

# Linguatca e Processamento de Texto Livre na Área da Saúde: Alguns Comentários e Sugestões

Liliana Ferreira e António Teixeira

Encontro 10 anos: Linguatca  
Setembro 2008  
Aveiro

- 1 O nosso trabalho
  - Introdução
  - MedAlert
  - Participação no II HAREM
  
- 2 As nossas sugestões

# Motivação: Rede Telemática da Saúde (RTS)

Desenvolvida e implementada pelo Consórcio RTS, que integra:

- o Hospital Infante D. Pedro e o Hospital Distrital de Águeda;
- a Sub-Região de Saúde de Aveiro;
- a Universidade de Aveiro como parceiro tecnológico.

A RTS implementa um "Processo Clínico Electrónico Regional" resumido, que agrega informação clínica do utente, como por exemplo:

- Cartas de alta de episódios;
- Boletins de análises clínicas;
- Relatórios de exames de imagiologia.

Sumário do episódio
Carta de Alta Administrativa
Não disponível
Carta de alta
Diagnósticos
Diário
Prescrição
Transferências
Sessões de Hospital de Dia

© 2004/2008 Rede Telemática da Saúde

# MedAlert - Sistema de Processamento de Linguagem Médica

## Objectivo:

Utilização de técnicas de extracção automática de informação aplicadas a textos médicos, de modo a inferir de forma automática irregularidades/dúvidas suscitadas pelas decisões tomadas pelos profissionais de saúde.

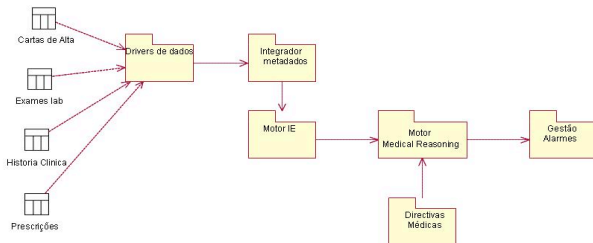


Figura: Arquitectura MedAlert

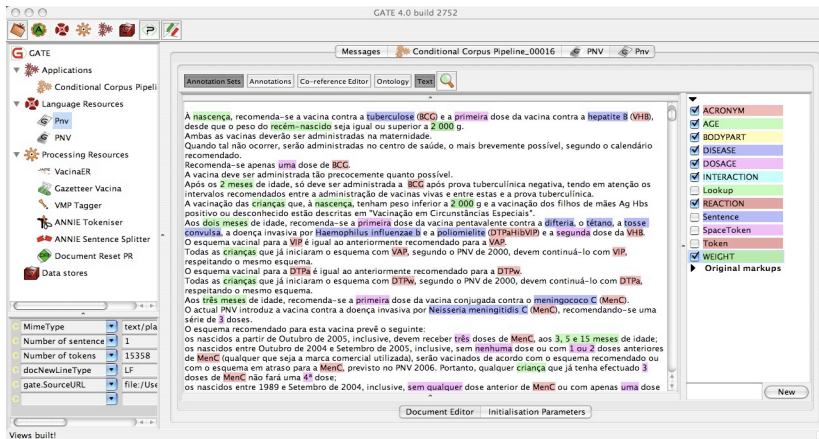


Figura: Interface Gráfica do GATE com os diferentes Recursos de Processamento.

**Tabela:** Resultados da tarefa de Extração de Informação para cada uma das entidades definidas.

	DOENCA	ACRONIMO	IDADE	PARTE_CORPO	DOSE	INTERACCAO	REACCAO	PESO
Saídas correctas	225	294	181	14	90	10	156	8
Parcialmente correctas	0	0	1	0	3	0	0	0
Em falta	0	0	6	0	0	0	0	0
<b>Total</b>	<b>225</b>	<b>294</b>	<b>188</b>	<b>14</b>	<b>93</b>	<b>10</b>	<b>156</b>	<b>8</b>
Abrangência	1,00	1,00	0,96	1,00	0,97	1,00	1,00	1,00
Precisão	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
<b>Medida F</b>	<b>1,00</b>	<b>1,00</b>	<b>0,98</b>	<b>1,00</b>	<b>0,98</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>

**Tabela:** Resultados gerais da tarefa de Extração de Informação.

Saídas Correctas	978
Parcialmente correctas	4
Em falta	6
Abrangência	0,991
Precisão	1,000
Medida F	0,995

# Representação Conceptual de Directivas

## Plano Nacional de Vacinação

### Frase

*À nascença, recomenda-se a vacina contra a tuberculose (BCG) e a primeira dose da vacina contra a hepatite B (VHB), desde que o peso do recém-nascido seja igual ou superior a 2 000g.*

Introdução: os regimes para todas as vacinas de administração orais ou cutâneas, serão abordadas nos capítulos seguintes e, eventualmente, em ilustrações complementares.

**Quadro 8. Vacinação neonatal. Crianças Disminuídas**

Vacina	Idade											
	0-15 dias	15 dias-1 mês	1-2 meses	2-3 meses	3-4 meses	4-5 meses	5-6 meses	6-11 meses	12 meses	12-18 meses	18-24 meses	24-36 meses
Tuberculose	BCG											
Hepatite B (VHB)												
Hepatite A (VHA)												
Difteria, Tétano e Coqueluche (DTP)												
Poliomielite												
Varicela												
Rotavírus												
Neumococo												
Meningococo												

© INSAPI, em colaboração com a OMS e a UNICEF. Última actualização em 2012. Para mais informações, consulte o site [www.dsa.gov.br](http://www.dsa.gov.br).

13

Circular Normativa Nº 0027 de 21/12/2009

**Vacinação no primeiro ano de vida**

A vacinação recomenda-se a vacina contra a tuberculose (BCG) e a primeira dose da vacina contra a hepatite B (VHB), desde que o peso do recém-nascido seja igual ou superior a 2 000 g.

Antes de aplicar qualquer vacina, o profissional de saúde deve verificar se a criança não possui nenhuma doença aguda ou febre. Quando se não estiverem presentes, a vacinação pode ser realizada. Quando houver alguma doença aguda ou febre, a vacinação deve ser adiada até o completo desaparecimento dos sintomas.

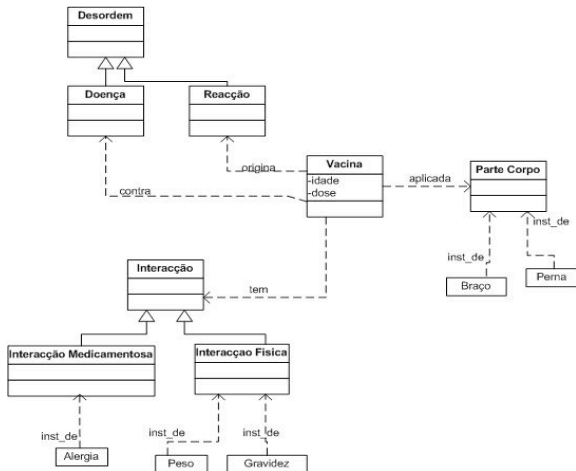
Recomenda-se apenas uma dose de BCG. A vacina deve ser administrada tão precocemente quanto possível. Após os 2 meses de idade, só deve ser administrada a BCG após prova tuberculínica negativa, sendo em situações de interesse.

### Frase com Padrões Linguísticos

*[idade] [vacina] {contra} [doença] ([acronimo]) E [dose] [vacina] {contra} [doença] ([acronimo]), {restriccao} [peso].*

# Plano Nacional de Vacinação

## Vista Parcial da Ontologia

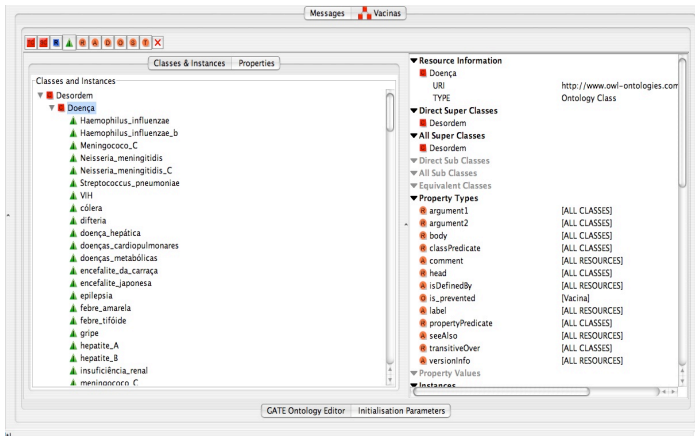




# Plano Nacional de Vacinação

## Introdução de instâncias

- A informação extraída do PNV automaticamente adicionada à ontologia:
  - associação entre a classe e a entidade identificada.



# Plano Nacional de Vacinação

## Relacionamentos entre instâncias

- Procura de Padrões Frequentes, usando Regras de Associação.

f1: {idade, acrónimo}

f2: {idade, idade, peso}

f3: {idade, dose, doença, acrónimo, dose, acrónimo}

...

- As Regras de Associação descrevem padrões que, neste caso, indicam entidades que são mencionadas em conjunto:
  - a *doença*  $\Rightarrow$  *vacina* (80%) indica que 4/5 vezes que uma doença é mencionada (*Tuberculose*) é seguida pela referência a uma vacina (*BCG*).
  - b O relacionamento *BCG* 'contra' *Tuberculose* pode ser inferido.

## Resultados

1	DOENÇA ⇒ ACRÓNIMO	79.68%
2	IDADE ⇒ ACRÓNIMO	69.73%
3	ACRÓNIMO ⇒ DOSE	20.83%

## Resultados

- |   |                   |        |
|---|-------------------|--------|
| 1 | DOENÇA ⇒ ACRÓNIMO | 79.68% |
| 2 | IDADE ⇒ ACRÓNIMO  | 69.73% |
| 3 | ACRÓNIMO ⇒ DOSE   | 20.83% |

## Trabalho Futuro

- 1 Explorar outras técnicas de extracção de relações semânticas, seguindo, por exemplo, o trabalho desenvolvido para o PAPEL.
- 2 ...

# REMMA - Reconhecimento de Entidades Mencionadas do MedAlert

- Desenvolvido em UIMA - *Unstructured Information Management Architecture*
- Principais recursos:



*almanaques*

+



*regras*

+



*wikipedia*

- A wikipedia disponibiliza os conteúdos da base de dados para cada língua em XML e os ficheiros SQL necessários para construir a BD.

*Dumps from any Wikimedia Foundation project:*  
*<http://download.wikimedia.org/>*

- Foi utilizada a versão portuguesa de Fevereiro de 2008, contendo cerca de 1 290 836 páginas.

### Exemplo

A "Universidade de Aveiro (UA)" é uma *universidade* pública portuguesa localizada em Aveiro.

- 1 O nosso trabalho
  - Introdução
  - MedAlert
  - Participação no II HAREM

- 2 As nossas sugestões

- 1 Produção de recursos tendo por base ambientes já existentes (ex.: GATE, UIMA):
  - a facilidade na adaptação a diferentes técnicas e áreas;
  - b maior estruturação dos recursos;
  - c independência e escalabilidade;
  - d maior divulgação da área (ex.: *aplicação ao ensino*.)
  
- 2 Inventariação dos módulos e sistemas existentes e disponíveis;



- Criação de directivas genéricas de eventos de avaliação:
  - a facilidade na criação de eventos de avaliação conjunta em domínios mais específicos
  - b ex.: Medicina;
  
- Exemplos:
  - a Construção de *Colecções Douradas*;
  - b métricas de avaliação;
  - c sítio electrónico.
  - d ...

# Linguateca e Processamento de Texto Livre na Área da Saúde: Alguns Comentários e Sugestões

Liliana Ferreira e António Teixeira

Encontro 10 anos: Linguateca  
Setembro 2008  
Aveiro

*Fim! Obrigado.*