



Framework Jornalismo 3.0

Edison Ishikawa



Objetivo

- Mostrar um esboço de framework para um Framework de Produção Jornalística Semântico
- Fornecer uma visão geral dos problemas de implementação para aplicações de Computação Semântica

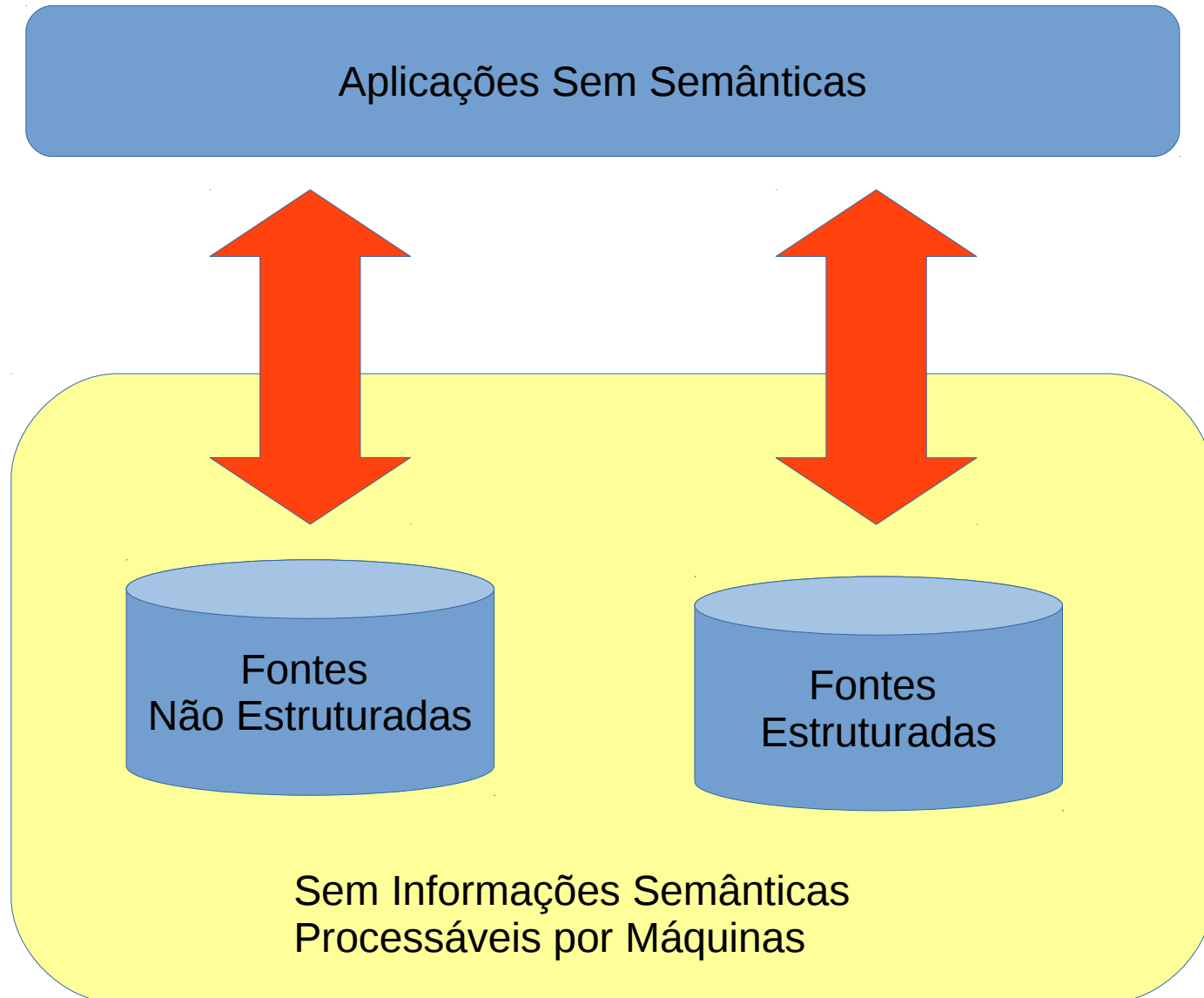


Sumário

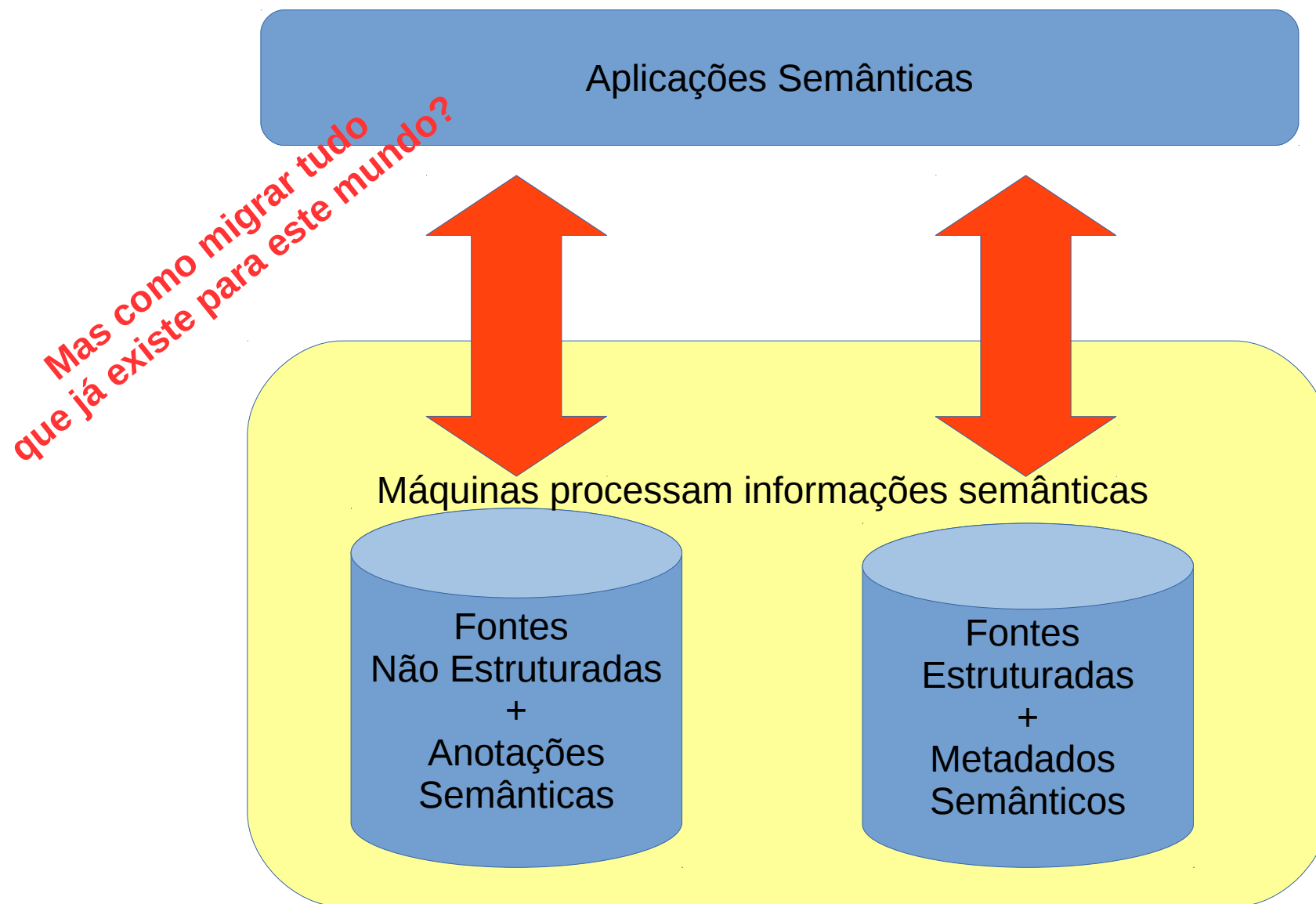
- Introdução
- Desenvolvimento
- Considerações Finais



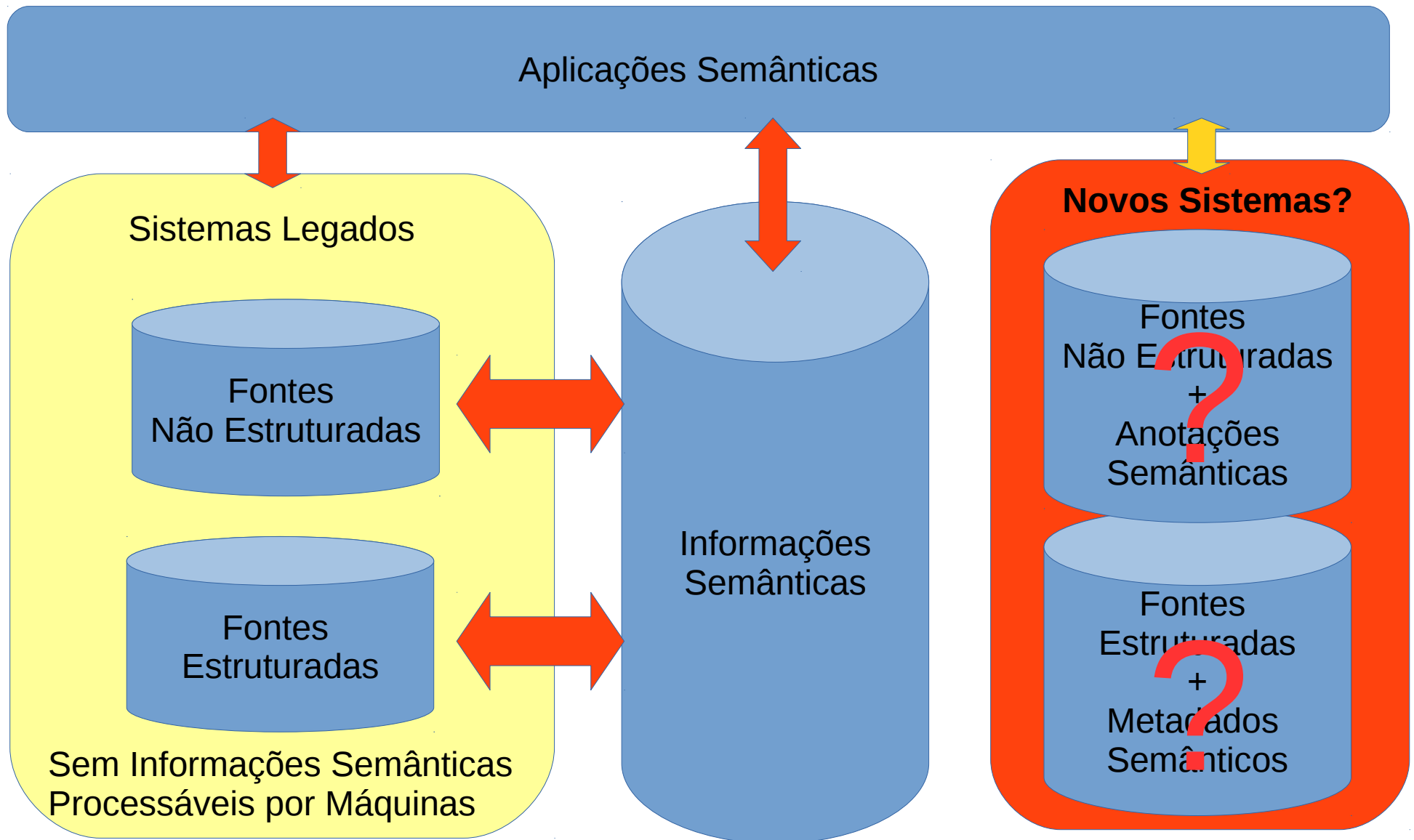
Sistemas de Informações Atuais Não Semânticos



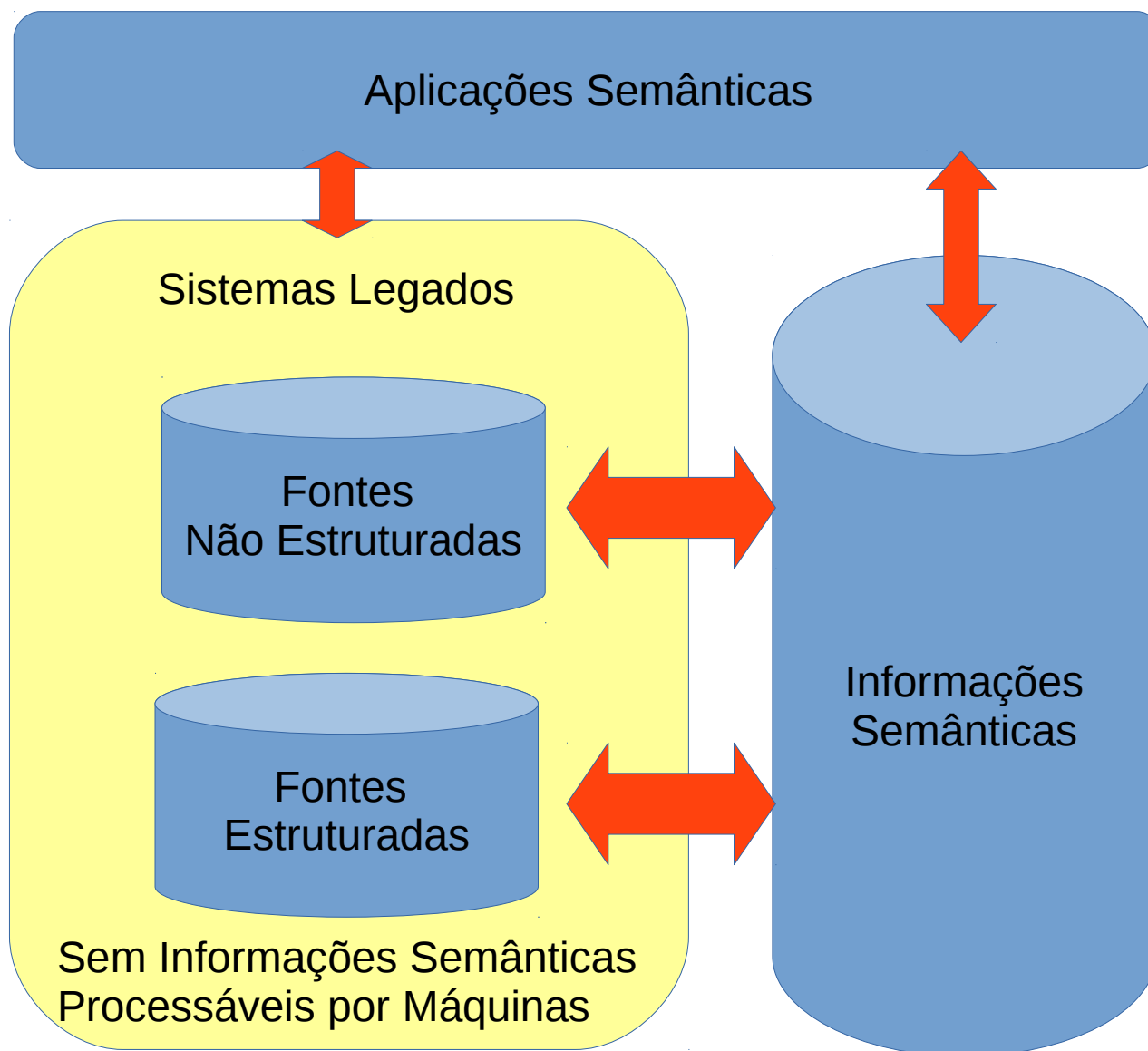
Sistemas de Informações Semânticos Idealizado



SI Semântico – Futuro?

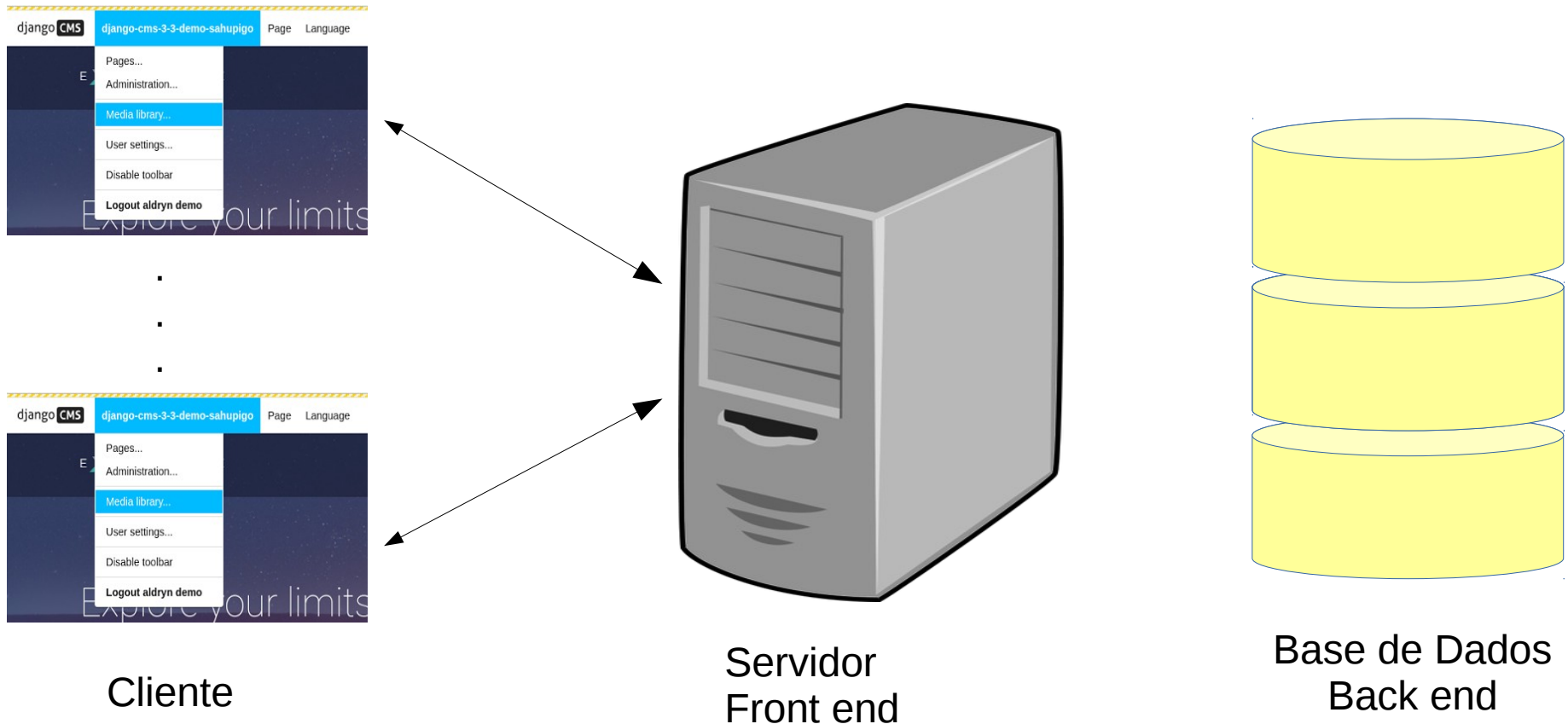


SI Semântico?



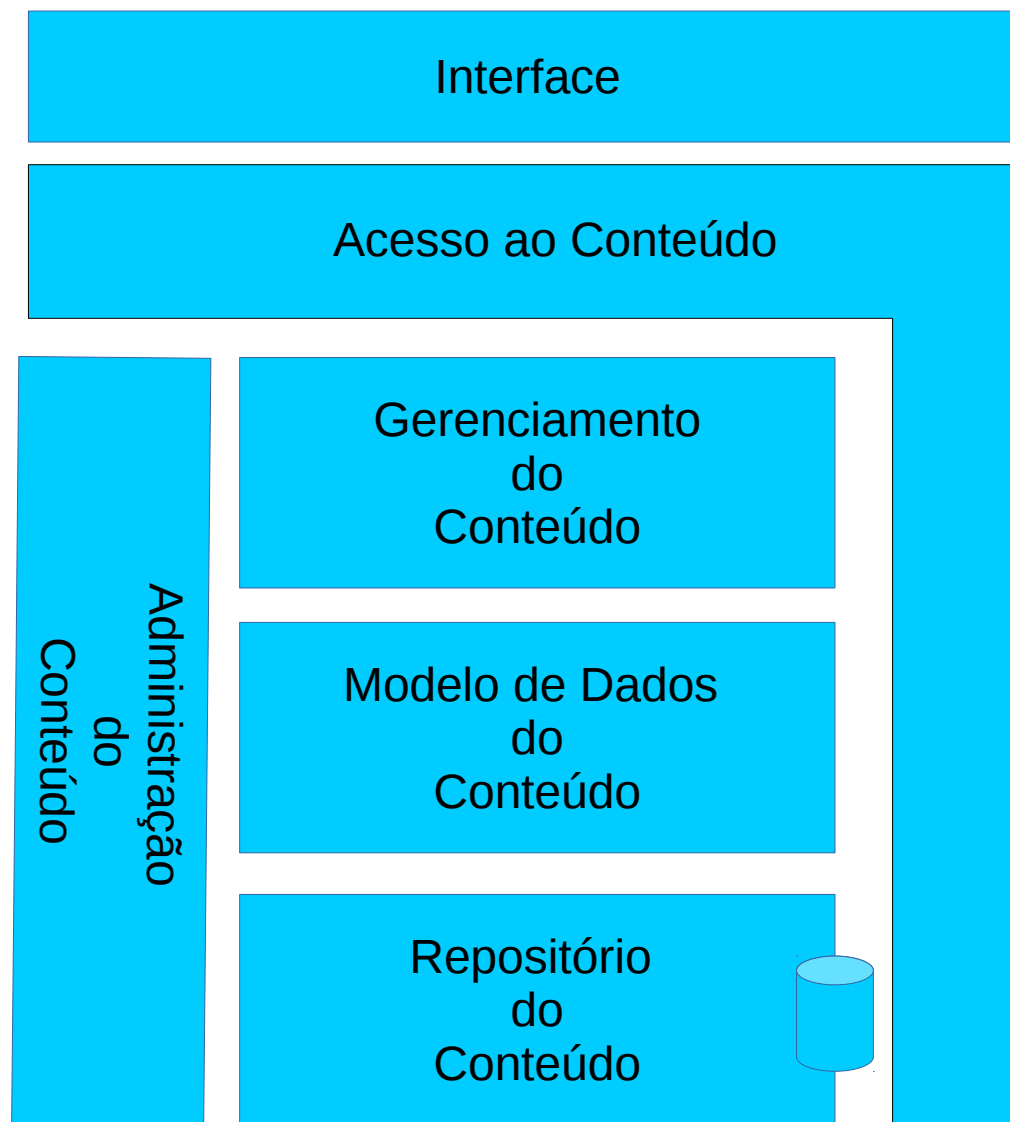
Content Management Systems CMS

- CMS tradicionais
 - Arquitetura em 3 camadas (3-tier)



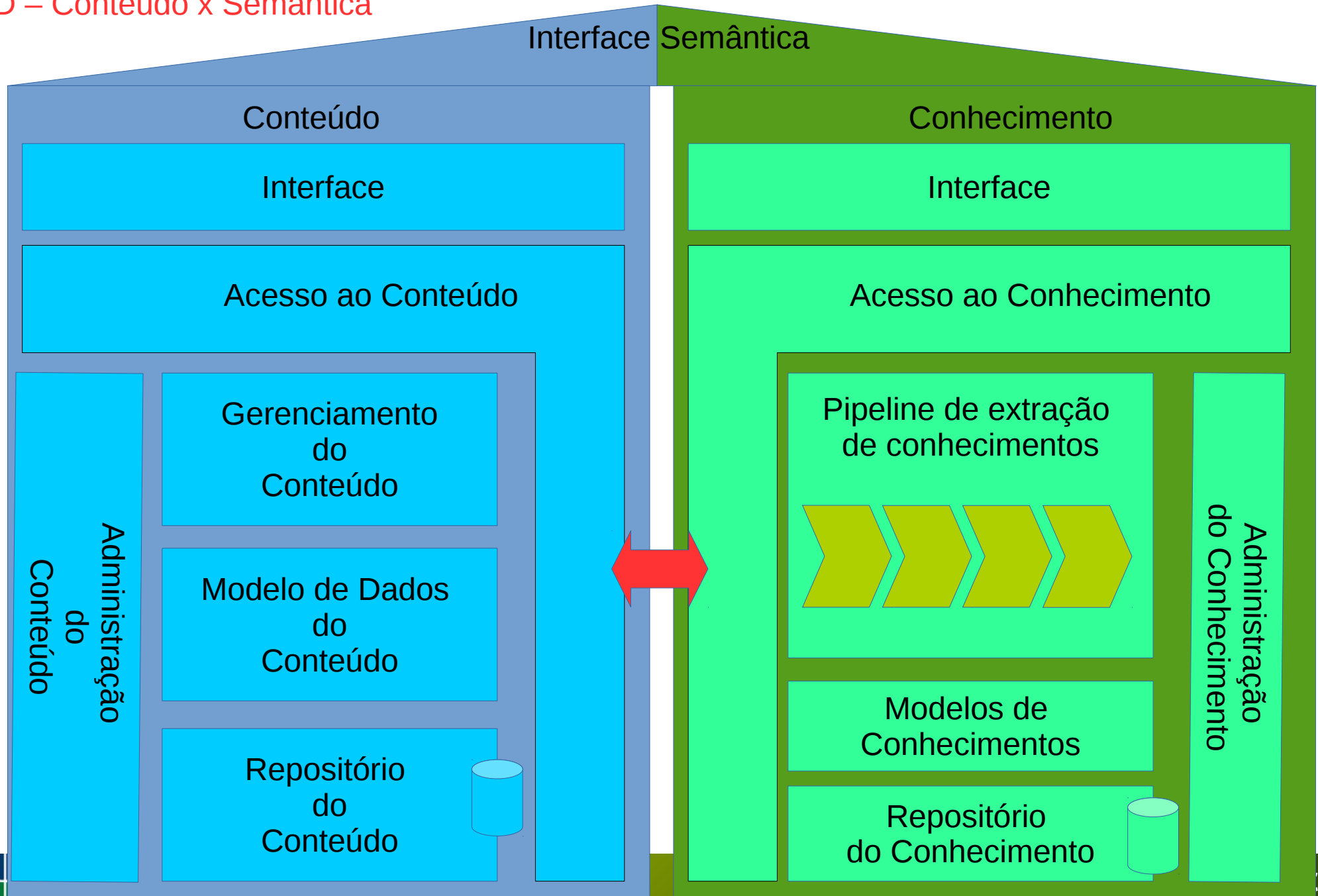
Arquitetura do Servidor CMS

1D - Conteúdo



Arquitetura de Referência Servidor CMS Semântico

2D – Conteúdo x Semântica



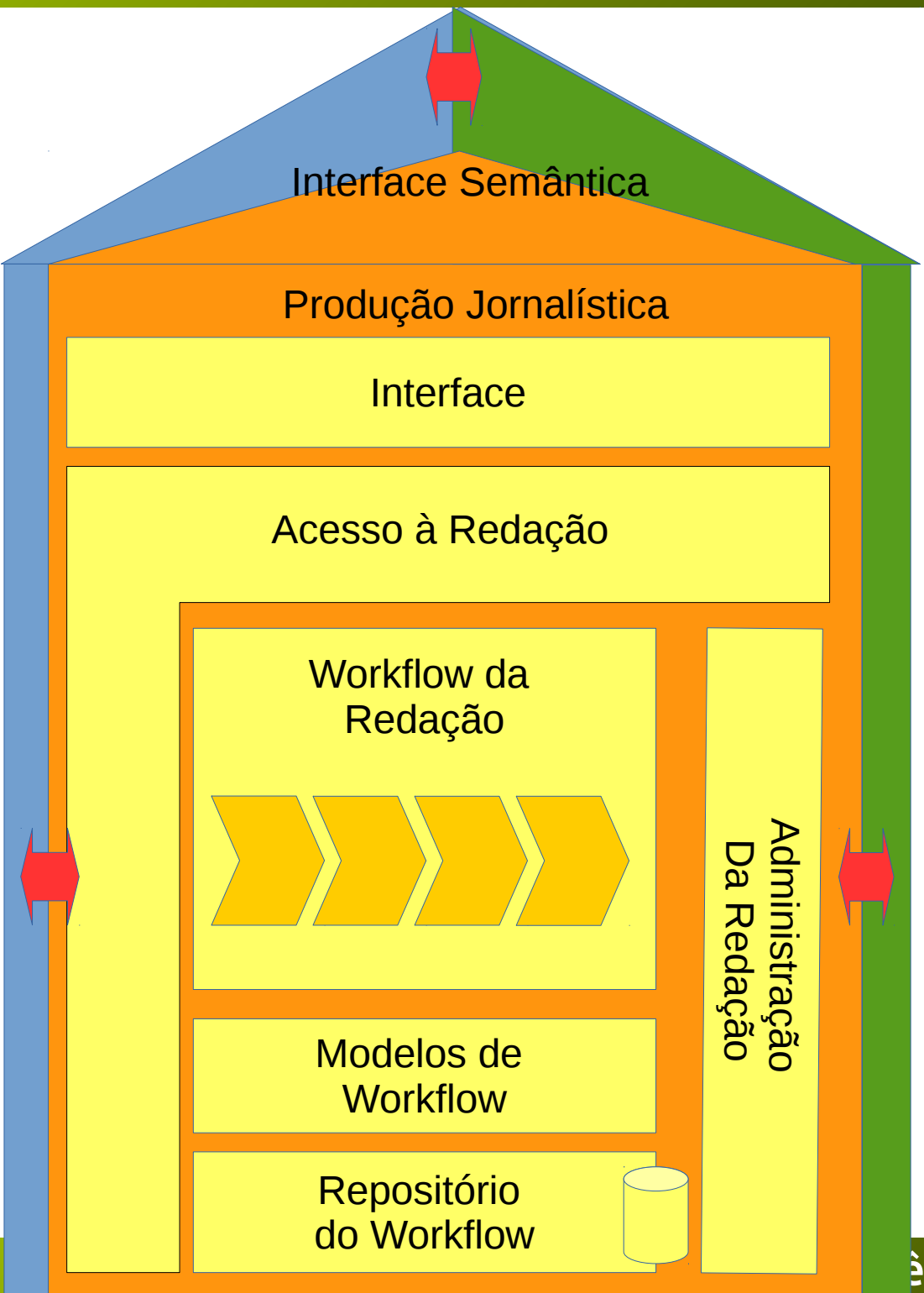
Pilha Jornalismo 3.0

- Adicionar capacidades C⁴
 - Cooperação
 - Coodenação
 - Comunicação (interna)
 - Conexão (comunicação externa)
- Workflow dinâmico



Arquitetura de Referência Servidor Produção Jornalística Semântico

ência da Computação



3D
Conteúdo
X
Semântica
X
Produção

Produção
=
Processo
=
Workflow
=
Qualidade
+
Confiabilidade

Workflow da Redação

- Atores
 - Chefe da Redação
 - Jornalistas
 - Assessor de Imprensa
 - ...
- Atividades
 - Definição de Pauta
 - Controle da qualidade
 - Revisão
 - Gramatical/Sintática
 - Semântica
- Formas de edição da Matéria
 - Processos
 - Individual
 - Colaborativo
 - Competitivo
 - Estados da Matéria
 - Iniciando
 - Em edição
 - Revisado
 - Arquivado
 - Publicado

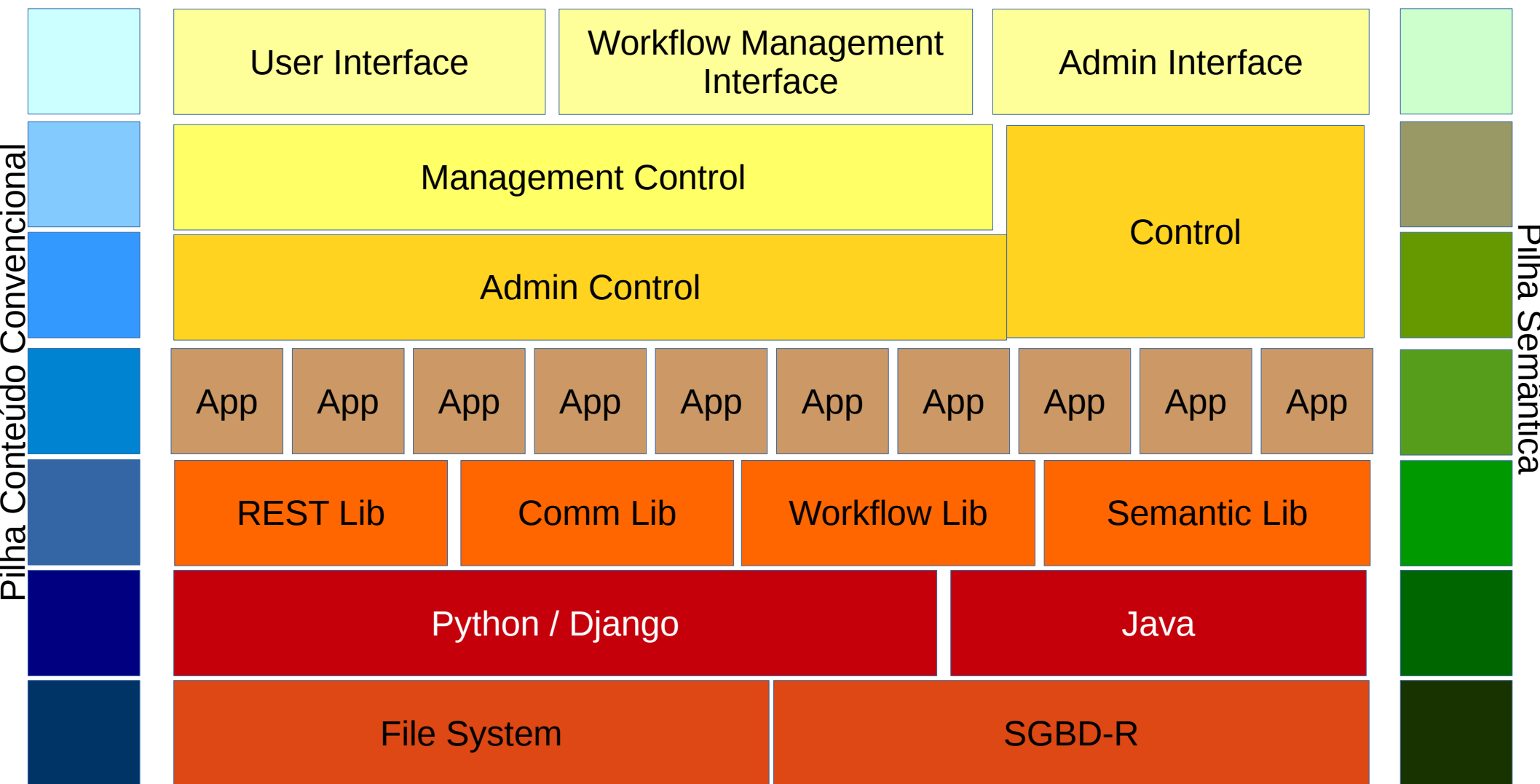


Workflow da Redação

- Cooperação
 - Na elaboração de artigo
 - Na elaboração de edição de jornal
- Coordenação
 - BPM de produção jornalística relaciona atores e tarefas para gerar qualidade e confiabilidade
 - Evitar pós-verdade
- Comunicação (Interna)
 - Dentro da própria aplicação para coordenação de atores e atividades
- Conexão(externa)
 - Alertas replicando comunicação interna para e-mail e App de Com Móveis (Telegram, SMS, WhatsApp)



Arquitetura do Framework de Produção Jornalística 3.0

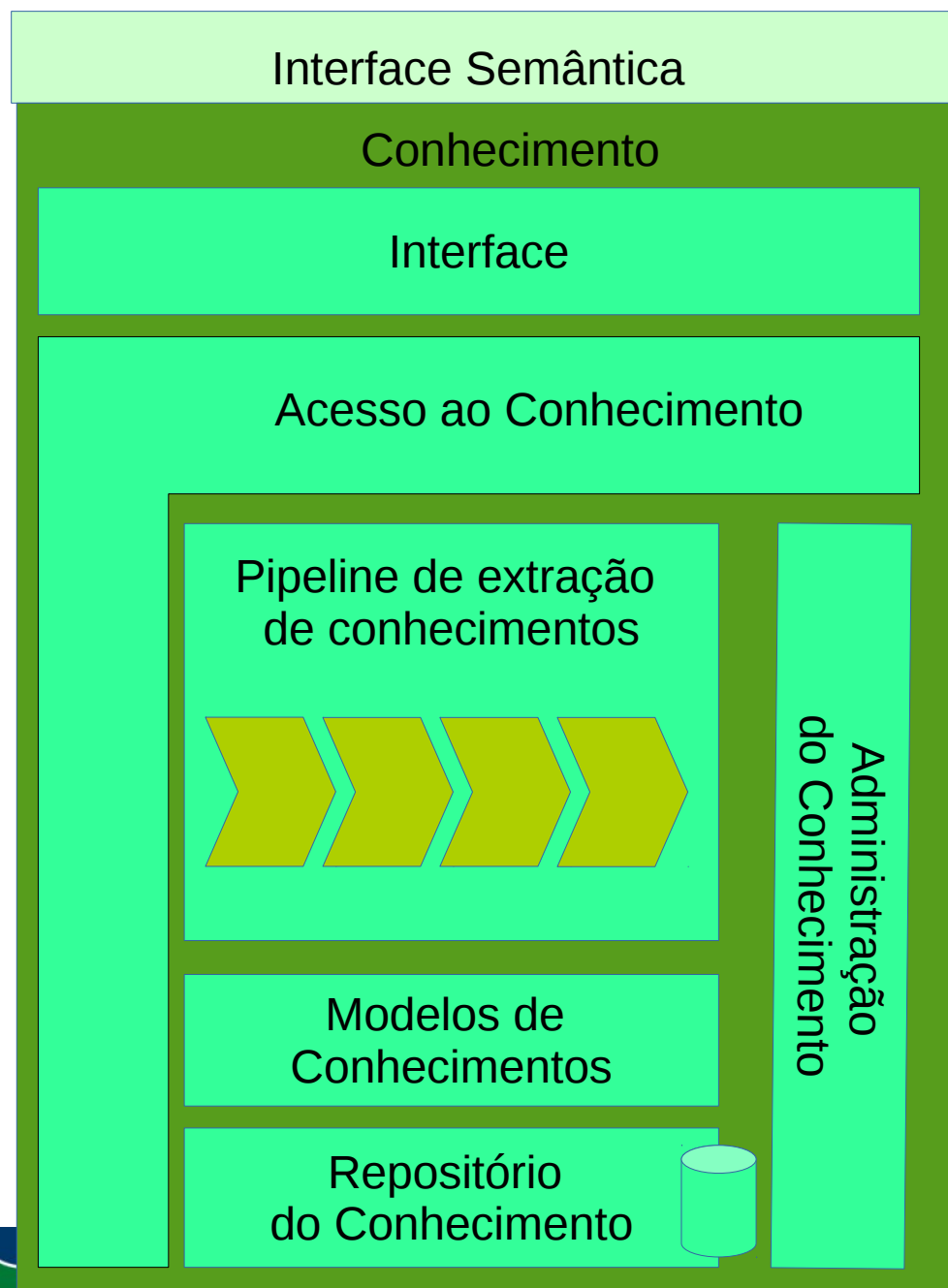


CMS Semântico p/ Jornalismo

- Inicialmente visão 2D
 - Conteúdo – Edição do Texto Matéria
 - Semântica – Edição do Conhecimento



Arquitetura de Referência Servidor CMS Semântico

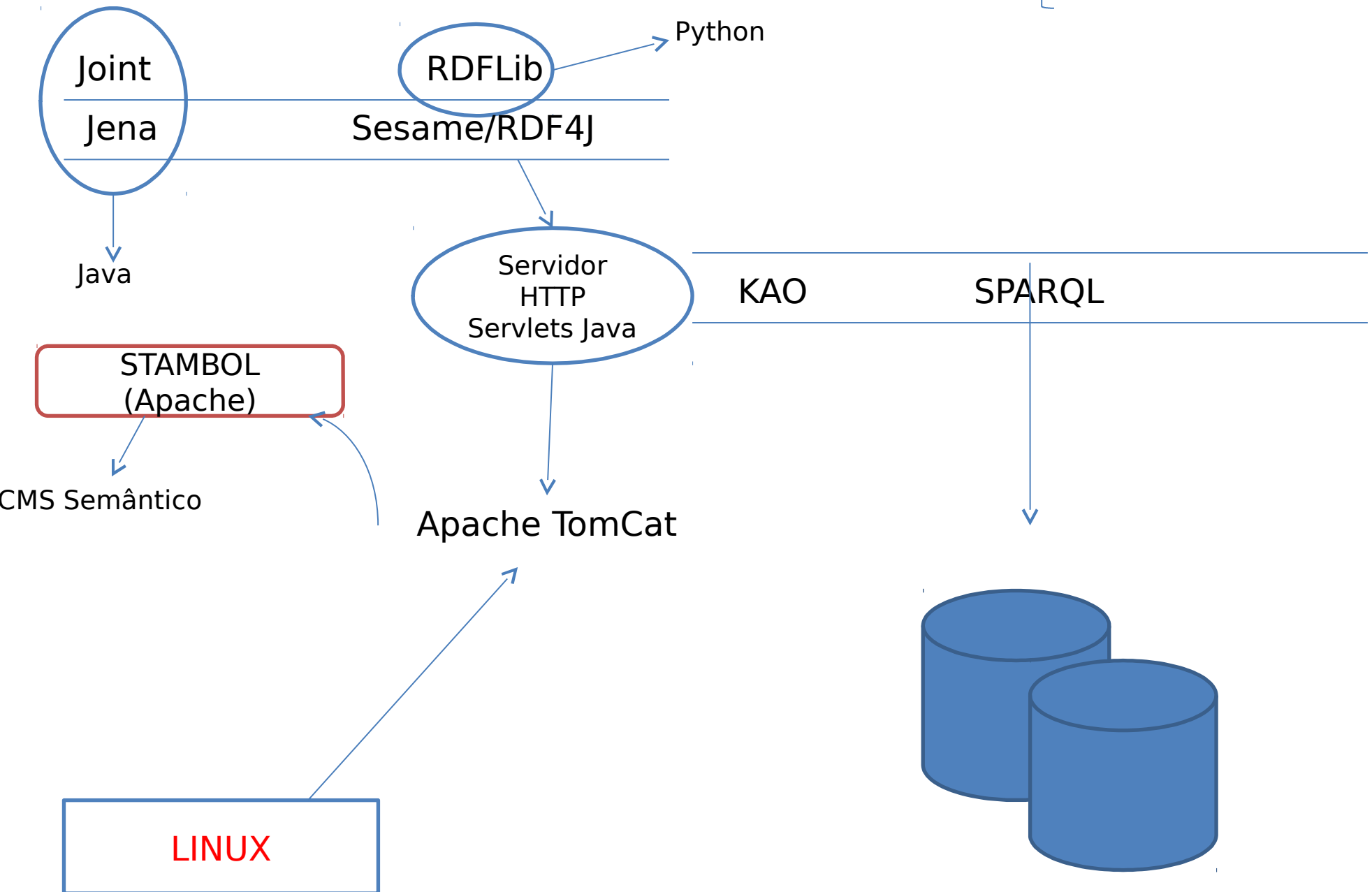


Interactive Knowledge Stack (IKS) Prova de Conceito

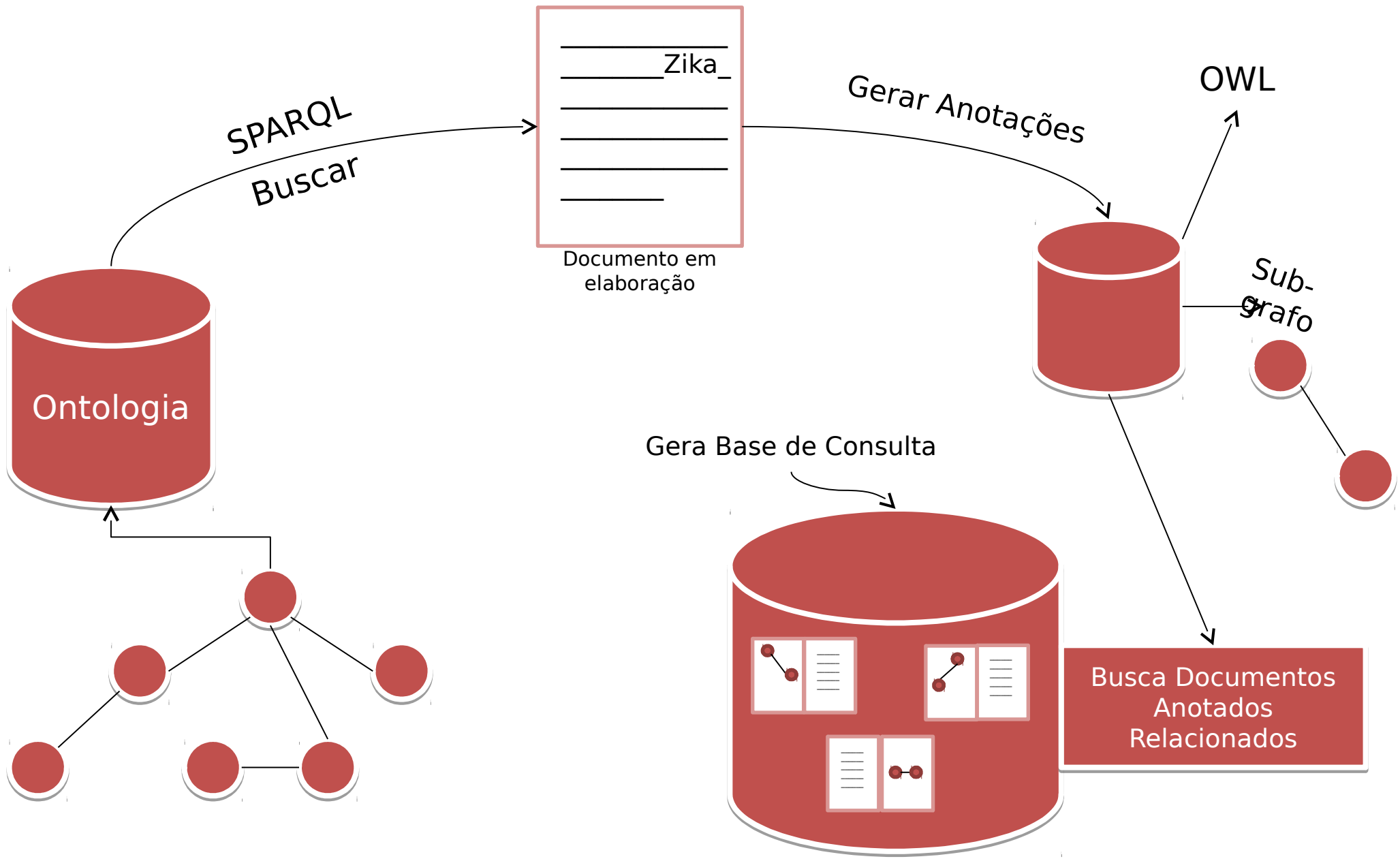


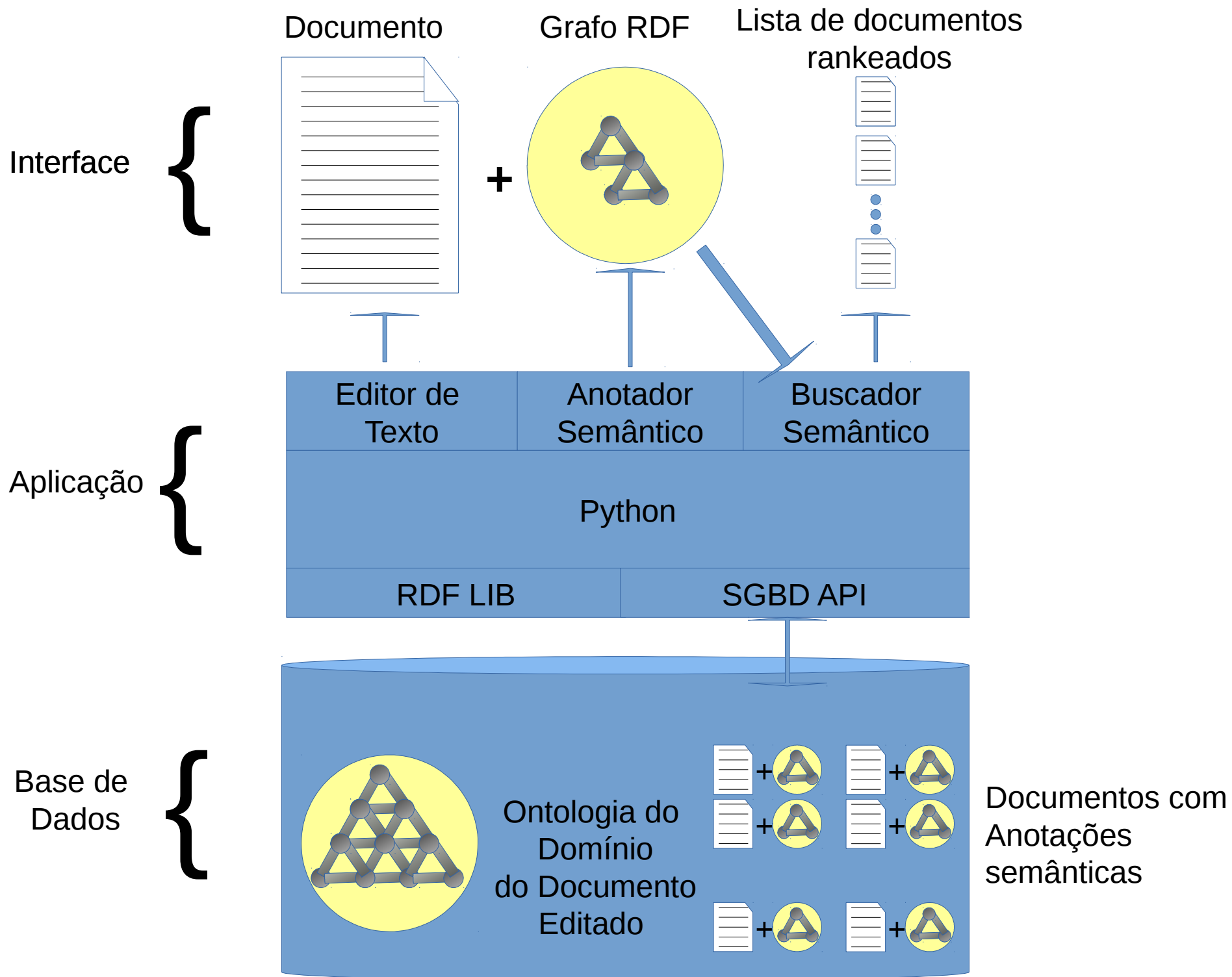
MODELO GERAL

SPARQL End Point { DBpedia
...



TRATAMENTO DE DOCUMENTOS





Anotações Semânticas

- Palavras-Chave
 - Zika
 - Vírus
 - Doença
 - Microcefalia
 - Vacina
- Agrego conhecimento a palavras chaves
 - Zika é um vírus
 - Zika é uma doença
 - Zika causa Microcefalia
 - Vacina para Zika

É o suficiente?



Interface Metáfora Página

Edição da Anotação
Semântica

Impeachment da Zica

Nova vacina promete erradicar...

Edição do Texto

- Zika é um vírus
- Zika é uma doença
- Zika causa Microcefalia
- Vacina para Zika

Zika



Máquina
de Busca

[HTML] Zika virus outbreak on Yap Island, federated states of Micronesia

[HTML] Ophthalmological findings in infants with microcephaly and presumable intra-uterus Zika virus infection

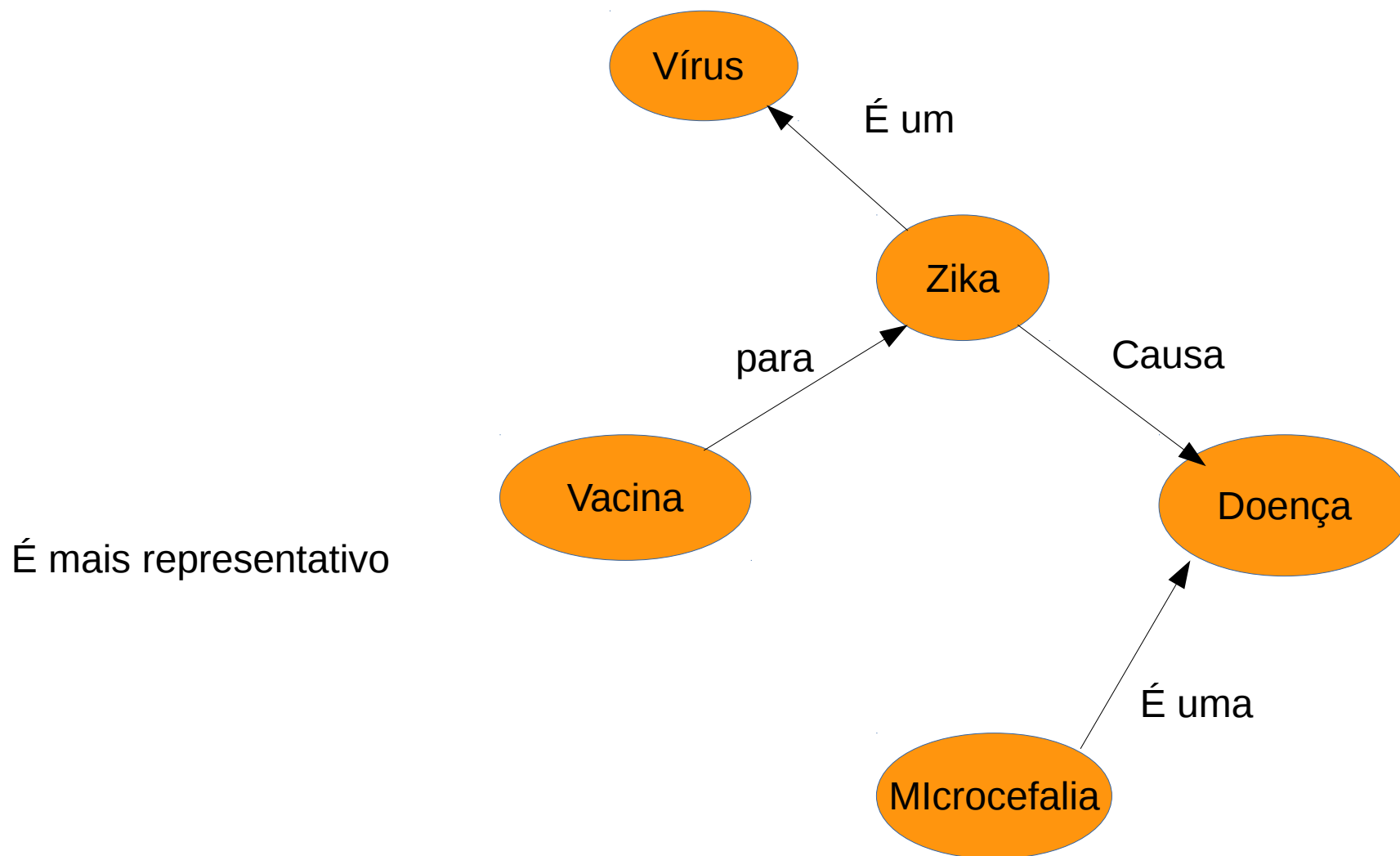
[PDF] Probable non-vector-borne transmission of Zika virus, Colorado, USA

Zika virus (I). Isolations and serological Specificity

[PDF] Potential sexual transmission of Zika virus

Resultados

Anotações Semânticas

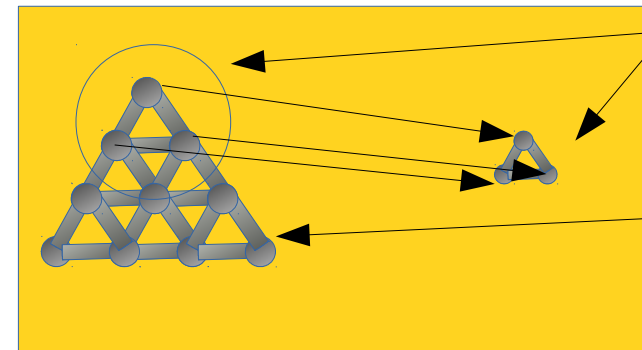


Então Interface Muda Metáfora muda?

Impeachment da Zica

Nova vacina promete erradicar...

Edição do Texto



Edição da Anotação Semântica

Fish Eye

Ontologia da Zica



Máquina de Busca

[HTML] Zika virus outbreak on Yap Island, federated states of Micronesia

[HTML] Ophthalmological findings in infants with microcephaly and presumable intra-uterus Zika virus infection

[PDF] Probable non-vector-borne transmission of Zika virus, Colorado, USA

Zika virus (I). Isolations and serological Specificity

[PDF] Potential sexual transmission of Zika virus

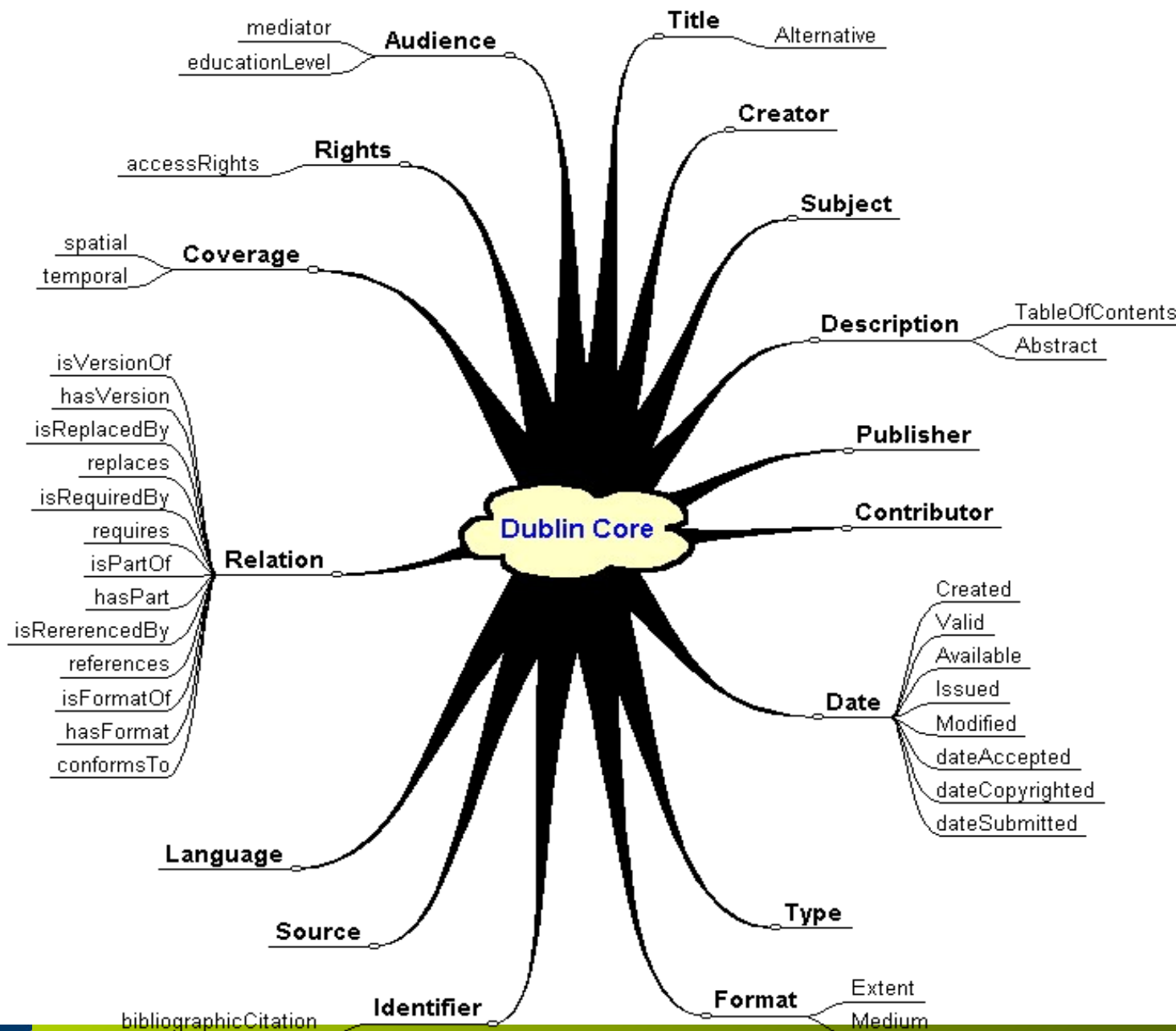
Resultados

2º Problema



- Realizo a busca segundo a anotação semântica gerado online
- Quero conteúdo que me ajude a gerar o texto e agregue “valor jornalístico” à notícia
 - Buscar só por Zika ou <s, p, o> diretamente relacionados vai trazer conteúdo com valor jornalístico?
 - Usar só ontologia da Zika não oferece bons resultados
 - Que outras ontologias usar para agregar valor?

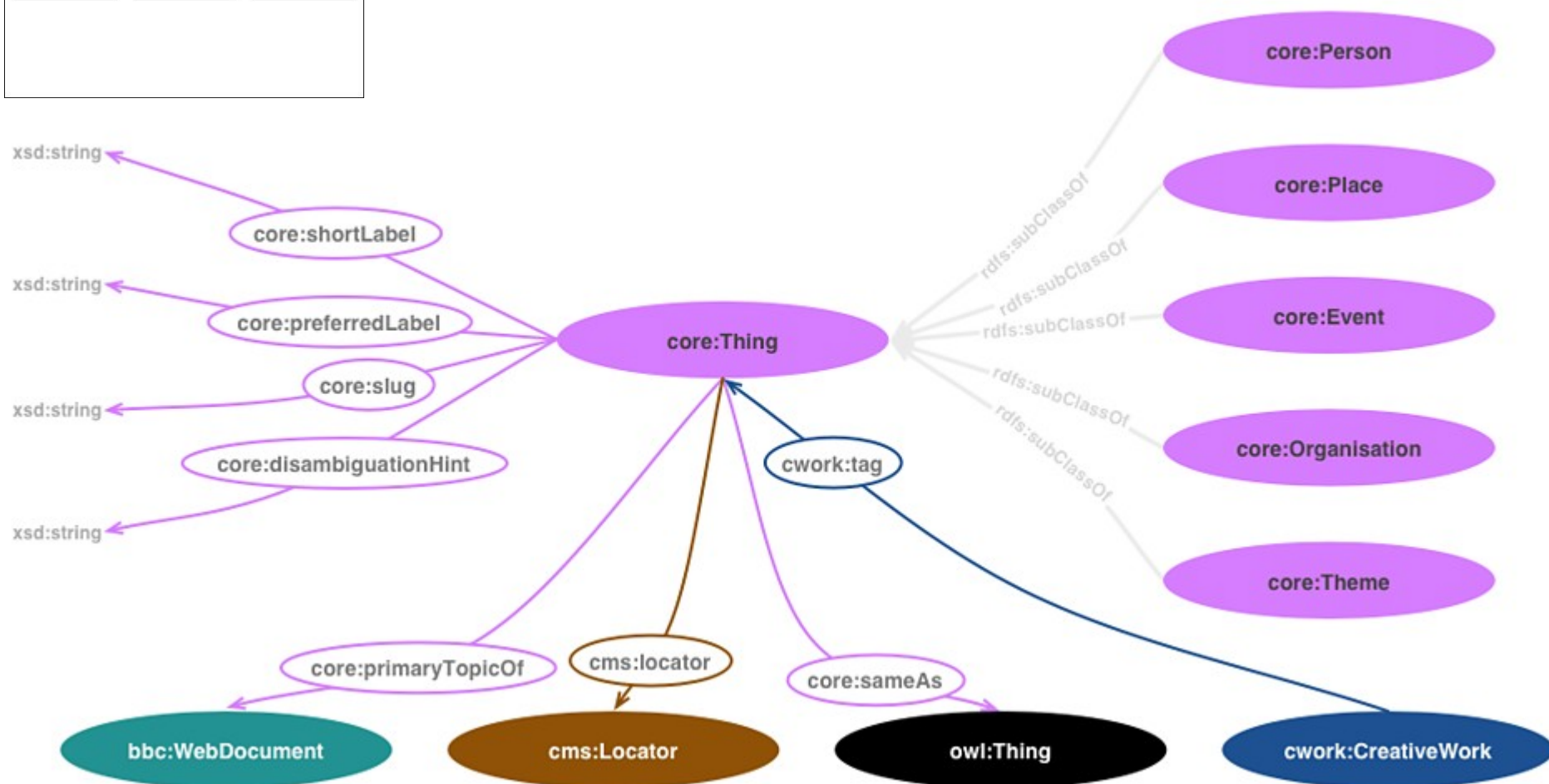
2º Problema



Resolve ontologia
da matéria



2º Problema



Resolve problema de classificação dos documentos.
Será que me ajuda a recuperar os documentos pelo valor da notícia?

O que compõe o “valor notícia”?

- valor subjetivo que determina a importância que um fato ou acontecimento tem para ser noticiado
- De acordo com o impacto (Wikipedia)
 - Amplitude - quanto maior o número de pessoas envolvidas...
 - Frequência - quanto menos frequente ...
 - Negatividade – más notícias vendem mais
 - Caráter inesperado - evento totalmente inesperado terá mais impacto do que um evento agendado e previsto
 - Clareza - eventos cujas implicações sejam claras vendem mais jornais do que aquelas que estão abertas a mais do que uma interpretação, ou cujo entendimento exija conhecimentos acerca dos antecedentes ou contexto desse mesmo evento
- De acordo com a empatia com a audiência
- De acordo com o pragmatismo da cobertura mediática

Fonte: 1965, Johan Galtung e Mari Holmboe Ruge



O que compõe o “valor notícia”?

- valor subjetivo que determina a importância que um fato ou acontecimento tem para ser noticiado
- De acordo com o impacto
- De acordo com a empatia com a audiência
- De acordo com o pragmatismo da cobertura mediática (Wikipedia)
 - Consonância - jornalistas têm esquemas mentais em que prevêem que determinado acontecimento pode vir a ocorrer. Previsão tem a ver com a experiência e rotina do jornalista. Se uma ocorrência corresponder às expectativas terá chances de publicação.
 - Continuidade: uma vez publicada, a notícia ganha uma certa inércia. Como a história já é pública existe maior clareza acerca da mesma. Isto cria um acompanhamento da notícia até que outras notícias mais importantes obriguem a deixar cair o assunto.
 - Composição: o arranjo das notícias por rubricas, seções ou cadernos deve ser equilibrado. Se um acontecimento internacional for importante terá de competir com o valor de outros acontecimentos internacionais. A importância não depende apenas do seu valor-notícia mas também do seu valor face a outras histórias.

Fonte: 1965, Johan Galtung e Mari Holmboe Ruge



O que compõe o “valor notícia”?

- valor subjetivo que determina a importância que um fato ou acontecimento tem para ser noticiado
- De acordo com o impacto
- De acordo com a empatia com a audiência (Wikipedia)
 - Personalização - as ocorrências que possam ser retratadas como ações de indivíduos atraem um maior interesse humano pela história relatada
 - Significado - este critério está relacionado com a proximidade geográfica e cultural que a ocorrência possa ou não ter para o leitor. Notícias sobre acontecimentos, pessoas e interesses mais próximos do leitor terão um maior significado para ele.
 - Referência a países de elite: notícias relacionadas com países mais poderosos têm maior destaque do que notícias relativas a países de menor expressão política e económica.
 - Referência a pessoas que integram a elite - histórias acerca de pessoas ricas, poderosas, influentes e famosas recebem uma maior cobertura noticiosa.
- De acordo com o pragmatismo da cobertura mediática

Fonte: 1965, Johan Galtung e Mari Holmboe Ruge



Anotação

- Resumindo
 - Ontologia do Assunto
 - Zika
 - Ontologia do Documento
 - DublinCore
 - Ontologia de Classificação do Documento
 - Para arquivamento e recuperação
 - Ontologia de Classificação Jornalística
 - Para valoração – adequado para ajudar a montar a pauta
 - Pragmática - Falta algo para colocar tudo em um contexto, em uma situação

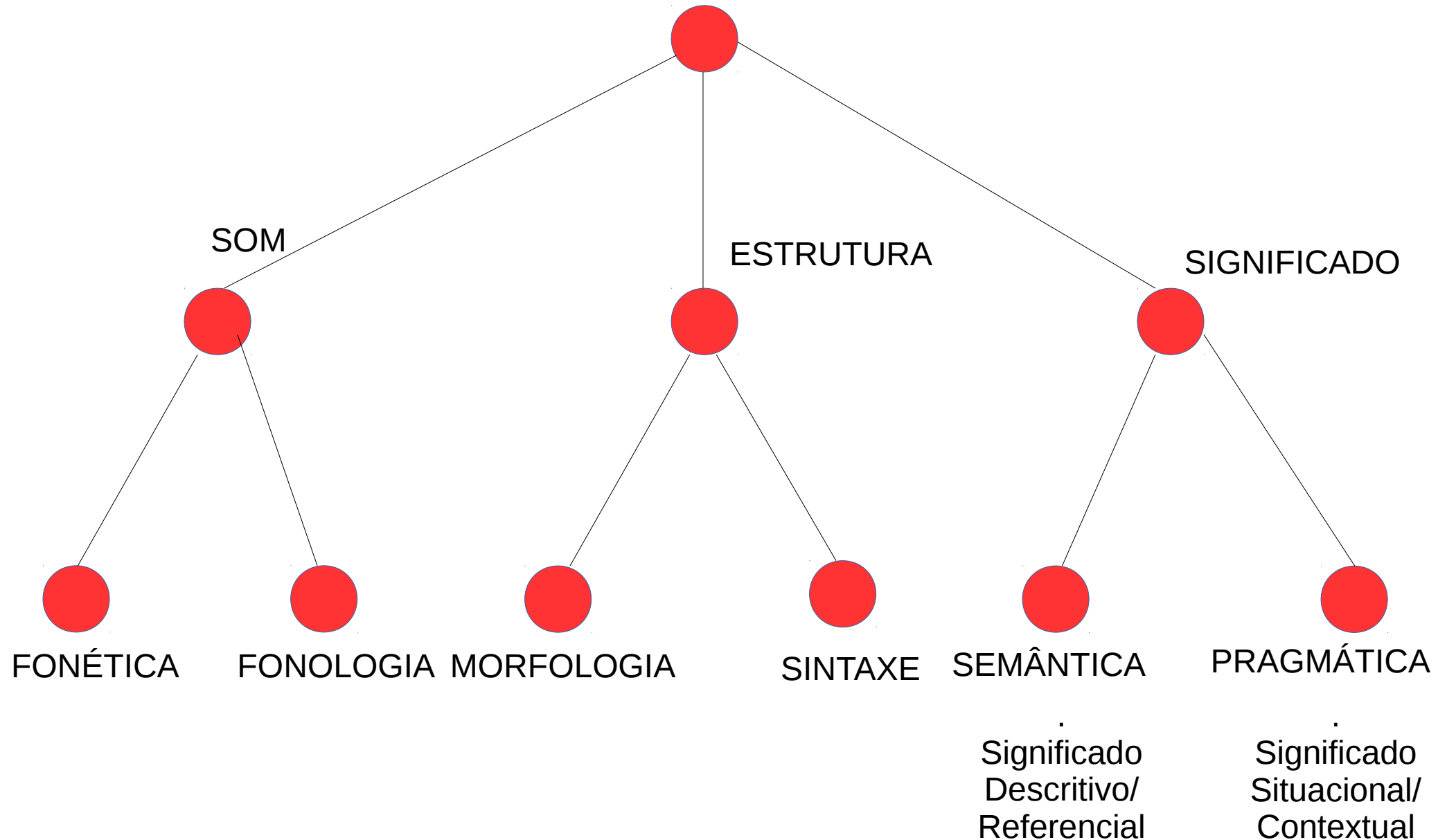


Pragmática

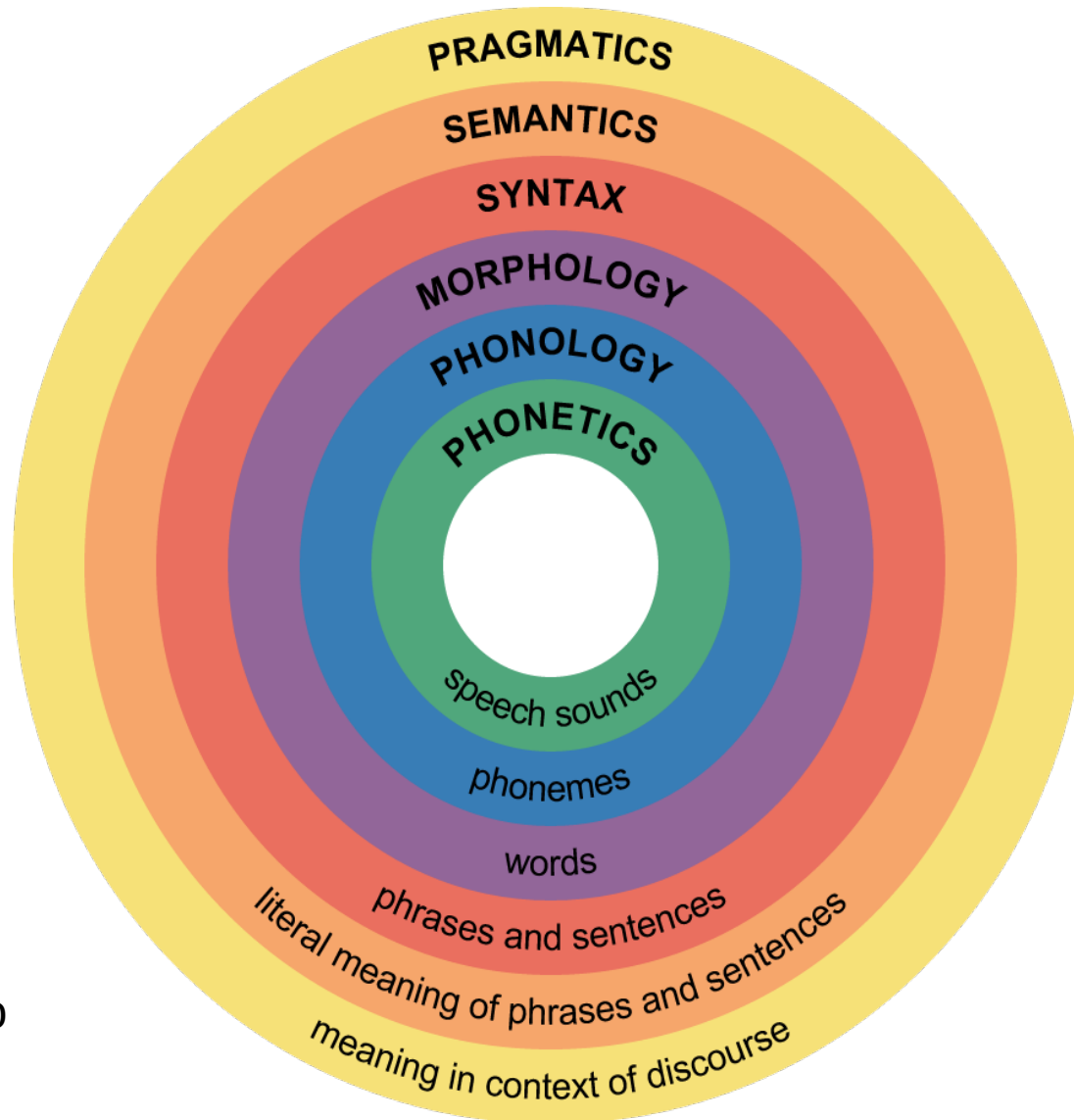
- Pragmática é o ramo da linguística que estuda a linguagem no contexto de seu uso na comunicação.
- A pragmática estuda os significados linguísticos determinados não exclusivamente pela semântica proposicional ou frásica, mas aqueles que se deduzem a partir de um contexto extralinguístico: discursivo, situacional, etc.



Linguística = Sistemas de Linguagem



Camadas de Descrição Linguística



Léxico=vocabulário



O que orienta o estilo da notícia

- Linha Editorial do veículo
- Ideologia
- Contexto social/político/econômico do momento
- Cultura Local
- Etc. . .



3º Problema

- Recuperar a informação de acordo com um contexto / situação (pragmática)
- O valor da informação depende da pragmática
- O que define um contexto / situação
 - Ontologia?
 - Valores?
 - Princípios
- Isto pode ser expressado por meio de uma metalinguagem?



MLEI

- MetaLinguagem da Economia da Informação
- Uma memória digital participativa comum ao conjunto da humanidade está em vias de constituição. Mas no início do séc. XXI, a exploração dessa memória por todos e por cada um é limitada por problemas de opacidade semântica, de incompatibilidade dos sistemas de classificação e de fragmentação lingüística e cultural. Na ausência de modelos computáveis, nós não conseguimos automatizar a maior parte das operações cognitivas de análise, de filtragem, de síntese e de interconexão das informações que permitiriam utilizar vantajosamente a imensa massa de dados que se nos oferecem.
Pierre Levy
- Como expressar (linguagem) o que a gente quer no contexto de um CMS semântico jornalístico para obter os resultados certos?



4º Problema

Persistência da Anotação

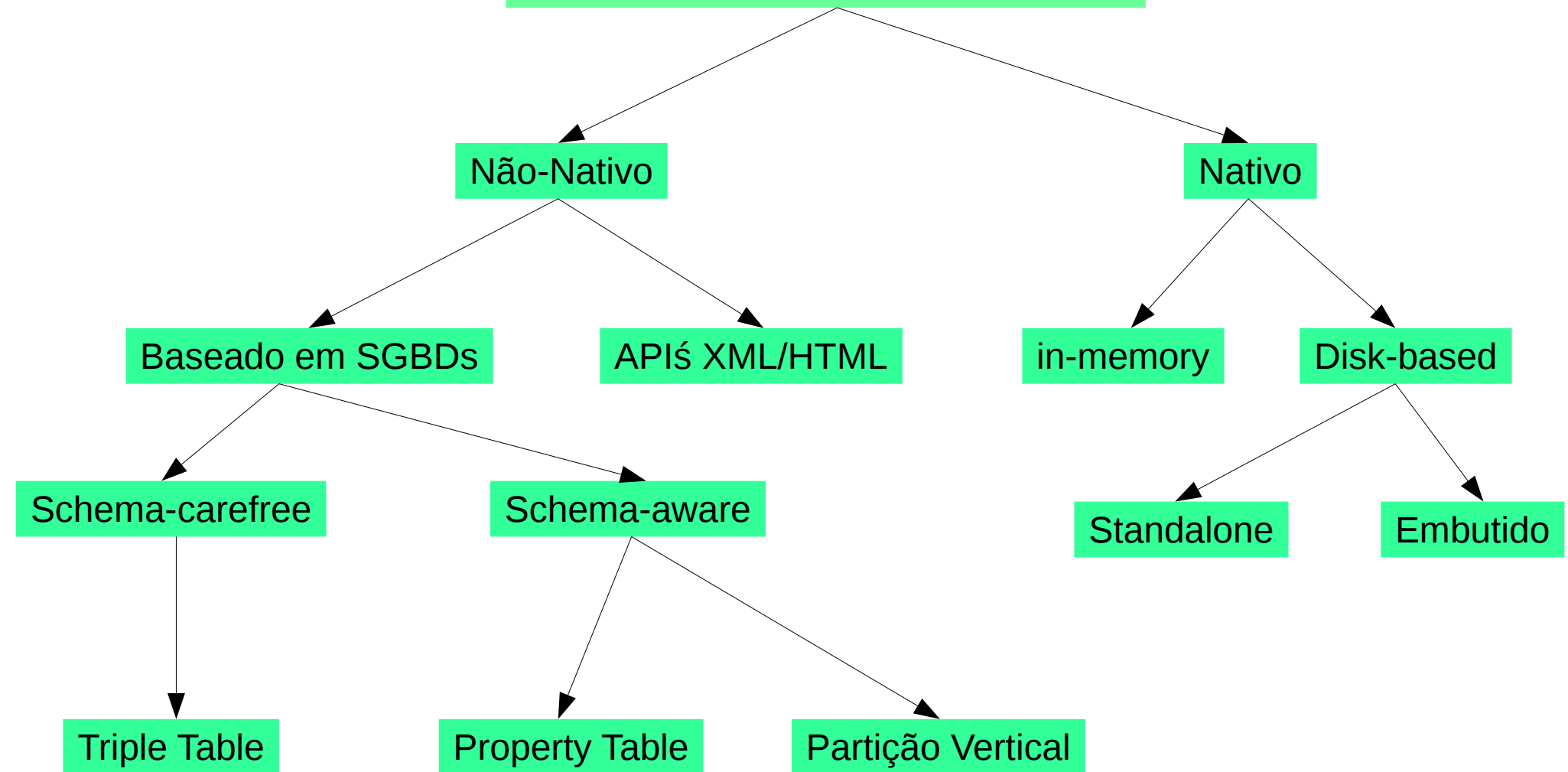
- Anotação RDF
- Como indexar Grafos RDF
- Características do BD
 - Mecanismo de armazenamento
 - Suporte a inferências
 - Suporte a atualizações
 - Escalabilidade
 - Distribuído x Centralizado

Fonte: Faye, D.C., Curé, O., Blin, G. A survey of RDF storage approaches. INRIA. 2012.



4º Problema

Técnicas de Armazenamento de RDF

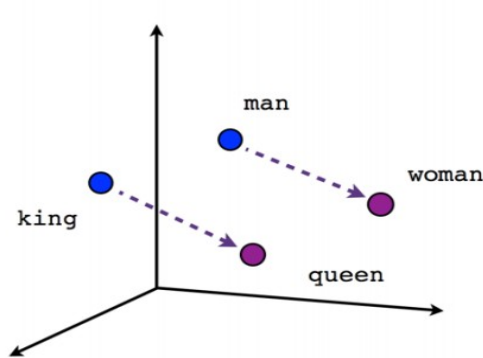


Fonte: Faye, D.C., Curé, O., Blin, G. A survey of RDF storage approaches. INRIA. 2012.

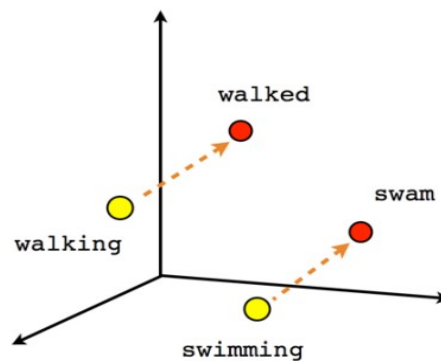


Como realizar inferências?

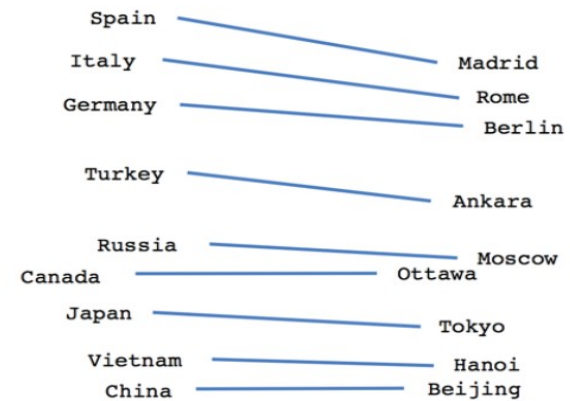
- SPARQL?
- DNN
 - Vetorização Semântica



Male-Female



Verb tense



Country-Capital

- Redução dimensional
- Exemplo
 - King – man = queen

Dados Abertos Linkados

- Dados Estruturados
 - SGBD
 - CSV
- Big Data
- Junção de diferentes bancos
 - Ontologias
 - Inferências
 - Indicadores relevantes
- Infográficos



Interface Semântica?

- O que é?
- Novo paradigma?
- Qual é a meta-função?
 - Da presença?



MLEI



Referências

- Fabian Christ, Benjamin Nagel. A reference Architecture for Semantic Content Management
- Pierre Lévy. A Esfera Semântica: tomo I, computação, cognição, economia da informação

