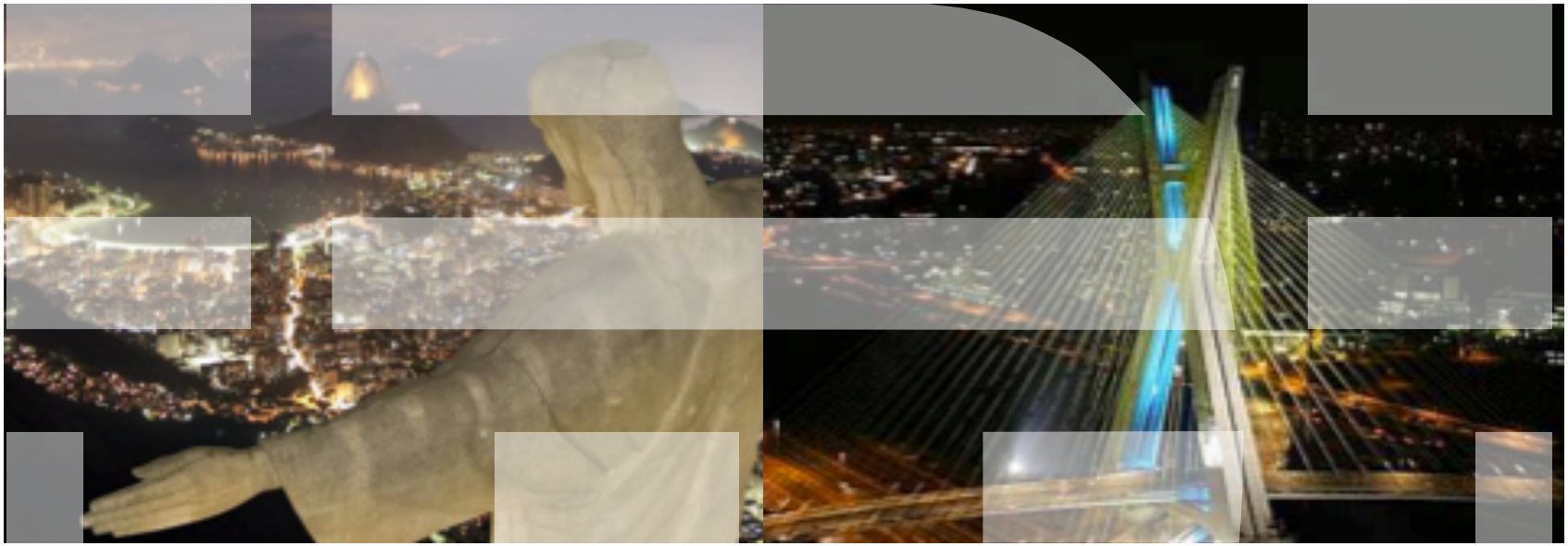
Challenges for Natural Language Processing in O&G Domain

IBM Research Brazil Lab
Alexandre Rademaker, PhD



IBM Research – Brazil: First of three new IBM Research labs in the Southern Hemisphere (2010)

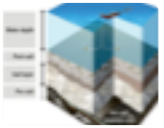
Mission: To be known for our science and technology and vital to IBM, our clients in the region, and globally

Provide differentiation for customer engagements in **Brazil and Latin America** and be a focal point for leveraging world-wide Research technologies and researchers.

IBM @ Tutoia Street
São Paulo

IBM @ Avenue Pasteur
Rio de Janeiro

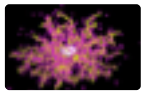
Strategic Focus Areas



Smarter natural resource discovery and logistics with emphasis on NR analytics. Optimization, SW technologies

Service systems with focus to reinvent large-scale service systems, operations, and enterprises.

Human systems with emphasis on Smart cities and Mobile for large-scale events



Smarter devices with emphasis in sensors and microfluidics for Health and Natural Resources

Core Competencies:

- Analytics & Optimization
- HPC & Computational Science
- Distributed Systems & Cloud Computing
- Mobile technologies
- Semiconductor Packaging
- Service Science
- Social Science, Design & Human Computer Interaction

Personal Background

- Master in Computer Science - Formal Methods, Rewriting Logic, Maude
  - "A Rewriting Semantics for a Software Architecture Description Language." Electronic Notes in Theoretical Computer Science 130: 345–77, 2005.
- PhD in Computer Science – PUC-Rio – Logics, Proof Theory
  - "A Proof Theory for Description Logics" Springer Briefs in Computer Science. http://dx.doi.org/10.1007/978-1-4471-4002-3 2010.
  - Ontology alignments via Category Theory
  - Internship at Microsoft Research (SAT Solver) and SRI (many projects, PVS)
- Lecture since 2008 (PUC-Rio and Getulio Vargas Foundation)
  - Data Structures, Mathematical Logic, Formal Languages and Automata, KRR, Functional Programming, Computational Semantics, Category Theory etc
- Joined IBM in Dec 2012 in the Natural Resources Software Technology Group.
- Research Agenda:
  - Lexical Resources, corpora and KBs - OpenWordnet-PT, NomLex-PT, UD Portuguese corpora, OWN-EN, contributions with SUMO etc
  - Information Extraction (PT and EN), NLU …
  - Logics and applications: in particular Law (iALC), also interested in ITP
  - Computational Grammars

Text to KB – why a "focus" on rule-based approach?

Rule-based Information Extraction is Dead!
Long Live Rule-based Information Extraction Systems!

Laura Chiticariu
IBM Research - Almaden
San Jose, CA
chiti@us.ibm.com

Yunyao Li
IBM Research - Almaden
San Jose, CA
yunyaoli@us.ibm.com

Frederick R. Reiss
IBM Research - Almaden
San Jose, CA
frreiss@us.ibm.com

Abstract

The rise of "Big Data" analytics over unstructured text has led to renewed interest in information extraction (IE). We surveyed the landscape of IE technologies and identified a major disconnect between industry and academia: while rule-based IE dominates the commercial world, it is widely regarded as dead-end technology by the academia. We believe the disconnect stems from the way in which the two communities measure the benefits and costs of IE, as well as academia's perception that rule-based IE is devoid of research challenges. We make a case for the importance of rule-based IE to industry practitioners. We then lay out a research agenda in advancing the state-of-the-art in rule-based IE systems which we believe has the potential to bridge the gap between academic research and industry practice.

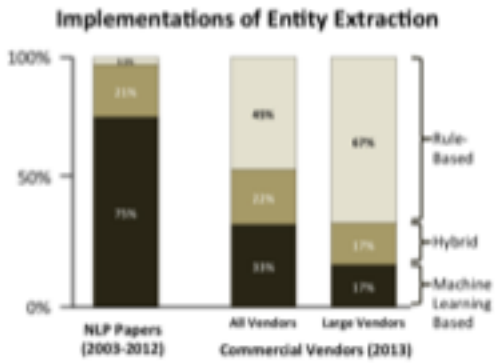


Figure 1: Fraction of NLP conference papers from EMNLP, ACL, and NAACL over 10 years that use machine learning versus rule-based techniques to perform

Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, pages 827–832.

deep vs shallow processing: small words can change everything

- The exact width of the Cochenour Thrust had never been well defined, being estimated to be between 50-150m within the Cochenour-Willans Mine (Hopson, 1994) and narrows towards the later Gold Eagle Deformation Corridor.
- There is no evidence for slumping or small-scale folding as relatively undeformed beds to either side of the structures can be traced laterally.
- Figures 3.10 to 3.15 (except 3.13a) show that these samples all contain a large percentage ofsericite / kaolinite alteration and no feldspar.
- In the case of fluid phase separation, the pre-separation fluid is likely to have been significantly enriched in CO2 (with XCO2 > 0.8) in order to produce the CO2-dominated fluid and to prevent the entrapment of the aqueous fluid.
- In rare cases, very low temperatures (i.e, <50ºC) of unmixing are recorded for some neonates.

Challenges of NLP/NLU in O&G domain

- The clients usually don't know what information they want from texts. They don't know how hard it is to process language.
- Corpora creation is hard. How we should select the 'most useful content'?
- What are the goals?
  - Entities and relations extraction (ABOX)? Terminology extraction (TBOX)? Both?
    "High-quality definitions are the exception rather than the rule in most of the corpora they [terminologists] work with." Meyer (2001)
  - Other downstream task such as QA and `deep' NLU?
    Even if first-order logic were sufficient for NL semantics, there is still a clash of compositionality between semantics and KR to be overcome. Semantic representations must respect the syntactic composition of the texts from which they are derived, to archive a general and systematic syntax-semantics mapping. Consequently, the semantic representations assigned to sentences tend to be more complex, and different, than the representations a knowledge engineer would assign on a case-by-case basis when targeting a particular knowledge base.
- What are the killer application for NLP/NLU in O&G?

Christopher D. Manning Inbox - Gmail Yesterday 16:07 Details CM

Re: [java-nlp-user] Regarding Core NLP abilities

To: [redacted], Cc: java-nlp-user@lists.stanford.edu, java-nlp-user

Hi Archana,

This question is much too vague to be answerable.

Probably, it's your job!

Chris.

> On Jun 6, 2017, at 2:27 AM, [redacted] wrote:
>
> Hi Team,
>
> We are working in a Search System project and using Core NLP for the natural language processing.
>
> However we wish to add intelligence to the existing system, So we would want to understand how to impart intelligence and Inference abilities to the existing system using Core NLP.
>
> Thanks in Advance!
> Regards,
> [redacted]
>
> _____
> java-nlp-user mailing list
> java-nlp-user@lists.stanford.edu
> https://mailman.stanford.edu/mailman/listinfo/java-nlp-user

O&G is very technical : we need experts

Zircon from a quartz-feldspar porphyry stock (Brewis porphyry) along Balmer Lake gives an age of 2726 ± 4 Ma, which is significantly younger than the previous age determination of 2742 +3/-2 Ma.

1. Zircon from a quartz-feldspar porphyry stock along Balmer Lake
2. quartz-feldspar porphyry stock along Balmer Lake
3. Brewis porphyry
4. quartz-feldspar porphyry stock

3 = 2 and for understand 2 and 1 we need to understand 4

- quartz = http://wnpt.brlcloud.com/wn/synset?id=14693733-n
- feldspar = http://wnpt.brlcloud.com/wn/synset?id=14864961-n
- porphry = http://wnpt.brlcloud.com/wn/synset?id=14996395-n
- stock = ? => https://en.wikipedia.org/wiki/Stock_(geology)

O&G is very technical : we need experts

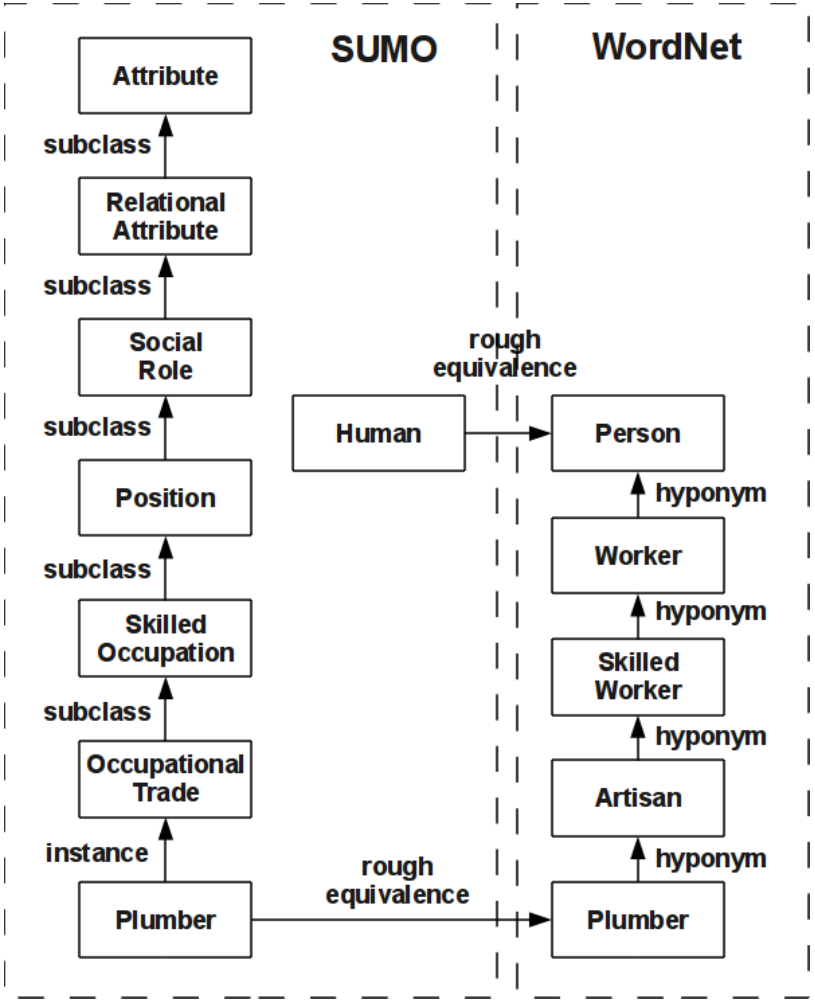
The Quartz Actinolite Zone comprises some of the samples with gold grades up to 701 g/t (internal Goldcorp data) but in general gold grades in this zone range from 0 to 30 g/t (average of 23.4g/t; Fig 3.9).

We want to recognize the indirect mention of the basic statements:

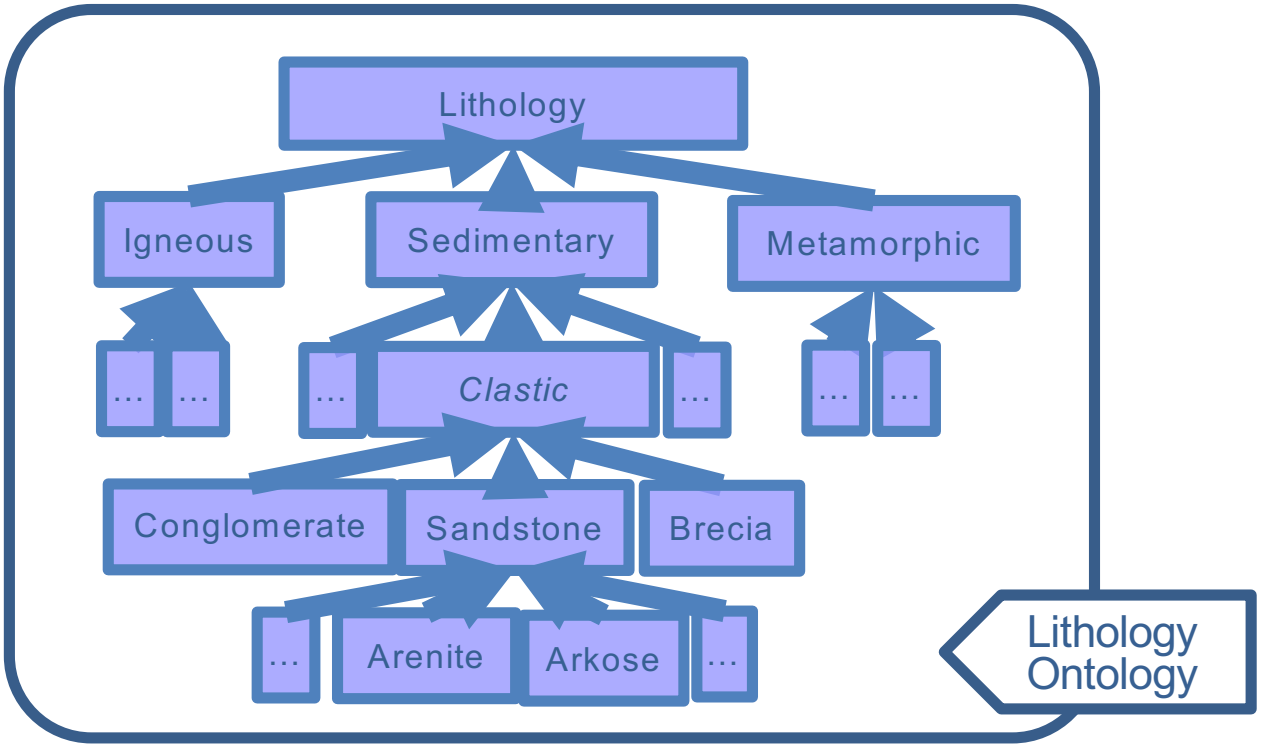
- SOME samples of "The Quartz Actinolite Zone" contains GG "up to 701 g/t"
- GENERAL "The Quartz Actinolite Zone" has range of GG from 0 to "30 g/t"
- GENERAL "The Quartz Actinolite Zone" has average of GG "23.4 g/t"

KB <~> Information Extraction

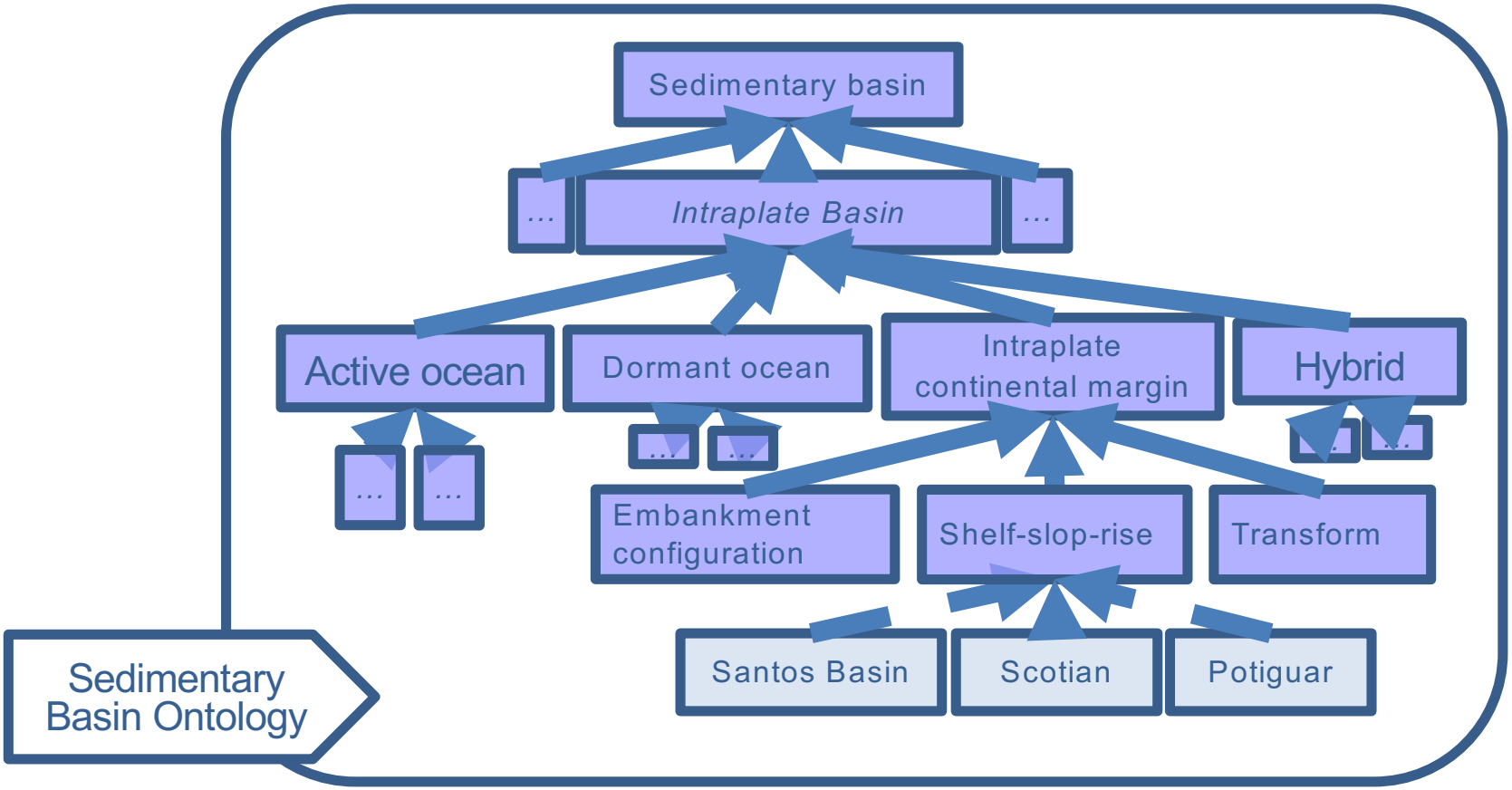
- Two-way road: Information Extraction consumes KB and produces KB.
- **Long tail information about Entities are in texts!**
- Lexical Resources: `lightweight' ontologies? Are they easier to maintain?
  - https://github.com/own-pt/own-en/blob/master/dict/noun.geosci.txt
- What are the good ontologies in O&G domain? Does this question make sense?
- Ontologies can be more expressive than DL/OWL, does it makes sense? Does it help? What are the good Upper Level Ontologies?
- "all grammars leak" (Edward Sapir) and ontologies too!
- What ad-hoc discretization decisions during domain modeling can we postpone?



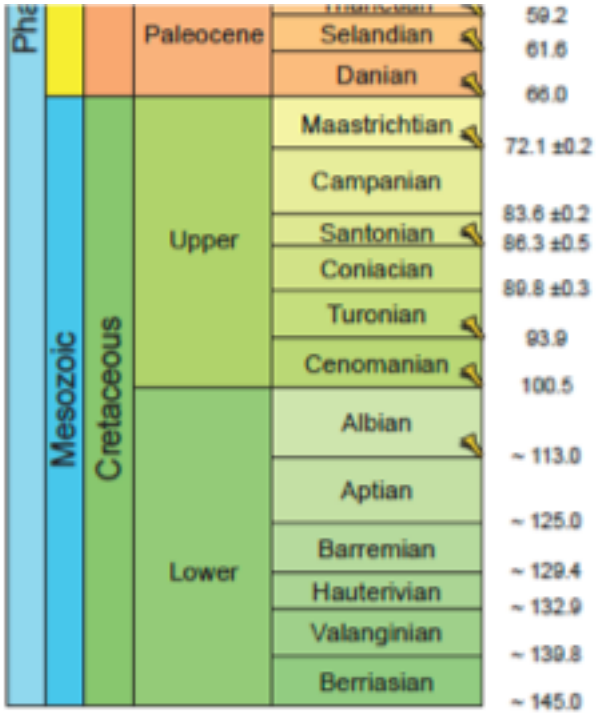
Galp and SLB ontologies (example)

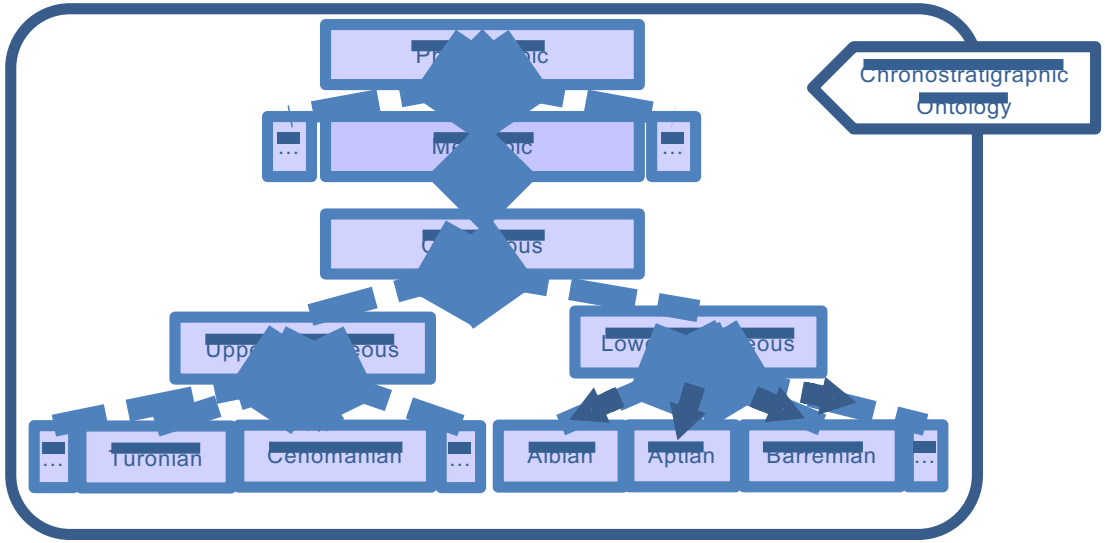


Galp and SLB ontologies (example)



Galp and SLB ontologies (example)





TKB in a nutshell

- Two possible approaches for information extraction
  - Annotation of entities and relations:
    - We will need to engage with successive refinements of the entities and relations model from the beginning
    - We will need specialists in the domain and train them to annotate documents in a consistent manner.
    - Time consuming
  - Sentences analysis for entities/relation extraction
    - Given a clean text files, robust parses (i.e. ESG) exists. We can create corpora and reuse the model.
    - Not all sentences will receive analysis and many will receive more than one.
    - Post-pone the discussion about which entities and relation we want to extract.
    - Mining the trees can be productive.
- The mining strategy is linguistically motivated: inspired by (Hearst, 1992) we assume that certain semantic relations have a linguistic realization, and therefore the inclusion of linguistic metadata such as part-of-speech, lemma, and syntactic information in the corpus is essential.

TKB in a nutshell

- Sentence Segmentation (split running text into sentences)
  - OpenNLP vs ESG (evaluating)
- IBM English Slot Grammar (parsing text to trees) (Lua wrapper), substituted statistical parsers (lack of corpora from the domain)
- Wordsense disambiguation (graph-based)
  - UKB with **GeoNames** ontology for geo named entities
  - UKB with **WN-EN** for WSD all remain open-class words (linking to **ontologies**)
- Prolog Rules:
  - Named entities using anchor words (fields, basin, rocks, deposits etc)
  - Quantities + units detection using dependencies patterns
  - Dependencies patterns to semantic relations and compounds
- Prolog to RDF
- Code restricted at https://github.ibm.com/brl-krr/tkb
- Demos, notebooks, dockers etc
- Common Lisp, Prolog, Python

ESG (1 of 13 trees) augmented with LF

There is no evidence for slumping or small-scale folding as relatively undeformed beds to either side of the structures can be traced laterally.

```
.--------------- subj(n)        there1(1)              noun pron sg def advnoun locnoun loc
o--------------- top            be(2,1,7)              verb vfin vpres sg vsubj absubj auxv
| .------------- ndet           no1(3)                 det sg no
| .------------- lconj          evidence1(4,u,5,u,u)   noun cn sg abst cognsa comm (latrwd 0.046150)
| | `----------- nobj(p)        for1(5,4,6)            prep pprefv nonlocp pobjp
| |   `--------- objprep(ing)   slump1(6,u,u)          verb ving vchng (nform slump)
`-+------------- pred(n)        or1(7)                 noun cn sg act cord cognsa comm abst process
| | .----------- nadj           small-scale1(8,9)      adj
| `------------- rconj          folding1(9,u,u)        noun cn sg act process (vform fold)
`--------------- vsubconj       as3(10,20)             subconj okadjsc oknounsc oknsubconj comparsc assc tosc
  |     .------- adjpre         relatively1(11)        qual
  |   .--------- nadj           un+deformed1(12,13)    adj
  | .-+--------- subj(n)        bed1(13,u)             noun cn pl tonoun physobj artf inst ent (latrwd 0.056520)
  | | `--------- nprep          to2(14,13,16)          prep pprefv motionp
  | |   | .----- ndet           either1(15)            det sg indef
  | |   `-+----- objprep(n)     side1(16,17,u)         noun cn sg location ent (latrwd 0.113050)
  | |     `----- nobj(n)        of1(17,16,19)          prep pprefn nonlocp
  | |       | .- ndet           the1(18)               det pl def the ingdet
  | |       `--- objprep(n)     structure1(19,u)       noun cn pl physobj abst property massn artf strct ent
  `-+----------- sccomp(bfin)    can1(20,13,21)         verb vfin vpres pl vsubj auxv
    `----------- auxcomp(binf)  be(21,13,22)           verb vinf absubj auxv
      `--------- pred(en)       trace1(22,u,u,13)      verb ven vpass vchng (nform tracing) (ernform tracer)
        `------- vadv           laterally1(23,22)      adv
```

ESG (1 of 13 trees) augmented with LF

There is no evidence for slumping or small-scale folding as relatively undeformed beds to either side of the structures can be traced laterally.

```
  there(e1,x1) pos(e1,EX)
be(e2,x2,x1,x7) vpres(x2) pos(e2,VBZ)
    no(e3,e7) pos(e3,DT)
    evidence(e4,x4,u,e5,u,u) pos(e4,NN)
      for(e5,x4,e6) pos(e5,IN)
        slump(e6,x6,u,u) pos(e6,VBG)
  or(e7,x7,x4,e9) pos(e7,CC)
      small-scale(e8,e9) pos(e8,JJ)
    folding(e9,x9,u,u) pos(e9,NN)
  as(e10,e2,e20) pos(e10,IN)
          relatively(e11,e12) pos(e11,RB)
        un+deformed(e12,x13) pos(e12,JJ)
      bed(e13,x13,u) pl(e13) pos(e13,NNS)
        to(e14,x13,x16) pos(e14,IN)
            either(e15,e16) pos(e15,DT)
          side(e16,x16,e17,u) pos(e16,NN)
            of(e17,x16,x19) pos(e17,IN)
                the(e18,e19) pos(e18,DT)
              structure(e19,x19,u) pl(e19) pos(e19,NNS)
    can(e20,x20,x13,e21) vpres(x20) pos(e20,MD)
      be_pass(e21,x21,x13,e22) pos(e21,VB)
        trace(e22,x22,u,u,x13) pos(e22,VBN)
          laterally(e23,e22) pos(e23,RB)
```

TKB (old) demos

- http://wnpt.brlcloud.com/kb-extraction/search?db=onepetro&term=*:*
- http://wnpt.brlcloud.com/kb-extraction/search?db=onepetrotr&term=petrobras
- http://wnpt.brlcloud.com/demo
- Jupyter Notebook

One example

- "They are the Arrecife Medio, Isla de Lobos, Tiburon, Bagre, Atun, Morsa, Escualo, Marsopa, and Carpa fields, and to date the estimated cumulative production is some 210 million BOE of light crude ranging from 30 to 40° API from the El Abra Middle Cretaceous limestone."

- The Cretaceous is divided into Early and Late Cretaceous epochs, or Lower and Upper Cretaceous series. In older literature the Cretaceous is sometimes divided into three series: Neocomian (lower/early), Gallic (**middle**) and Senonian (upper/late). A subdivision in eleven stages, all originating from European stratigraphy, is now used worldwide. In many parts of the world, alternative local subdivisions are still in use. (https://en.wikipedia.org/wiki/Cretaceous)

- Demo NLU
- Notebook

TKB in context

- There are more than just text processing
- PDF to text
- Images processing and classification
- Knowledge Explorer during a specify task (images annotation).

The assessment of the quality of TKB

- Comparing ESG with statistical parsers? No available corpora for the specific domain.
- Understanding ESG and its underline linguistic theory comparing to UD?
- Evaluate ESG in a downstream task
  - SRL?
  - Information Extraction

ESG vs UD

- UD “content words first” principle. In ESG the head of the PP are the prepositions.
  - In UD, copula verbs and its subject have the predicate as its head.
- UD rejects the argument vs adjunct distinction using core arguments and oblique modifiers. ESG has complement slots (obligatory or optional) and adjunct slots (can be filled multiple times).
- MWE. In UD we have special relations (flat, compound and fixed). ESG treat MWE as single token called “normal lexical multiword”. Another alternative is the “structural lexical multiword” (transparent analysis). Finally, the “quote nodes” (artificial nodes).
- Coordination and tokenization: UD consider each parenthesis a token. In cases where parenthesis define a conjunction, ESG takes a pair of parenthesis as a single word, the head of the conjunction.
- Other cases …

ESG vs UD: ellipses

UD “promoting” words to the position of the elided one, or using the special “orphan” dependency relation. It is not clear how ESG treats cases of ellipsis, as it is not clearly documented.

```
The average porosity ranges around 30 – 35%, and permeability
2000 – 10000 mD.
_____________________________________________________________________________
  .--------- ndet       the1(1)                 det sg def the ingdet
  .--------- nadj       average1(2,3)           adj nqual
.----------- subj(n)    porosity1(3,u)          noun cn sg abst property massn
o----------- top        range3(4,3,u,u,u)       verb vfin vpres sg vsubj
`----------- vprep      around1(5,4,106)        prep badobjping
  | .------- lconj      30(6,u)                 noun num sg pl sgpl
  `-+------- objprep(n) -(106)                  noun cn pl notfnd cord
    | .----- lconj      35%(7,u)                noun num sg pl sgpl
    `-+----- rconj      and1(8)                 noun cn pl notfnd cord
      | .--- lconj      permeability1(9,u,u)    noun cn sg abst property massn
      | | `- nprop      2000(10,u)              noun num sg pl sgpl yr
      `-+--- rconj      -(110)                  noun cn sg notfnd cord massn property abst
        | .- nadj       10000(11,u)             noun num sg pl sgpl
        `--- rconj      mD.(12)                 noun propn sg notfnd
_____________________________________________________________________________
```

ESG vs UD: quantities

The “as 1 Darcy” as filling a slot for comparatives (avcompar), it is analyzed as an adverb of “make”, being unrelated to “permeability”, which is in a separate dependency subtree. None of the four analyses simultaneously make “as high as 1 Darcy” a subtree by itself and a dependent of “permeability”.

```
The permeabilities (as high as 1 Darcy) make them effective
reservoirs.
_____________________________________________________________________________
  ._____ ndet       the1(1)              det pl def the ingdet
._______ subj(n)    permeability1(2,u,u) noun cn pl abst property massn
| ._____ advpre     as2(3)               qual pre badattrib soqual
.-+_____ vadv       high2(4,8,u)         adv erest partf
| `_____ avcompar   as1(5,4,7)           prep pprefv nonlocp pobjp asprep
|   | .- nadj       1(6,u)               noun num sg
|   `--- objprep(n) Darcy1(7)            noun propn sg h anim gname sname
o_______ top        make1(8,2,11,u,u,u)  verb vfin vpres pl vsubj (nform ..) (ernform maker makeover)
| ._____ ndet       them2(9)             det pl def
| ._____ nadj       effective1(10,11,u)  adj
`_______ obj(n)     reservoir1(11,u)     noun cn pl locn physobj inst ent container (latrwd 0.023080)
_____________________________________________________________________________
```

ESG vs UD: appositions

- UD appos vs nmod relations
  - An appositional modifier of a noun is a nominal immediately following the first noun that serves to define, modify, name, or describe that noun. It includes parenthesized examples, as well as defining abbreviations in one of these structures.
  - The nmod relation is used for nominal dependents of another noun or noun phrase and functionally corresponds to an attribute, or genitive complement.
- The syntactical analysis of ESG seems to raise similar doubts. There are the following distinct noun adjunct slots, but no more informative descriptions:
  - nnoun (“the boat house”)
  - nappos (“John, my brother”, “Paris, the capital of France”) and
  - nprop (“company X”)

ESG vs UD: appositions

This area contains the Tupi discovery (Lula and Cernambi Fields)
announced as having up to 8 BBbls of oil reserves.

```
  .─────────── ndet        this1(1)              det sg def
.─────────── subj(n)     area1(2,u)            noun cn sg location ent (latrwd 0.278820)
o─────────── top         contain1(3,2,106,u)   verb vfin vpres sg vsubj (nform containment)
|   .─────── ndet        the1(4)               det sg def the ingdet
|   .─────── nadj        Tupi1(5)              noun propn sg h cpropn capped liv ent lang comm
| .─────────── lconj      discovery1(6,u,u)     noun cn sg act accomp speechact (vform discover)
`─+─────────── obj(n)     ((106)                noun cn sg act cord accomp speechact
  | .─────── lconj      Lula(7)               noun propn sg capped notfnd
  `─+─────── rconj      and1(8)               noun propn pl glom h capped cord
  | `─────── rconj      Cernambi Fields1(10)  noun propn sg glom h capped performer entertainer
  `─────────── nnfvp      announce1(11,u,106,u) verb ven (nform announcement) (ernform announcer)
  | `─────── vsubconj   as3(12,13)            subconj okadjsc oknounsc oknsubconj assc tosc
  |   `───── sccomp(ing) have2(13,u,17)        verb ving ingprep sta badvenadj supportv
  |     |   .- nadv      up to13(15,16,u)      prep amtmod sepmw
  |     | .── nadj      8(16,u)               noun num sg pl sgpl
  |     `──── obj(n)     BBbls(17)             noun propn sg notfnd
  `─────────── nprep      of1(18,106,20)        prep pprefn nonlocp
    | .────── nnoun      oil1(19,u)            noun cn sg physobj massn sbst artf ent
    `─────── objprep(n)  reserve2(20,u)        noun cn pl h plmod physobj abst massn anim
```

ESG: incomplete

Potential porosities are high, the result of complex diagenetic histories.

```
o──────── top         incomplete(11)    incomplete
| .────── nadj        potential1(1,2)   adj capped
`──────── u           porosity1(2,u)    noun cn pl abst property massn
`──────── u           be(3,4,6)         verb vfin vpres pl q vsubj absubj auxv
  `────── subj(n)     high3(4,u,u)      noun cn sg abst property massn (latrwd 0.061050)
  | .──── ndet        the1(5)           det sg def the ingdet
  `─+──── pred(n)     result4(6,7,u)    noun cn sg (latrwd 0.065270)
    `──── nobj(n)     of1(7,6,10)       prep pprefn nonlocp
      | .- nadj       complex1(8,10,u)  adj
      | .- nadj       diagenetic1(9,10) adj
      `── objprep(n) history1(10,u,u)  noun cn pl abst (latrwd 0.065270)
```

ESG: citations

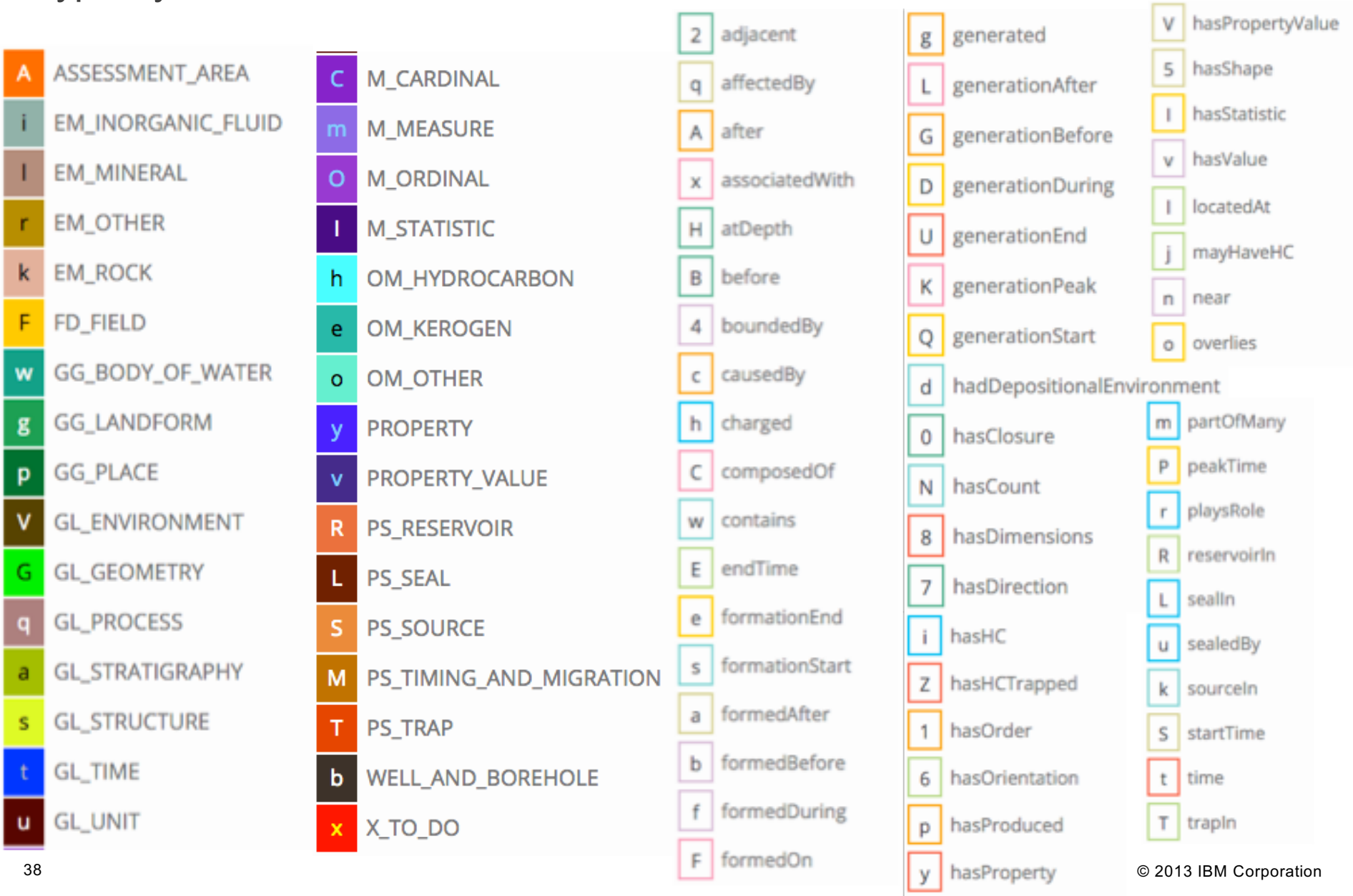
```
The high potential porosities make them effective reservoirs
(Guardado et al., 2000).
```

```
  .-----> ndet    the1(1)              det pl def the ingdet
  .-----> nadj    high1(2,4,u)         adj erest nqual lmeasadj
  .-----> nadj    potential1(3,4)      adj
.-------> subj(n) porosity1(4,u)       noun cn pl abst property massn
o-------> top     make1(5,4,108,u,u,u) verb vfin vpres pl vsubj badvenadj supportv badnen vchng
| .-----> ndet    them2(6)             det pl def
| .-----> nadj    effective1(7,u,u)    adj
| .-----> lconj   reservoir1(8,u)      noun cn pl locn physobj abst artf inst wlocn ent
`-+-----> obj(n)  ((108)               noun cn pl locn physobj abst artf inst cord wlocn
  | .---> lconj   Guardado(9)          noun propn sg notfnd
  | | `-> nadjp   et al.1(10,u)        adv
  `-+---> rconj   ,(111)               noun sg yr cord notfnd
    `---> rconj   2000(12,u)           noun num sg pl sgpl yr
```

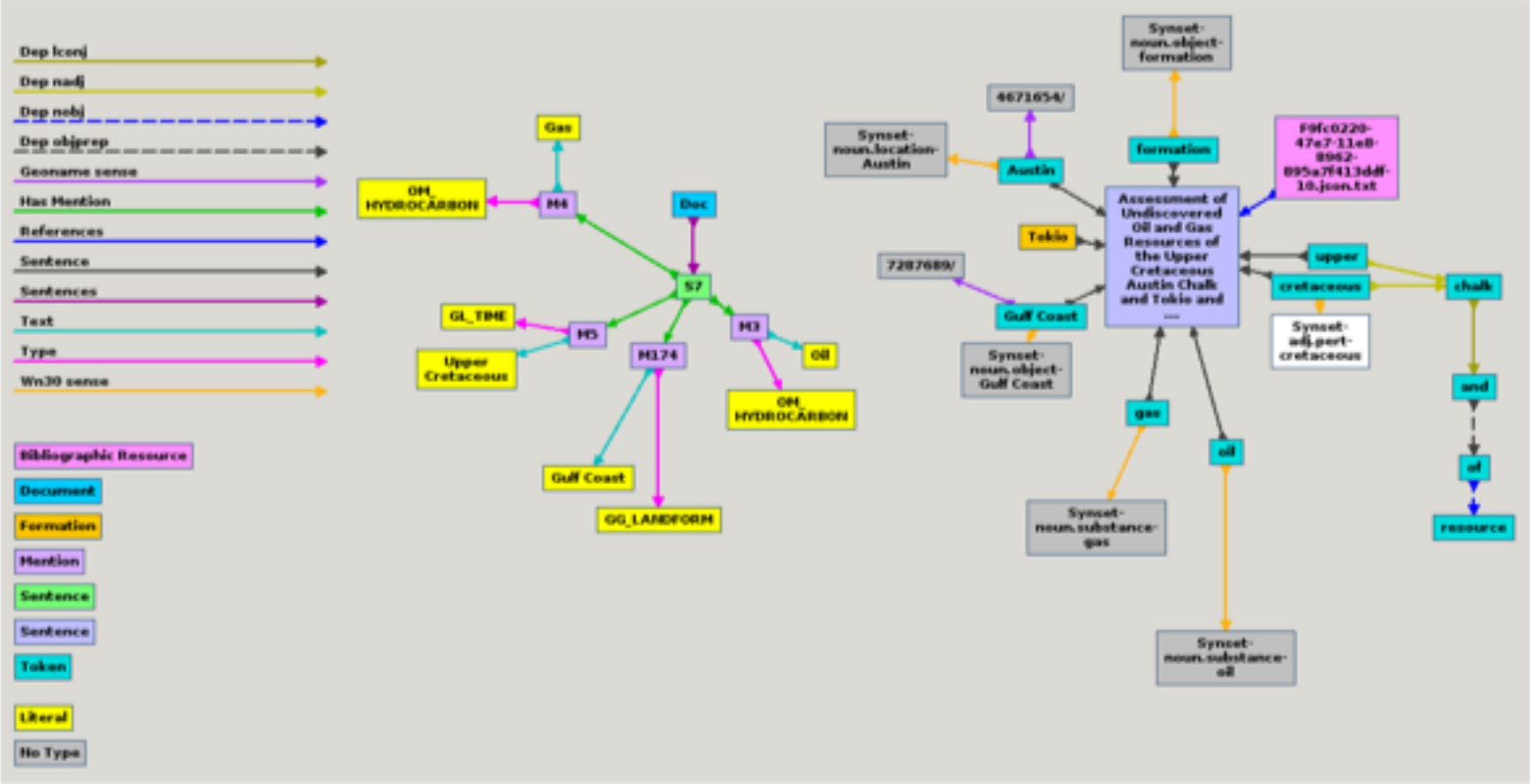
Previous Project

- The `golden set' (GS) were randomly selected from a corpus of 1298 publicly available English language geological reports, published by the United States Geological Survey (USGS), Geological Survey of Canada (GSC), and British Geological Survey (BGS). 155 text passages (7007 sentences) relevant to petroleum systems were extracted. Multiple occurrences of the same entity in a document were annotated as co-references.
- The documents were annotated by individuals with a background in geology, all with oil industry experience. In total, 38,322 mentions of the 32 entity types were annotated. Inter-annotator agreement for mentions reached 0.84, and documents annotated by more than one annotator were adjudicated to arrive at a final version.
- 32 entities types drawn from the GeoSciML, and expanded with petroleum system and exploration concepts. The entity types can be broadly categorized as physical (earth materials, organic materials), geographical, geological including geological time, petroleum system, field development, and property/measurement.
- The type system also defines 653 relations between these entity types, such as `formedDuring', `causedBy', and `composedOf'. But only 53 relations occurs in the golden set.
- **Hard to compare different data models!**

Type System

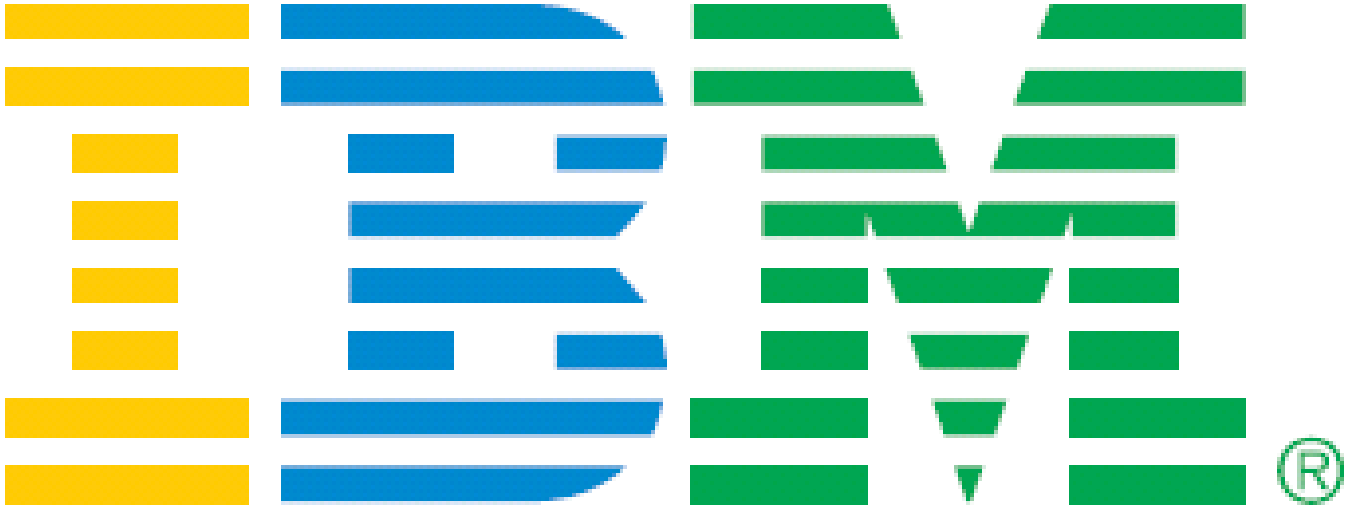


Evaluating TKB using annotated data (Gruff)



Conclusion

- Architecture:
  - file based ~> UIMA vs micro-services possible sharing data using a MQ service?
  - Turn TKB less dependent from RDF/Semantic Web?
  - Extract versus annotations inline, visualization, debug
- ESG gives deep analysis (plus LF) and we are using only the syntactic information
  ~> relations extraction can be very productive after we solve it.
- ESG vs HPSG:
  - robust semantic representation
  - community, tools
- We need time for design and run a performance and error analysis in a compiled and representative corpus for the domain, given a well-defined IE goals.
- Thesaurus, lexical resources are necessary.
- Small projects doing specific tasks (text entailment) with specific corpora (SICK) and resources help a lot to clarify ideas.



IBM Research – Brazil
http://www.research.ibm.com/brazil/