HITEX: Um Sistema em Desenvolvimento para Historiadores e Arquivistas

José Nuno Oliveira [1]

1 Introdução

Pretende este artigo descrever os objectivos e motivação do HITEX, um sistema informático para registar, organizar e gerir informação de natureza histórica. O sistema está a ser desenvolvido para o Arquivo Distrital de Braga (ADB/UM), em colaboração com o Departamento de Informática da Universidade do Minho, com o apoio da Companhia IBM Portuguesa e da Fundação C. Gulbenkian.

A estrutura deste texto é a seguinte: começa-se por analisar alguns temas em aberto referentes ao tratamento automático do registo histórico. Segue-se a descrição da evolução do projecto HITEX e uma referência ao modelo matemático que se escolheu para nele representar a informação. Segue-se uma breve descrição da estrutura do sistema e da sua interface com o historiador. Termina-se com uma revisão de outros trabalhos nesta área de investigação e propostas de trabalho para o futuro.

1.1 Os Computadores e o 'Métier' de Historiador

Muitos problemas da organização da sociedade moderna escapam, ainda hoje, ao tratamento por computador. Embora muitos outros conheçam já uma informatização bem sucedida (a gestão bancária é disso um bom exemplo), sente-se que existe uma distância apreciável entre a tecnologia disponível e a enorme complexidade de alguns problemas que se gostariam de resolver por computador. Exemplos típicos desses problemas não triviais são o *processo jurídico*, a *análise linguística* e a *investigação em História*, para citar apenas alguns.

Acontece que uma das mais severas limitações do computador é a sua incapacidade para lidar com a *ambiguidade* do discurso humano, o que se sente claramente sempre que se tenta automatizar uma actividade humana. De facto, os computadores são incapazes de entender o mundo que os cerca e apenas o conseguem fazer na medida em que se fôr capaz de criar *modelos estilizados* desse mundo, livres de toda a subtileza ou componente subjectiva. Portanto, qualquer solução automática para um dado problema "humano" pressupõe a concepção para ele de um modelo não-ambíguo. Modelos não ambíguos dizem-se *formais* na medida em que possuem uma estrutura matemática *objectiva*, exactamente aquela que pode ser emulada por um autómato — por um programa de computador, por exemplo.

A emulação que se acaba de referir é, contudo, afectada por outro "salto epistemológico" entre seres humanos e computadores — por mais rápidos e eficientes que sejam estes últimos, não passam de máquinas grotescas cuja "linguagem" nos é difícil de ler e de entender. É por isso que os projectistas de 'software' sentem a necessidade de dividir os seus projectos em duas

[1] Dept. de Informática, Universidade do Minho, Rua D.Pedro V, 88-3º, 4700 Braga

fases, primeiro uma dita de *especificação* (cujo objectivo é construir modelos matemáticos dos problemas) e a de *implementação* a seguir (onde se constroem os programas respectivos).

Em resumo, a aplicação fiável das tecnologias da informação a áreas como a investigação em História vê-se confrontada com, pelo menos, dois níveis de complexidade:

- "desambiguação" (= necessidade de construir especificações);
- implementação (= necessidade de construir programas).

O escrever-se "pelo menos" refere-se à necessidade, em História, de algo mais, como consequência da natureza *evolutiva* do próprio conhecimento histórico. De facto, será muito difícil encontrar uma especificação *estável* e *coerente* do conhecimento de um historiador, sujeito este como está a permanente revisão [2]. Para piorar as coisas, a maior parte das questões que o historiador gostaria de ver respondidas por computador, envolvem o processamento de enormes quantidades de dados, possivelmente incoerentes.

Esta falta de estabilidade na informação histórica é determinante na escolha do tipo de tecnologia informática adequada ao 'métier' de historiador, eliminando à partida as ferramentas convencionais de processamento assentes em *bases de dados* rígidas. Por um lado, torna-se necessário transportar para 'run-time' os próprios modelos da informação histórica, para assim poderem ser modificados ao longo da vida de uma aplicação. A necessidade de processamento interactivo (*i.e.* a orientação manual do historiador) é assim óbvia. Por outro lado, é desejável o acesso eficiente à fonte de informação original (*e.g.* manuscritos, livros de registos, notas tabeliónicas *etc.*), o que sugere ambientes 'multi-media', envolvendo processamento de imagem. Para grandes acervos documentais, deve-se recorrer a dispositivos de grande capacidade de memória como são os *discos ópticos*. Finalmente, a tradição da investigação histórica em publicar edições críticas de fontes sugere a conveniência de se ter à mão ferramentas para *edição electrónica*. Como se mostra mais adiante, estes são exactamente os aspectos aos quais o sistema HITEX pretende dar resposta.

1.2 O Síndrome da Incompatibilidade

É hoje bem conhecido o síndrome de *incompatibilidade* que afecta a fragmentação descontrolada de recursos computacionais em termos de máquinas pessoais. De facto, se "cada um fizer a coisa à sua maneira" proliferarão modelos incompatíveis no mesmo domínio de trabalho, tornando impossível qualquer esforço de integração futura.

O tópico *"Standardization and exchange of machine readable data in the historical disciplines"* posto à discussão na recente conferência da 'Association for History and Computing' (AHC'91) em Odense, Dinamarca, exprime preocupação sobre o facto de esse síndrome estar já a afectar a comunidade dos historiadores, onde muitas bases de dados pessoais se têm construído isoladamente, com pouca ou nenhuma preocupação de compatibilidade [3].

[2] Note-se que qualquer "especificação estável e coerente" depende do universo de informação que se considera pertinente para reconstituir determinado problema. De forma nenhuma se pretende aqui insinuar falta de competência lógica por parte dos historiadores. O que se pretende é justificar a necessidade de ferramentas suficientemente flexíveis para o difícil tratamento mecânico dos vestígios históricos, a compasso com a evolução das maneiras de fazer a História.

[3] O leitor encontra em [Thaller 1991] um diagnóstico exaustivo deste problema, acompanhado de uma lúcida exposição sobre padronização de informação em História.

Re-utilizar é a "palavra mágica" que a indústria de computadores acena hoje contra esta crónica epidemia de incompatibilidade. Neste campo, indústrias como a electrónica, a de produção 'software' *etc.* têm aprendido que devem produzir *componentes re-utilizáveis*, quer dizer, componentes padrão com interfaces (especificações) bem definidas, o que torna possível a sua substituição relativa. Assim, podem-se construir grandes sistemas por mera integração de componentes do mercado, em lugar de os construir a partir do nada, sendo óbvio o aumento em produtividade.

Como se verá adiante, o sistema HITEX tenta adoptar este tipo de estratégia ao propor a noção de *componente re-utilizável com interesse histórico*. Tais componentes arquivam-se no HITEX segundo uma *taxonomia padrão* de conceitos históricos (classes) que exprime a ordem de *subsunção* sobre o conhecimento histórico. Adiante se verá também que todo o componente histórico é tolerante com a informação *incompleta*, um aspecto adicional de muita relevância prática no que diz respeito a aquisição *incremental* da informação histórica.

2 Antecedentes do HiTeX e Evolução do Projecto

O projecto HITEX foi lançado em Janeiro de 1989 na sequência de directivas da gestão do Arquivo Distrital de Braga com respeito a uma utilização efectiva de recursos computacionais a ele doados pela Companhia IBM Portuguesa. Por um lado, sentia-se a necessidade de uma infraestrutura computacional para a re-estruturação incremental e acesso aos vastos fundos documentais de que o Arquivo é depositário [4]. Por outro lado, a política editorial do Arquivo sentia necessidade de melhorar a sua tecnologia de publicação, pretendendo com ela lançar uma nova série editorial de críticas de fontes. Pretendia-se ainda que o 'software' a desenvolver fosse capaz de absorver e publicar os relatórios emitidos por bases de dados já existentes [5].

Um aspecto final a ter em conta era a "arquitectura aberta" que se pretendia para o sistema, por forma a comunicar com ferramentas já em uso no Arquivo (*e.g.* processadores de texto, bases de dados) evitando custos de formação desnecessários.

Assim, a filosofia de base do HITEX foi criar uma *arquitectura aberta* capaz de trocar informação com as ferramentas em uso em cada ambiente onde venha a instalar (desde que essas ferramentas sejam igualmente *abertas*!).

A primeira fase do projecto HITEX concentrou-se no projecto de uma linguagem de fácil utilização para transcrição documental — que se designou por *formato* HITEX — cuja descrição se pode encontrar em [Pereira 1991]. Uma das tarefas mais trabalhosas na edição crítica de fontes é a compilação manual de *índices* que os historiadores tradicionalmente usam como bases de dados textuais — por exemplo, *índices cronológicos* (que registam a ocorrência de informação cronológica na fonte), *índices toponímicos* (que registam a ocorrência de topónimos ou de outra informação geográfica), *índices antroponímicos* (que registam todas as referências a pessoas), *etc.* Assim, no sistema HITEX foi previsto um módulo para compilação *automática* destes índices.

[4] Por exemplo, os Cartórios do Cabido, Mitra e Sé de Braga, para cuja re-estruturação se havia obtido apoio da Fundação Calouste Gulbenkian.

[5] Por exemplo, o da base de dados de *Inquirições de Genere* cuja digitação arrancou em 1988 e que já regista cerca de trinta mil processos de ordenação (séculos XVII a XX).

O sistema LaTeX [Lamport 1986] para preparação de texto foi escolhido como ferramenta de base para construir a correspondente parte do sistema HITEX, por várias razões:

- LaTeX baseia-se numa linguagem tipográfica estruturada, de grande capacidade expressiva;
- LaTeX está disponível em praticamente todos os recursos computacionais, desde IBM-PC compatíveis aos ambientes MACINTOSH e UNIX;
- LaTeX produz resultados gráficos (em TeX) de muito boa qualidade.

A segunda fase do projecto dedicou-se à tarefa da *representação do conhecimento*. Note-se que os índices produzidos pelo HITEX podem ser já considerados como uma representação (textual) de conhecimento histórico. Contudo, esses índices são registos "planos" de informação que ignora a textura interligada e multifacetada do conhecimento humano. Mais ainda, esses índices terão sempre uma qualidade necessariamente dependente do seu autor.

Assim, o sistema HITEX evoluiu no sentido de encarar tais índices não como fontes de conhecimento, mas antes como produtos de um sistema de representação de conhecimento. Quer dizer, esperava-se a produção *automática* desses índices sob a forma de relatórios textuais de alguma base de conhecimento previamente construida. O projecto de uma tal base foi influenciado por dois paradigmas bem conhecidos na tecnologia informática actual — as *redes semânticas* [Chouraqui 1986] da Inteligência Artificial, e a *programação por objectos* [Goldberg & Robson 1983]. O modelo resultante será abordado mais à frente.

O protótipo mais actual do sistema corre sobre MS/DOS e foi escrito na linguagem SMALLTALK[Goldberg & Robso de programação por objectos. A interface com o utilizador deste protótipo contém já o acesso a ficheiros de imagens documentais obtidas a partir de uma unidade de digitalização ('scanning') e gravadas num sistema de disco óptico. A Figura 1 mostra o aspecto do 'ecran' depois de se ter pedido o acesso às primeiras linhas de um *folio* do primeiro volume do *Índice das Gavetas do Cabido de Braga*.

3 Sobre o Modelo de Dados HiTeX

Como atrás se disse, a arquitectura da base de conhecimento do HITEX (BCH) é semelhante à de uma rede semântica hierárquica. A informação histórica regista-se sob a forma de "grãos" de informação univocamente acedidos por identificadores únicos.

Depende do utilizador a decisão sobre quão específico um dado componente é. Seja por exemplo o seguinte fragmento de uma fonte [6]:

> *Certidão da doação que o arcebispo de Braga D.Martinho de Oliveira fez ao Cabido de Braga [...] Ano de 1300.*

A análise meticulosa deste fragmento de fonte revela a presença dos seguintes grãos de informação:

- *D. Martinho de Oliveira* era o Arcebispo de Braga em 1300;

[6] Início do primeiro volume do Índice das Gavetas do Cabido de Braga.

Figura 1: Acesso a imagens documentais no protótipo do HITEX.

- Este arcebispo doou qualquer coisa ao *Cabido de Braga*, nesse mesmo ano;
- O Cabido guardou uma *certidão* dessa doação;
- Essa certidão está arquivada no fundo documental designado *Gavetas do Cabido*;
- Uma referência a essa certidão é a que se pode encontrar no fol.1, vol.1 do correspondente *Índice*, compilado no século XVIII.

Como registar estes componentes de forma flexível e incremental?

Primeiro é preciso saber o que é que de facto um componente histórico "é". O HITEX segue a linha dos sistemas ditos *de objectos* segundo a qual se associam *classes* ou *tipos* a cada componente. Vamos então tentar *classificar* os componentes acima:

- *Martinho de Oliveira* → *Arcebispo*;
- *cabido de Braga* → *Cabido*;
- *doação* ... → *Contrato*;
- *certidão da doação* ... → *Certidão*.

A questão é, agora, "o que é uma classe?". Por exemplo, o que é uma *certidão*? Toda a certidão é, concerteza, um *documento*. Mas é um documento especial — um documento *institucional*. Outros documentos podem existir que não sejam institucionais, *e.g.* cartas particulares. Assim, a introdução de mais algumas classes ajudam a "enquadrar" esta nossa noção (para já algo imprecisa) de certidão:

$$Documento \left\{ \begin{array}{l} Particular \left\{ \begin{array}{l} \vdots \\ Carta \\ \vdots \end{array} \right. \\ Institucional \left\{ \begin{array}{l} \vdots \\ Notarial \left\{ \begin{array}{l} \vdots \\ \textit{Certidão} \\ \vdots \end{array} \right. \\ \vdots \end{array} \right. \\ \vdots \end{array} \right. \tag{1}$$

Em resumo, acabamos de esboçar uma *taxonomia de classes* documentais, afinal da mesma maneira como, por exemplo, as espécies vêm sendo classificadas pelos biólogos. É claro que (1) não é necessariamente a melhor taxonomia para documentos. Mas é um ponto de partida, e uma das características do HITEX é permitir a revisão deste tipo de conhecimento, ainda que dentro de certos limites. Note-se que toda a taxonomia tem uma *ordem de subsunção* inerente, segundo a qual se pode falar de *subclasses* e de *superclasses*. No exemplo acima, *Certidão* é subclasse de *Notarial*, *Documento* é superclasse de *Particular*, *etc.*

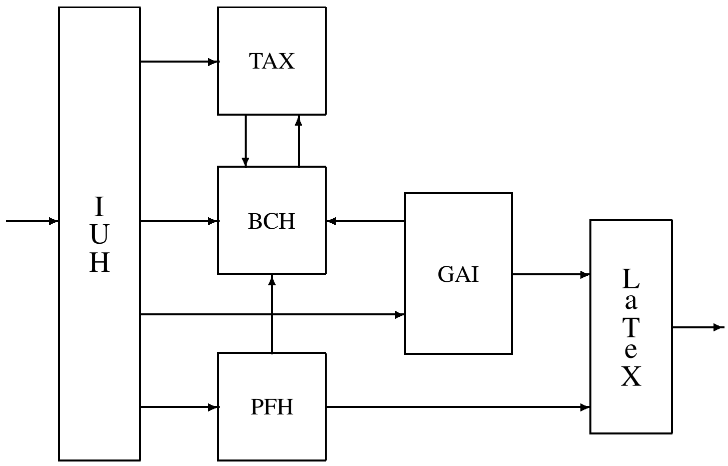

Figura 2: Diagrama de blocos do protótipo do HITEX.

Não cabe aqui prosseguir com esta análise do modelo HITEX, que o leitor pode encontrar discutido em pormenor na referência [Oliveira *et al* 1991]. Interessa é referir que cada classe é caracterizada por *atributos* (propriedades) que são afinal os *dados* que se podem arquivar sobre os componentes dessa classe, também designados por *instâncias*. No facto de valores de atributos poderem referir outros componentes reside a estrutura *interligada*, ou em *rede*, da base do HITEX. Um dos aspectos mais relevantes desta ordenação de classes é o facto de todo o tipo de informação disponível numa classe ficar automaticamente disponível (diz-se que *por herança*) nas suas subclasses, o que confere grande economia e flexibilidade ao modelo.

4 Arquitectura do Sistema HiTeX

A Figura 2 mostra o diagrama de blocos estruturais do actual sistema HITEX. O bloco IUH (Interface do HITEX com o Utilizador) é responsável pela gestão da interacção com o utilizador, tal como se ilustra na secção seguinte. Grande parte dessa interacção tem a ver com a criação e manutenção da Base de Conhecimentos (BCH) e correspondente taxonomia (TAX) onde, como se viu, toda a informação é registada.

Os outros blocos estruturais do HITEX têm a ver com o tratamento de fontes (transcrição documental). PFH é o Processador do Formato HITEX atrás referido, produzindo código LaTeX que pode ser impresso em impressora normal, 'laser' ou 'phototypesetter'. Finalmente, o bloco GAI (Gerador Automático de Índices) pode ser activado para se gerarem automaticamente os índices toponímico, cronológico *etc.* Mais uma vez, o texto dos índices é produzido em LaTeX, e "colado" — se pretendido — ao texto já processado, para tudo ser impresso. A referência [Pereira 1991] contém uma ilustração do processo de geração dos índices, tomando como base o primeiro volume do *Índice das Gavetas* atrás referido.

5 O HiTeX Visto pelo Utilizador

O protótipo corrente do HITEX dá ao utilizador primitivas para:

- gestão da base de conhecimento;
- acesso a *facsimiles* de documentos arquivados em memória óptica;
- transcrição documental, actividades que o utilizador pode invocar integrada ou isoladamente.

A gestão da base de conhecimentos permite actividades como:

- navegação sobre o diagrama da taxonomia do sistema;
- procura de componentes;
- criação de um novo componente da classe correntemente seleccionada na taxonomia;
- preenchimento de novos dados referentes ao componente correntemente seleccionado;
- especialização de um dado componente;
- adição de uma nova classe à taxonomia;
- adição de mais atributos a uma dada classe;
- interpelação assistida da base de conhecimentos;
- *etc.*

As Figuras 1 e 3 mostram 'ecrans' do protótipo disponível em MS/DOS. A janela superior esquerda mostra a taxonomia do sistema. Nas janelas laterais faz-se o acesso e a gestão de atributos e instâncias. Toda a conversação com o utilizador é feita através de menus. A janela inferior tem a ver com transcrição de fontes. Esta janela não é mais do que um simples editor de texto, capaz de comunicar com a base de conhecimentos sempre que se quiser. É de notar que uma fonte pode ser transcrita usando outros editores de texto, sendo posteriormente importada para esta janela com facilidade desde que não contenha caracteres de controlo obscuros. Este editor tem a vantagem de permitir referências cruzadas em modo semi-automático: deixando o cursor na posição pretendida o utilizador pode, entretanto, pesquisar a base procurando os componentes que lhe interessarem. Um simples toque no rato (*cf.* a opção `to text` na Figura 3) será suficiente para que os identificadores dos componentes seleccionados sejam inseridos automaticamente na janela de texto, imediatamente após a posição do cursor. *Relembra-se que tais identificadores têm por objectivo marcar a ocorrência de ítens que se destinam a índices.*

A transcrição de uma fonte é um texto que deve obedecer ao padrão que é o *Formato* HITEX. No protótipo corrente ainda não está disponível a edição dirigida pela respectiva sintaxe. A activação do bloco PFH agarra no ficheiro corrente de transcrição, e passa-o pelo tradutor do formato gerando código LATEX, assim como um ficheiro interno de comandos que se pode fazer correr em 'batch' para actualizar referências cruzadas na base de dados do HITEX, o que apenas afecta os componentes envolvidos.

Figura 3: Tratamento de referências cruzadas no editor do HITEX.

6 Outros Trabalhos Relacionados com o HiTeX

Há muito que os computadores são utilizados pelos historiadores. Referências bibliográficas como, por exemplo [AHC 1989, Amado *et al* 1990], mostram a utilização de bases de dados tradicionais e de pacotes estatísticos. Contudo sentem-se aí críticas às limitações da tecnologia clássica quando confrontada com a complexidade da informação relevante para o historiador.

O problema da falta de re-utilização e troca de dados entre historiadores é abordado extensivamente em [Thaller 1991], que se concentra aliás no difícil problema da *padronização*. Há muitos pontos comuns entre as motivações do HITEX e a análise de Thaller, sobretudo no que diz respeito à necessidade de uma formalização de *padrões abstractos* dos tipos de dados presentes nas fontes históricas, independente da sua representação em diferentes ambientes de processamento. O projecto TEI [ACH 1991] de um padrão textual assente num 'standard' internacional (SGML) é um avanço significativo nesta área.

Note-se que, apesar do seu aparente sucesso, o paradigma do "hipertexto" [ACM 1988] não chega a ser uma solução para o historiador, dada a sua natureza "fechada" e a falta de uma semântica rigorosa [Fountain *et al* 1990]. Para além disso *"a experiência do passado ensinou-nos que qualquer 'standard' preso a um nível tecnológico específico está condenado a falhar"* [Thaller 1991].

O desenvolvimento do HITEX pode inserir-se nas muitas iniciativas de I&D para conceber ferramentas mais úteis à disciplina da História. Uma abordagem semelhante à do HITEX é a do sistema MICROCOSM [Fountain *et al* 1990]. Partindo de uma análise crítica do paradigma do hipertexto, a abordagem MICROCOSM baseia-se na constatação de que "*hypermedia links in themselves are a valuable store of knowledge*". Tal como no HITEX, e ao contrário dos sistemas de hipertexto, relacionamentos complexos não são registados sobre as imagens dos próprios documentos, mas antes separados numa base de conhecimento independente. Aquilo que em MICROCOSM se chama *universo geral de discurso* assemelha-se à noção de taxonomia no HITEX. O projecto MICROCOSM já produziu duas implementações alternativas, uma sobre MS-WINDOWS e outra escrita na linguagem ACTOR [Fountain *et al* 1990].

Outro sistema que partilha a motivação do HITEX chama-se LACE [Rahtz *et al* 1989]. Este é mais um sistema tipo *hipertexto aberto*, que corre em estações de trabalho SUN e se baseia no LaTeX, em NEWS (um sistema de gestão de janelas baseado em POSTSCRIPT), e num servidor de 'hypermedia'.

Outro trabalho relacionável com o HITEX é o de Barroca [Barroca 1991] em bases de dados orientadas a objectos para a arqueologia. A sua ênfase é mais em *método* que em projecto de um sistema, mas de modo semelhante à disciplina que o HITEX sugere para a gestão incremental de dados históricos.

7 Conclusões e Perspectivas para Futuro

O projecto HITEX está em curso no Arquivo Distrital de Braga e chegou ao fim da sua segunda fase de desenvolvimento após a construção de um protótipo escrito em SMALLTALK e C. Embora a interface com o utilizador esteja bastante mais aperfeiçoada do que a de um protótipo anterior [Pereira 1991], é difícil prever quão longe se pode progredir com este protótipo no que diz respeito ao tamanho da base de conhecimentos. O tratamento e publicação dos seis volumes do *Índice das Gavetas do Cabido* (que já se encontram todos transcritos) está em curso, o que constituirá um teste definitivo ao sistema.

A terceira fase do projecto começará logo que o protótipo seja testado extensivamente e avaliado pela equipa de historiadores do projecto. Prevê-se a necessidade de aumentar a capacidade de dedução automática do sistema (veja-se aqui o que é proposto por Hsieh [Hsieh 1990]) assim como melhoramentos na produção de índices textuais (a geração automática de linguagem natural requerirá investimento adicional em linguística computacional [Ravera 1990]).

Nesta última fase planeiam-se progressos significativos em integração do sistema sobre tecnologia convencional, *e.g.* a sua montagem sobre bases de dados relacionais, implementação sobre o 'standard' CDS-ISIS da Unesco [UNESCO 1989], acesso ao SPSS *etc.* Uma novidade desta integração consistirá na aplicação sistemática de um cálculo de implementação, desenvolvido independentemente pelo autor [Oliveira 1990], com o objectivo de colmatar o desafio implementacional do HITEX.

Agradecimentos

O autor deseja agradecer todo o apoio prestado pela Dr. Maria Assunção Vasconcelos (Directora do ADB) e da Sr^{a} D. Leónida Gomes (técnica do ADB). Agradece ainda ao Dr. Nuno Madureira (ISCTE) as suas sugestões e comentários sobre a primeira versão deste texto.

O 'software' HITEX está a ser desenvolvido para o Laboratório IBM de Microcomputadores do ADB. Agradece-se ainda ao núcleo do INESC da Universidade do Minho o acesso à sua rede de 'workstations' SUN.

O Projecto HITEX tem recebido apoios da Fundação Calouste Gulbenkian via financiamento Nr.E/75/88.

Referências

[AHC 1989] Actes du IV^{e} Congrès "History and Computing'89" *L'Ordinateur et le Métier d'Historian.* CNRS, Travaux et Documents 2, Maison des Pays Ibériques 1990.

[ACH 1991] ACH-ACL-ALLC *Guidelines for the Encoding and Interchange of Machine Readable Texts.* Report TEI P1, Draft Version 1.1, June 1991.

[ACM 1988] Association for Computing Machinery. *Special Issue on Hypertexts.* ACM Comm., 31(7), Jul. 1988.

[Amado *et al* 1990] Amado J.P., Cardoso J.C., Neves A.M. *Hϵródoto 1.0 — Estação de Trabalho Informática em História e Arquivística.* Fac. de Ciências Sociais e Humanas da U.N.Lisboa, 1990.

[Barroca 1991] Barroca L. *Object-oriented Database Design in Archæology.* Dept. of Computer Science, University of Minho, 1991.

[Chouraqui 1986] Chouraqui E. *Representation des Connaissances et Comprehension du Language Naturel.* Informatica e Archivi, Torino 17–19 Guigno 1985; Min. per i Beni Culturali e Ambientali, Publicazioni degli Archivi di Stato, Roma, 1986.

[Fountain *et al* 1990] Fountain A., Hall W., Heath I., Davis H. *MICROCOSM: An Open Model for Hypermedia With Dynamic Linking.* Technical Report CSTR 90-12, Dept. of Electronics and Comp. Science, Univ. of Southampton, 1990.

[Goldberg & Robson 1983] Goldberg A., Robson D. *Smalltalk'80: the Language and Its Implementation.* Addison-Wesley 1983.

[Hsieh 1990] Hsieh D. *A Logic to Unify Semantic Network Knowledge Systems with Object-oriented Database Models.* Technical Report SRI-CSL-90-15, SRI International, Dec. 1990.

[Lamport 1986] Lamport L. LaTeX — *A Document Preparation System.* Addison-Wesley Publishing Company, 5^{th} edition, 1986.

[Oliveira 1990] Oliveira J.N. *A Reification Calculus for Model-Oriented Software Specification.* Formal Aspects of Computing, Vol.2, 1-23, 1990, Springer-Verlag.

[Oliveira *et al* 1991] Oliveira J.N., Araújo A.S., Silva A.M. *Historical Records Processing in the HiTeX System.* To appear in Proc. of AHC'91 — 6^{th} International Conference of the Association for History & Computing 1991, Odense, Denmark, Aug. 28-30.

[Pereira 1991] Pereira, A.M. *HiTeX — Um Sistema para Transcrição Documental em Larga Escala.* Relatório de Estágio, LESI, U.Minho, Braga 1991.

[Rahtz *et al* 1989] Rahtz S., Carr L., Hall W. *New Designs for Archæological Reports.* Science and Archæology, Vol.31, 20-34, 1989.

[Ravera 1990] Ravera S.B. *Information-Based Linguistics and Head-Driven Phrase Structure.* In *Natural Language Processing*, Lecture Notes in Artificial Intelligence, Vol.476, 55-101, Springer-Verlag, 1990.

[Thaller 1991] Thaller M. *The Need for Standards: Data Modelling and Exchange.* in *Modelling Historical Data: Towards a Standard for Encoding and Exchanging Machine Readable Texts*, 1–18, Greenstein D.I. (Ed.), Max-Planck-Institut für Geschichte, St. Katharinen, 1991.

[UNESCO 1989] UNESCO *Mini-micro* CDS/ISIS *Reference Manual.* Version 2.3, March 1989, UNESCO, 7 Place de Fontenoy, Paris.

1. Ficha Técnica.

O protótipo corrente do HITEX *corre em* MS/DOS *sobre* SMALLTALK*/V R2.0 (© 1986 Digitalk Inc.) e* LaTeX *(© American Mathematical Society) e pressupõe a aquisição prévia destes sistemas. O núcleo do sistema* HITEX *cabe numa diskete de 3.5in, exigindo um mínimo de 640 Kb de memória* RAM*, 40 Mb de disco e placa gráfica* VGA.

2. Resenha Histórica.

O projecto HITEX *foi lançado em 1989 com apoio da Companhia* IBM *Portuguesa, da Fundação C. Gulbenkian e do Departamento de Informática da Universidade do Minho.*

O HITEX *encontra-se em fase de protótipo, tendo já sido apresentado ao* IPA *e na conferência* AHC'91 *(Dinamarca, Agosto de 1991).*

3. Contactos.

Dr. Maria Assunção Vasconcelos, Direcção do Arquivo Distrital de Braga, Universidade do Minho, Largo do Paço, 4719 Braga Codex, Tel. 053–612234, Fax. 053–616936.

Curriculum vitæ do autor.

José Nuno Oliveira (n.1955) obteve os graus de M.Sc. (1981) e Ph.D. (1984) em Ciências da Computação em Manchester, Reino Unido. É actualmente professor associado do Departamento de Informática da Universidade do Minho, e investigador do INESC*, núcleo de Braga.*

Tem desenvolvido actividade de investigação no Arquivo Distrital de Braga, em regime de colaboração.