

# Propostas de Tese para o Mestrado em Engenharia Informática (MEI)

Grupo de Processamento de Linguagens  
DI/CCTC  
Universidade do Minho

(Pedro Rangel Henriques)  
ano lectivo 14/15

## 1 CIDOC-CRM para o Museu da Pessoa

*Supervisor: Pedro Rangel Henriques + Ricardo Martini*

Resumo:

No contexto do doutoramento do Ricardo Martini e de um projeto integrado do Perli de Processamento de Linguagens e Conhecimento desenvolveu-se um estudo aprofundado da norma CIDOC-CRM para construção de Ontologias na área dos museus. O estudo foi aplicado a um fundo documental do Museu da Emigração.

Neste projeto de tese de mestrado pretende-se aplicar a mesma abordagem a fundos documentais variados que o nosso grupo possui relativos ao Museu virtual da Pessoa com o objetivo de criar uma ontologia para a organização dessa informação, seu armazenamento em triplos RDF num triple-store e construção de um sítio web para interrogação desse repositório.

## 2 Sistema de Querying and Answering para a área das Linguagens de Programação

*Supervisor: Pedro Rangel Henriques + Maria João Varanda Pereira*

Resumo:

A tecnologia de sistemas Q&A (Querying and Answering systems) está cada vez mais em voga para suportar a construções de portais web de atendimento automático ao cliente final.

A ideia fundamental é permitir que o interessado coloque ao sistema questões formuladas em formato livre em língua natural (neste caso em Português) e o sistema, após reconhecer a pergunta busque no seu repositório de informação a resposta mais adequada e a apresente, também em língua natural, ao utilizador.

Neste projeto pretendemos estudar a área, ou seja investigar os recursos e as técnicas mais vulgarmente usadas pela comunidade científica para resolver este problema, e depois construir um sistema Q&A para o domínio da Programação — assim, uma ou mais, Linguagens de Programação, LPs, a escolher com o candidato (p.ex. C, Java, Perl, Python, etc.), devem ser consideradas na construção do repositório de conhecimento. O utilizador final poderá então colocar ao sistema questões diversas ligadas com a sintaxe e a semântica da LP em causa e o sistema fornecer-lhe-á a resposta adequada à dúvida colocada (ou mais do que 1 resposta, se tal for o caso). Se necessário o sistema deve ser capaz de entrar em diálogo com o utilizador, permitindo-lhe ir colocando questões sucessivas de maneira a ir-se aproximando progressivamente da clarificação total da sua dúvida.

## 3 Aplicação de projecções de código para uma compreensão faseada

*Supervisor: Pedro Rangel Henriques + Maria João Varanda Pereira*

Resumo:

O termo "projeção de código" é normalmente usado no contexto das ferramentas que se baseiam nas árvores de sintaxe abstracta de programas. Neste tipo de sistemas é possível gerar diferentes vistas sobre um mesmo programa dependendo da projeção aplicada. Vários IDE's já proporcionam funcionalidades semelhantes e o primeiro objectivo deste trabalho será fazer uma resenha dos diversos tipos de projeções que poderão ser úteis e que, por ventura, já estarão implementados em editores e outro tipo de ferramentas. No artigo *Source Code Projections in Integrated Development Environments*, de Jan Juhár e Liberios Vokorokos, são considerados quatro tipos de projeções: *projeções textuais* que são aplicadas sobre a notação textual dos programas (funcionalidade incluída normalmente em editores); *projeções estruturais* baseadas na estrutura do programa (árvore de sintaxe abstracta); *projeções de pesquisa e localização*; e *projeções de domínio específico* que permitem o mapeamento de conceitos do domínio do problema com o código (normalmente o mapeamento entre a interface e o código é um exemplo deste tipo de projeção). Assim sendo, propõe-se a criação de uma DSL que permita a especificação de vários tipos de projeções e a sua integração num editor. O resultado final será proporcionar uma forma expedita de aplicar projeções e de gerar visualizações faseadas sobre o código.

## 4 Detecção do Perfil do Programador por Análise de Código

*Supervisor: Pedro Rangel Henriques + Maria João Varanda Pereira*

Resumo:

Na sequência do interessante trabalho pioneiro de Emília Pietriková et al, *Profile-driven Source Code Exploration*, pretende-se neste projeto de mestrado estudar as capacidades de análise estática de código (Java ou C) para perceber o perfil dos programadores.

A ideia base é analisar programas de um programador (profissional ou estudante) e retirar uma série de métricas que possam ser comparadas com padrões para aferir a qualidade ou habilidades do dito programador. As instruções usadas, os padrões repetitivos ao longo do código, a indentação, os identificadores, os comentários, tudo isso ajudará a caracterizar o perfil que se pretende analisar.

## 5 Visualização ontológica de Programas

*Supervisor: Pedro Rangel Henriques + Maria João Varanda Pereira*

Resumo:

Para apoio à complexa atividade de Compreensão de Programas (área de investigação conhecida por *PC – program comprehension*, na qual o nosso grupo de investigação vem trabalhando há anos), propõe-se um tema de tese que consiste em criar visualizações de um programa (o programa em análise) baseadas na apresentação da ontologia do Domínio da Linguagem (a qual deveria ser inferida automaticamente a partir de respetiva gramática) e na apresentação da respetiva ontologia populada (extraída de cada programa).

Note-se que já em trabalho de mestrado anterior resolvemos, com sucesso e interessantes resultados, o problema de derivar uma gramática para uma linguagem concreta a partir da ontologia do Domínio do Problema. Agora neste trabalho pretende-se investigar o oposto: saber se é possível extrair a ontologia a partir da gramática e como o fazer sistematicamente.

Será ainda tema de investigação a procura de soluções de visualização que realmente ajudem a compreender o código em análise.

## 6 Visualização de Redes Sociais dinâmicas

*Supervisor: Pedro Rangel Henriques, Alda Lopes Gançarski*

Resumo:

As diferentes tecnologias utilizadas nas plataformas da Web 2.0 constituem os media sociais. Estes permitem que redes sociais (RS) se constituam a partir de um conjunto de utilizadores que partilham interesses ou que têm ligações directas, como profissionais, familiares ou de amizade. A visualização é uma forma importante de ajudar os utilizadores a perceber o funcionamento e evolução das RS e correspondentes comunidades latentes. Devido à grande escala de utilização das RS, a visualização de um grande número de utilizadores e suas interações torna-se difícil.

O objectivo desta tese é propor um sistema de visualização de RS, tendo em conta as ligações e a geo-localização dos utilizadores, bem como a dimensão temporal. O sistema de visualização deve ser desenvolvido de forma adequada para cada tipo de RS (blogs, wikis, RS de contacto, sites para partilha de recursos, sites para partilha de links).

## 7 SELF-Storify (R)—Sistema de Apoio à Construção de Narrativas Autobiográficas

*Supervisor: Pedro Rangel Henriques + Margarida Rangel Henriques*

### Resumo:

A construção de narrativas autobiográficas está na base da capacidade do ser humano de construir o significado das suas experiências. O significado pessoal abstraído de cada situação vivida, será por sua vez a base das escolhas futuras. De facto, as expectativas e a motivação com que o ser humano se envolve nas situações reflete o seu passado e a forma como "conta" a si e aos outros o sucedido. Por sua vez, a visão que cada um vai construindo de si mesmo, emerge do significado que atribui às experiências do quotidiano. A construção da narrativa implica descrever um acontecimento passado, clarificando as emoções e os pensamentos que irão contribuir para uma melhor compreensão dos comportamentos e expectativas futuras. Para além disso, a clareza quanto ao significado de uma determinada experiência permite a libertação em relação à mesma.

Estudos empíricos têm vindo a documentar que a capacidade de organizar narrativamente as experiências pessoais promove outros domínios, tais como a emergência e consolidação do "self" (Fivush, 1991; Nelson, 1996), a memória autobiográfica (Fivush, Haden, & Adam, 1995), a identidade cultural (Gone, Miller, & Rappaport, 1999) e a integração na comunidade (Wiley, Rose, Burger, & Miller, 1998).

Assim neste contexto e para apoio à atividade de construção de significados com vista a interpretar as experiências do dia a dia, favorecendo uma compreensão fluída das vivências e evitando estados de confusão e bloqueio emocional, os Psicólogos precisam de uma ferramenta de SW que suporte a identificação de cenários, de sensações (sensorialidades aliadas à experiência), de emoções, de pensamentos e, finalmente, de metáforas enquanto interpretação do acontecimento. Trata-se de um sistema que ajude a recriar cenários, acontecimentos, emoções, pensamentos e sentidos (significados) vivenciados por adolescentes em acompanhamento clínico (ou mesmo pela população em geral). O sistema deve ser versátil no sentido de fazer a identificação a partir de afirmações/descrições autonomamente narradas pelo dito adolescente, ou apresentadas como resposta a questões (mais vagas ou mais específicas) formuladas pelo sistema. O sistema poderá também aceitar frases livres ou propor um leque de alternativas (tipo checklist) para o seu interlocutor escolher a que melhor se adequa à sua situação ou sentimento.

Adicionalmente, pretende-se que essa ferramenta permita criar uma rede (um grafo) conectando acontecimentos ao longo da vida, base de uma autobiografia. Este complemento promove o auto-conhecimento, o sentimento de continuidade e segurança base de uma identidade.

## 8 Dicionário da Academia de Ciências de Lisboa

*Supervisor: Alberto Simões*

### Resumo:

Com esta proposta de dissertação pretende-se um trabalho em quatro vertentes:

**Vertente tecnológica:** Análise da base de dados não relacional (noSQL) eXist-DB e da sua plataforma para o desenvolvimento de aplicações baseadas em tecnologias XML;

**Vertente educacional:** Apresentação de forma pedagógica dos métodos de desenvolvimento de aplicações sobre eXist-DB.

**Vertente científica:** Análise e discussão de vantagens/desvantagens do desenvolvimento baseado em tecnologias XML vs o desenvolvimento de aplicações Web tradicional.

**Vertente prática:** Implementação de um sistema baseado em tecnologias W3C para a edição do dicionário da academia de ciências de lisboa.

Tecnologias relevantes: XML, XML Schema, CSS, XSLT, XForms, XQuery e XPath (entre outras).

## 9 DSL para descrição de Aventuras Gráficas

*Supervisor: Alberto Simões*

### Resumo:

Pretende-se desenvolver uma linguagem de domínio específico que permita a definição de uma estrutura base de um jogo típico de "Aventura Gráfica". O processamento desta linguagem deverá gerar um projeto "boilerplate" em Unity que posteriormente deverá ser editado pelo programador no sentido de ajustar detalhes do jogo.

O projecto poderá verter apenas na especificação da linguagem ou na especificação e processamento da linguagem, com a construção de um protótipo, de acordo com a real exequibilidade do projeto.