

**Nome:** Paulo Jorge Machado Oliveira

**Orientadores:** Prof. Doutor Pedro Rangel Henriques (DI – UM)

Prof. Doutora Maria de Fátima Rodrigues (DEI – ISEP/IPP)

**Título:** Detecção e Correção de Problemas de Qualidade em Dados Relacionais

**Resumo:** A análise e exploração de dados para extrair informação e/ou conhecimento que suportem ou auxiliem as actividades de gestão, constituem um factor de sucesso das organizações na sociedade actual. No entanto, vários problemas podem afectar a qualidade dos dados. Estes problemas possuem um efeito negativo nos resultados extraídos, afectando a sua correção e utilidade. Neste contexto, é muito importante detectar e corrigir os problemas de qualidade dos dados. Nesta comunicação apresenta-se:

- Uma taxionomia de problemas de qualidade dos dados, em que estes se encontram organizados por níveis de granularidade de ocorrência.
- Um método de detecção e correção dos problemas de qualidade em que as operações são automaticamente sequenciadas (criação do grafo de dependências), seguindo uma aproximação ascendente (*bottom-up*) por nível de granularidade. O método intercala operações de detecção com operações de correção, efectuando a sua execução de modo incremental.
- Duas linguagens declarativas que permitem ao utilizador especificar, respectivamente, as operações de detecção e correção.
- Uma aproximação que permite a interoperabilidade das operações de detecção e correção entre bases de dados diferentes, através de uma camada ontológica que suporta a sua especificação a um nível conceptual. Esta camada de abstracção torna as operações independentes dos esquemas das bases de dados e permite a sua reutilização.