- Handschuh, Staab, Ciravegna. S-CREAM - Semi-automatic CREAtion of Metadata (2002) <http://citeseer.nj.nec.com/529793.html>
- F. Ciravegna, A. Dingli, D. Petrelli, Y. Wilks: User-System Cooperation in Document Annotation based on Information Extraction. Knowledge Engineering and Knowledge Management (Ontologies and the Semantic Web), (EKAW02), 2002.
- M. Vargas-Vera, Enrico Motta, J. Domingue, M. Lanzoni, A. Stutt, and F. Ciravegna. MnM: Ontology driven semi-automatic or automatic support for semantic markup. In Proc. of the 13th International Conference on Knowledge Engineering and Knowledge Management, EKAW02. Springer Verlag, 2002



Sharing, Reuse and Visualisation: biblio

- Petrelli, Mazumdar, Dadzie, Ciravegna: Multi Visualisation and Dynamic Query for Effective Exploration and Annotation of Semantic Data, Proceedings of the 8th International Semantic Web Conference, Washington DC, 25-28 October 2009
- Dzbor, M. - Domingue, J. B. - Motta, E.: Magpie - towards a semantic web browser. 2nd International Semantic Web Conference (ISWC), Sanibel Island, Florida, USA, 2003.
- Lanfranchi, V., Ciravegna, F., Petrelli, D.: Semantic Web-based Document: Editing and Browsing in AktiveDoc, Proceedings of the 2nd European Semantic Web Conference , Heraklion, Greece, 2005.
- Fluit, C., Sabou, M., van Harmelen, F. Ontology-based Information Visualization. In: Geroimenko, V. and Chen C. (eds.) Springer (2003).
- Shneiderman, B. The eye have it: A task by data type taxonomy of information visualization. In Bederson, B. and Shneiderman, B. The craft of information visualization. Morgan Kaufman (2003)
- Paul Mutton and Jennifer Golbeck: Visualization of Semantic Metadata and Ontologies”, in Proceedings of the Seventh International Conference on Information Visualization, 2003
- Stuckenschmidt, H. et al. Exploring Large Document RepositThe DOPE Project, IEEE Computer Society, May/June (2004).
- schraefel, m.c., Wilson, M., Russell, A., Smith, D. mSpace: Improving Information Access to Multimedia Domains with MultiModal Exploratory Search. CACM, 49 (4) 47-49 (2006)
- Fluit, C., Sabou, M., van Harmelen, F. Ontology-based Information Visualization. In: Geroimenko, V. and Chen C. (eds.) Springer (2003).
- Geroimenko, V., Chen, C. (eds.) Visualizing the Semantic Web, Springer, (2003)
- Mutton, P., Golbeck, J. Visualization of Semantic Metadata and Ontologies. 7th International Conference on Information Visualization, IEEE Computer Society (2003)



The University Of Sheffield.

Thank You!

Fabio Ciravegna

20



Director of Research & Innovation in the Digital V
[University of Sheffield](#)

Professor of Language and Knowledge Technolo
[OAK Group, Department of Computer Science,](#)
[University of Sheffield](#)

Director of X-Media,
an Integrated Project Funded by the European Commissio

Director of K-Now Technologies Ltd
a University of Sheffield Spin-off Company

See my [contact page](#) for email and snail mail addresses
I also make available my [FOAF profile](#)

Recent News

- The final review of the [X-Media](#) Project was held on the 25th and 26th of March. The project achievements was de "excellent". The project produced, among several outputs, 3 spin-out companies and 2 patents. Follow-up investr one of the industrial partners to further develop Sheffield's technology will exceed £580,000 over the next 2 years 2010)

fabio@dcs.shef.ac.uk
<http://www.dcs.shef.ac.uk/~fabio/>