



UNIVERSIDADE DE COIMBRA

# **TIME LINK: A EVOLUÇÃO DE UMA BASE DE DADOS PROSOPOGRÁFICA**

**João Manuel Silva Antunes Neves de Carvalho**

**Faculdade de Letras da Universidade de Coimbra**

# **Time Link: a evolução de uma base de dados prosopográfica**

**João Carvalho**

Dissertação de mestrado em Património Europeu, Multimédia e Sociedade de Informação, apresentada à Faculdade de Letras da Universidade de Coimbra, sob a orientação do Professor Doutor Joaquim Ramos de Carvalho.

## Agradecimentos

Este trabalho não teria sido possível sem a preciosa ajuda do professor Doutor Joaquim Ramos de Carvalho, coordenador do mestrado em Património Europeu, Multimédia e Sociedade de Informação, e orientador.

Gostaria também de expressar uma palavra de gratidão aos professores das disciplinas curriculares do mestrado.

Agradeço ainda ao Filipe e à Mafalda o apoio.

## Índice Geral

<b>Índice de figuras.....</b>	<b>VI</b>
<b>Índice de tabelas.....</b>	<b>VIII</b>
<b>Introdução .....</b>	<b>1</b>
<b>1. Os serviços genealógicos .....</b>	<b>3</b>
1.1. Os programas de aplicação.....	3
1.2. Os sítios na Web .....	8
1.3. As redes sociais online.....	14
<b>2. MHK: um sistema para a reconstituição de comunidades históricas.....</b>	<b>17</b>
2.1. Base de dados (SGBD) .....	17
2.2. Inserção de dados (Kleio).....	20
2.3. A interface gráfica de utilizador (GUI).....	21
2.4. Algumas bases de dados: exemplos concretos .....	23
<b>3. A informação genealógica .....</b>	<b>25</b>
3.1. Registos paroquiais .....	25
3.2. Registos notariais.....	26
3.3. Nobiliários .....	27
3.4. Outras fontes.....	27
<b>4. Os modelos de negócio .....</b>	<b>28</b>
<b>5. Actores e casos de uso (requisitos).....</b>	<b>34</b>
5.1. Actores e objectivos (actors and goals) .....	34
5.2. Casos de uso (Use cases) .....	35
<b>6. Time Link: o desenho de um novo sistema .....</b>	<b>39</b>
<b>6.1. A plataforma .....</b>	<b>39</b>
<b>6.2. A interface de utilizador.....</b>	<b>41</b>
<b>6.3. Os módulos .....</b>	<b>42</b>
6.3.1. Famílias e genealogias .....	42
6.3.2. Identificação colectiva/cooperativa de pessoas .....	48
6.3.3. Integração de mapas (Google Maps).....	52
6.3.4. Comentários.....	54
6.3.5. Introdução de dados por utilizadores .....	56
6.3.6. Suporte a múltiplos idiomas .....	57

6.3.7. URL's únicos.....	58
6.3.8. Interligação entre diferentes bases de dados .....	59
6.3.9. Factores de infra-estrutura: instalação e registo de utilizadores .....	60
<b>Considerações finais.....</b>	<b>63</b>
<b>Bibliografia .....</b>	<b>65</b>
<b>Anexo I. Estatísticas de acesso a <i>timelink.fl.uc.pt</i> .....</b>	<b>70</b>

## Índice de figuras

Figura 1. Visitantes únicos no portal Ancestry.com, entre Abril de 2008 e Abril de 2009.	11
Figura 2. Visitantes únicos no portal FamilySearch.org, entre Maio de 2008 e Maio de 2009.	12
Figura 2. Visitantes únicos no portal MyHeritage.com, entre Maio de 2008 e Maio de 2009.	16
Figura 3. Estrutura de dados (Carvalho, 1997)	19
Figura 4. Exemplo de uma transcrição de registo paroquial (acto de baptismo).	20
Figura 5. Página de entrada	22
Figura 6. Página de Pesquisa	22
Figura 7. Ficheiros Kleio	23
Figura 8. Página de Sistema	23
Figura 9. Ficha de família	44
Figura 10. Árvore genealógica	45
Figura 11. Vista de costados	46
Figura 12. Listagem com caixas de verificação.	49
Figura 13. Conjunto de ocorrências disponíveis para serem identificadas como pertencendo a uma mesma pessoa.	50
Figura 14. Listagem com ocorrências identificadas e com a distinção de indivíduos através das cores.	50
Figura 15. Ficha biográfica ou de pessoa real com indicação do utilizador responsável.	51
Figura 16. Separador de 'Identificações' da Pessoa Real.	51
Figura 17. Visualização com as referências espaciais da base de Eiras-Coimbra, recorrendo à Google Maps API.	53
Figura 18. Tabela de comentários de utilizador	55
Figura 19. Ecrã com lista de comentários na página de entrada	55
Figura 20. Ecrã com um comentário agregado a uma entidade da base de dados.	56
Figura 21. Exemplo de ficheiro de tradução.	58
Figura 22. Exemplo do ficheiro de configuração	59
Figura 23. Indicação da base de dados seleccionada actualmente e das restantes bases de dados acessíveis para o utilizador.	60
Figura 24. Formulário de registo de utilizadores online.	61
Figura 25. MHK/Time Link Portable a iniciar e respectivo ícone na barra de tarefas do Windows.	62
Figura 26. Total de visitas e visualizações de páginas, entre Março de 2009 e Outubro de 2010 (Google Analytics)	70

Figura 27. Distribuição espacial das visitas (Google Analytics)..... 71

## Índice de tabelas

Tabela 1. Volume de mensagens da <i>mailing list</i> ROOTS-L. ....	9
Tabela 2. Número de entradas para uma pesquisa por genealogia no Google (2009). ....	10
Tabela 3. Percentagem de registos preservados por distrito (Dic. Hist Port.) . ....	26
Tabela 4. Assinaturas disponibilizadas pelo portal Ancestry. ....	29
Tabela 5. Custos de assinatura do portal Geni.com. ....	30
Tabela 6. Exemplo da tabela de tercetos para os actos de baptismo . ....	47
Tabela 7. Novas tabelas para armazenamento da informação genealógica. ....	47
Tabela 8. Tabela de entities . ....	57
Tabela 9 Número de visitas a <a href="http://timelink.fl.uc.pt">timelink.fl.uc.pt</a> por país entre Março de 2009 e Outubro de 2010 (Google Analytics) . ....	70

## Introdução

A transformação de uma aplicação de gestão de bases de dados prosopográficas, desenvolvida em meio académico e utilizada por investigadores profissionais, num produto mais abrangente e apelativo, acessível a um público mais vasto, é o objecto do presente trabalho.

A aplicação é conhecida, na sua versão original, pela sigla MHK, que desdobrada resulta em Micro-History with Kleio (ou Micro-História com Kleio<sup>1</sup>), e resulta do trabalho desenvolvido, na Faculdade de Letras da Universidade de Coimbra, ao longo de vários anos, pelo Doutor Joaquim Ramos de Carvalho. A primeira versão tem origem na sua dissertação de doutoramento, apresentada à Faculdade de Letras da Universidade de Coimbra, em 1997, sob o título *Comportamentos Morais e Estruturas Sociais numa paróquia de Antigo Regime (Soure, 1680-1720)*, da qual persistem o modelo de dados e a forma de registo e importação das fontes históricas. A aplicação foi depois utilizada noutros trabalhos de investigação de deram origem a dissertações de mestrado e doutoramento e a novas bases de dados, de que são exemplo Soure, Lousã e Eiras.

Tomando consciência de um crescente interesse pelas questões da genealogia, espelhado no aparecimento de novos serviços online, equacionou-se, pela primeira vez, a conversão desta ferramenta de investigação numa aplicação acessível na Internet. Ambicionou-se levá-la a novos públicos e, deste modo, permitir que outros pudessem retirar vantagem do manancial de informação que até então permanecia nos computadores pessoais de cada investigador.

O MHK, apesar de concebido e implementado como uma aplicação Web, nunca estivera verdadeiramente acessível na grande rede, na Internet. Para que essa ambição se concretizasse, seria necessário efectuar importantes transformações, especialmente nos componentes de visualização da informação, mas também em certos aspectos mais fundamentais da implementação. São estas transformações que se descrevem no corpo do presente documento.

Desde logo no primeiro capítulo, identificam-se “os serviços genealógicos” existentes no mercado, desde as aplicações informáticas mais tradicionais, passando pelos *websites* dedicados ao tema e terminando as redes sociais online dedicadas à genealogia.

O segundo capítulo - “MHK: um sistema para a reconstituição de comunidades históricas” - descreve aquele que é ponto de partida tecnológico, ou seja, o MHK na sua forma original. Descrevem-se, de forma relativamente sucinta, o modelo de dados, os mecanismos de transcrição e inserção de dados e, por fim, a interface gráfica de utilizador (GUI).

No terceiro capítulo - “a informação genealógica” - procuram identificar-se as principais fontes históricas portadoras de informação genealógica, tecendo-se breves considerações históricas sobre a origem dessas mesmas fontes e identificando a sua localização ou possíveis localizações.

O capítulo quarto - “os modelos de negócio” - faz um elenco de modelos de negócio explorados por entidades ligadas à disponibilização de conteúdos e prestação de serviços *online*, especialmente conteúdos e serviços ligados à genealogia.

---

<sup>1</sup> Kleio é uma notação de transcrição de documentos apresentada com maior detalhe no segundo capítulo.

Por fim, nos dois últimos capítulos, entra-se definitivamente nos aspectos fundamentais da transformação e adaptação do MHK. Elencam-se os requisitos que deverão guiar essa transformação e expõem-se os pormenores, devidamente ilustrados, da nova versão. O MHK assumirá não apenas uma nova forma mas também uma nova identidade enquanto Time Link.

## 1. Os serviços genealógicos

Em 2007 o New York Times publicou um artigo com o título “Latest Genealogy Tools Create a Need to Know” (Rosen, 2007) onde a genealogia é apresentada como um grande negócio e uma actividade popular. Nesse artigo, ainda que focado na realidade norte-americana, revela-se a existência de um interesse geral da população pelas suas origens ancestrais, interesse que gera um amplo conjunto de recursos online, desde aplicações informáticas no sentido mais tradicional, motores de pesquisa, novas redes sociais ou mesmo serviços de investigação profissional<sup>2</sup>.

Na mesma linha, um estudo recente (Liberge, 2008-2009) da responsabilidade do governo Holandês, veio reafirmar o interesse de uma parte importante da população (42%) pela sua história familiar e local. Esta valorização manifesta-se, sobretudo, a partir dos 30 anos, com autonomização dos indivíduos e a constituição de novos agregados familiares. Nesse estudo, e reportando-se ao caso Holandês, conclui-se que no caso dos Países Baixos, que contam com mais de 16 milhões de habitantes, essa percentagem corresponderia a 6,7 milhões de consumidores, potenciais ou reais, deste tipo de produtos (Liberge, 2008-2009). São sinais reveladores, desde logo, de apetência pelo tema.

Ao longo das próximas páginas, procurar-se-á expor de forma mais concreta esse interesse potencial, nas suas diferentes manifestações e tendo como pano de fundo os modelos de negócio adoptados e possíveis públicos-alvo. A abordagem focar-se-á, essencialmente, em três grandes blocos: uma breve reflexão sobre programas de aplicação *desktop* destinados à recolha, organização, processamento e apresentação da informação de carácter genealógico; os sítios na World Wide Web dedicados à genealogia, onde se incluem também os motores de pesquisa especializados e os portais estatais; e ainda as redes sociais assentes na genealogia e integradas na denominada Web 2.0.

### 1.1. Os programas de aplicação

O cruzamento da genealogia com o mundo dos computadores dá-se muito cedo na história da computação pessoal. Na verdade, ainda antes do aparecimento dos primeiros computadores pessoais já o acesso a *mainframes*<sup>3</sup> era utilizado para algum tipo de trabalho com informação genealógica: “and before the invention of the home computer, I entered onto 80-column punch cards the names and limited information about 200 or so of my ancestors. I did this after hours in my employer’s data center. I then used the employer’s mainframe computer that cost hundreds of thousands of dollars to sort the data and to print a few crude reports” (Eastman, 2002). Assim, seria de esperar que com o lançamento dos primeiros computadores pessoais, nos anos 70 do século XX, desde logo estes fossem utilizados como auxiliar da investigação genealógica. Numa primeira fase utilizando recursos dos próprios equipamentos - como os primeiros processadores de texto e as primeiras impressoras na impressão de relatórios - e posteriormente criando programas de aplicação especificamente destinados a esse fim. O Genealogy: Compiling Roots and Branches - também conhecido por GCRAB - é um dos primeiros programas de genealogia do conhecimento geral. Escrito por John J. Armstrong na linguagem Microsoft BASIC surge

---

<sup>2</sup> <http://www.genealogiportuguesa.com>

<sup>3</sup> Mainframe, termo inglês para computador de grandes dimensões, utilizado no processamento de grandes volumes de dados.

impresso na revista "Personal Computing Magazine". Corria o ano de 1979, e envolvia a introdução manual do código fonte<sup>4</sup> na máquina. A aplicação podia ainda ser adquirida numa disquete, por 250 dólares.

A “explosão” dos programas de aplicação especializados na investigação genealógica dá-se nos anos 80. Nessa década surgem aplicações cuja popularidade perdurará no tempo, alguns resistindo até à segunda metade da década de 90, ou mesmo até à primeira década do novo milénio, quando uma nova revolução começa a insinuar-se.

O ROOTS, da responsabilidade da CommSoft, foi o primeiro programa de tratamento de informação genealógica a atingir alguma popularidade e que, ao longo das sucessivas versões, foi introduzindo muitas das funcionalidades que encontramos nas actuais aplicações (Nurse, 1996). A primeira versão, denominada ROOTS89, foi lançada em Abril de 1981, destinada às máquinas Heathkit H-89. A aplicação foi ainda adaptada para sistemas CP/M, sob o nome ROOTS/M e, finalmente, para os actuais IBM/PC com o nome ROOTS II.

A generalidade destes primeiros programas de aplicação são ainda consideravelmente dispendiosos, sobretudo atendendo à relação entre funcionalidade e preço: não permitem a exportação ou importação de informação e atingem frequentemente valores na ordem das várias centenas de dólares. Não obstante, a década de 80 ficará ainda marcada pelo aparecimento de outros programas, um alargamento da oferta que será acompanhado por melhorias funcionais nas aplicações e preços mais competitivos (Eastman, 2002).

Outras entidades desempenharam um papel importante no desenvolvimento de aplicações informáticas e de estratégias de recolha da informação. A Igreja de Jesus Cristo dos Santos dos Últimos Dias representa um caso peculiar de interesse pela informação genealógica. Decorrente das suas crenças e valores, esta instituição religiosa desde cedo investiu na criação de ferramentas informáticas e mesmo de sítios na Web dedicados à genealogia e à família: “The Family History Department began the massive project of automating its data and gathering further records from the genealogical community in automated format in the early 1970s” (Allen et al., 1994). O Personal Ancestral File, actualmete na versão 5, é uma ferramenta gratuita dedicada à organização da informação genealógica, desenvolvida pela Igreja Mórmon desde Abril de 1984 (Allen et al., 1994). A comunidade Mórmon desenvolveu ainda um trabalho intensivo de recolha de registos (paroquiais) sob a forma de microfilme, sendo responsável pela microfilmagem de inúmeros arquivos em Portugal (Público, 2007).

Retomando o tema principal deste capítulo, persiste actualmente uma ampla oferta de aplicações informáticas dedicadas à genealogia. Algumas surgiram ainda na década de 80 e foram conquistando o apreço dos utilizadores. Servem de exemplo o programa de aplicação da responsabilidade da Igreja Mórmon, o Personal Ancestral File, anteriormente referido, ou o Family Tree Maker, analisado com maior detalhe mais à frente neste capítulo. Uma pesquisa por programas de aplicação relacionados com genealogia revela um conjunto extenso: Ancestral Quest, Atlas Family Historian, Brother's Keeper, Family Historian, Family Tree Builder, Family Tree Maker, Genbox Family History, Gramps, Legacy Family Tree, MacFamilyTree, The Master Genealogist, Ancestral File, Personal Ancestral File, Personal Historian, Reunion, RootsMagic, entre outros. É ainda possível encontrar ferramentas especializadas em determinadas tarefas – o desenho de genogramas e árvores

---

<sup>4</sup> Código fonte: instruções introduzidas no computador, conhecidas como linguagem de programação, e que são a origem dos programas informáticos.

genealógicas – como o GenoPro para ambientes Windows, ou mesmo uma aplicação Web conhecida como Free Online Pedigree Tool. Ao longo dos próximos parágrafos proceder-se-á a uma análise de algumas aplicações genealógicas, assumindo como critério de selecção não apenas a popularidade que gozam mas também as histórias complexas que podem revelar.

No final da década de 80, a Banner Blue Software lança para sistemas DOS o Family Tree Maker (FTM), actualmente um dos mais populares programas de genealogia, detido pela The Generations Network (TGN) - a empresa que controla o Ancestry.com e o MyFamily.com. A Banner Blue dedicara-se desde a sua fundação, em 1984, à criação de software para a gestão de diagramas em empresas, comercializando um programa denominado Org Plus: "a product that automates the tedious task of creating and editing corporate organization charts"<sup>5</sup>. Compreendendo que muitos utilizadores utilizavam a aplicação para criar as suas próprias árvores familiares, a Banner Blue decide produzir uma aplicação destinada especialmente à área da genealogia. É o início de uma longa história, para a empresa produtora e para a aplicação, que ficará marcada por múltiplas aquisições e que culminará no momento presente enquanto marca da TGN. A sucessão de aquisições tem início em 1995, quando a Banner Blue é adquirida pela Broderbund que, por sua vez, será adquirida pela The Learning Company Inc. em 1998. Logo no ano seguinte, a Mattel compra a The Learning Company Inc. e numa parceria com a A&E Television Networks, Hearst New Media and Technology, Thomas H. Lee Company e Weston Presidio Capital cria uma divisão dedicada à genealogia, a Genealogy.com LLC, que será responsável pelo Family Tree Maker e outros produtos destinados ao campo da genealogia. Em 2001, a A&E Television Networks, detentora dos canais Biography e The History Channel, decide comprar a totalidade da Genealogy.com LLC. Por fim, em 2003, a MyFamily.com Inc., detentora do Ancestry.com, MyFamily.com e RootsWeb.com, compra a Genealogy.com LLC<sup>6</sup> e, com esta aquisição, o Family Tree Maker passa a ser uma marca da MyFamily.com Inc. – empresa entretanto renomeada para The Generations Network.

Apesar deste movimento contínuo de aquisições e de transmissão de propriedade, o Family Tree Maker continua a afirmar-se junto dos utilizadores e a versão 2006, já enquanto marca da TGN, é uma das versões de maior sucesso, recebendo o prémio CODiE Awards na categoria "Best Consumer Productivity Solution". A consulta da loja Amazon.com reflecte também o reconhecimento dos utilizadores que atribuem uma classificação média de 4 estrelas (em cinco possíveis). No entanto, a versão 2008 da aplicação representa uma ruptura no percurso evolutivo da aplicação, implicando a sua rescrita total, recebe uma nova interface e novas formas de interacção. Esta versão é alvo de críticas frequentes por parte dos utilizadores, que lamentam a perda de funcionalidades, a menor resposta do programa e mesmo corrupção de dados. Esta reacção reflecte-se também na avaliação média na Amazon, onde a versão 2008 não atinge as 3 estrelas: 50 utilizadores em 85 atribuíram uma estrela de cinco possíveis. A versão 2009 parece recuperar a confiança dos utilizadores, não só pelas classificações médias atribuídas (aproxima-se novamente das 4 estrelas), mas também pelas posições que consegue no topo de vendas da Amazon: é líder nos EUA onde ocupa os 3 primeiros lugares com as edições FTM Essentials, FTM Platinum e FTM Deluxe, situação que se repete no Reino Unido com as versões FTM Platinum 2009, FTM Deluxe 2009 e o pacote especial "Who Do You Think You Are? Family Tree Maker". Os preços variam de acordo com a versão e edição. A edição FTM

---

<sup>5</sup> [http://familytreemaker.genealogy.com/co\\_bkgnd.html](http://familytreemaker.genealogy.com/co_bkgnd.html)

<sup>6</sup> <http://genealogy.about.com/cs/ukrecords/a/monopoly.htm>

Platinum 2009 é uma das mais dispendiosas na Amazon, listada a 99,99 dólares, antes do desconto atribuído pela loja. As edições Essentials e Deluxe são listadas a 29,99 e 69,99 dólares, respectivamente, antes de desconto. No Reino Unido, a edição Platinum é vendida por 58,71 Libras (38,97 Libras depois do desconto da Amazon).

O título The Master Genealogist (TMG) é desenvolvido, desde 1993, pela Wholly Genes Software, empresa fundada em 1992 por Bob Velke em Columbia no estado de Maryland nos E.U.A. (Clunies, 2002). Além do TMG a Wholly Genes comercializa o Family Tree SuperTools, uma ferramenta dirigida aos utilizadores de outros programas de genealogia, que permite o acesso a novas funções de visualização e gestão da informação. A empresa é ainda responsável pela tecnologia proprietária GenBridge, que permite aceder directamente a informação constante noutras bases de dados de outros programas. A tecnologia, apresentada como uma alternativa ao formato GEDCOM, é utilizada nos vários produtos da Wholly Genes como o The Master Genealogist ou o Family Tree SuperTools, e ainda comercializada como um SDK (Kit de Desenvolvimento de Software) para utilização por outros produtores de software.

O TMG, actualmente na versão 7, é uma ferramenta especialmente complexa, no entanto, poderosa e flexível, frequentemente referida por utilizadores avançados (Reigel 2005). Na última versão, a edição Silver, é vendida online por 34 dólares sob a forma de download ou por 39,95 dólares em CD-ROM, enquanto a edição Gold custa 59 dólares através de download ou 69,95 em CD-ROM.

A MyHeritage.com, cujo website é analisado mais à frente neste capítulo, possui também a sua própria ferramenta genealógica para ambientes Windows, o Family Tree Builder (FTB). A empresa tem sede em Israel e a sua fundação remota ao ano de 2003. Em 2007 foi anunciada a aquisição da Pearl Street Software, empresa responsável pelo programa Family Tree Legends que está na origem do FTB. O Family Tree Builder está disponível gratuitamente, sob a forma de download. O Family Tree Legends ainda é comercializado na Amazon, sob a forma do conjunto Create Your Own Family Tree Genealogy Suite (PC), ocupando o 4º lugar nas vendas da Amazon britânica.

Uma característica que sobressai num primeiro contacto com o FTB é um interface visualmente agradável e amigável. A aplicação incorpora os recursos habituais, como a exibição de árvores, integração de recursos multimédia (fotografias, vídeo), publicação na Web, importação GEDCOM, e também funcionalidades menos comuns, como reconhecimento facial em fotografias, pesquisa baseada em semelhanças fonéticas (Daitch-Mokotoff Soundex) e suporte a múltiplos calendários com conversão automática. No respectivo website são ainda anunciadas funcionalidades para utilizadores mais avançados como a personalização da interface, a gestão de factos em conflito ou a gestão de fontes e citações.

O MacFamilyTree, ou MacStammbaum em alemão, foi criado em 1998 por Stefan Sicurella. O desenvolvimento da aplicação foi posteriormente assegurado pela OnlyMac Software. A Synium Software GmbH, com sede em Mainz, é responsável pela aplicação desde a versão 3, adquirindo controlo total sobre a aplicação na versão 5.7.

Em Novembro de 2007 é apresentada a versão 5 da aplicação.

Surge associada à MacFamilyTree.com, um serviço de alojamento web que permite colocar as árvores familiares online, podendo optar-se pelo acesso público ou reservado.

---

<sup>7</sup> <http://www.macgenealogy.org/2007/08/14/macfamilytree-5-sneak-preview/>

A Leister Productions, Inc., fundada em 1984 por Frank Leister, é a empresa responsável pelo Reunion, uma outra aplicação genealógica para sistemas Mac OS. Até à versão 4 da aplicação coexiste uma versão para Windows, cuja propriedade é vendida à Sierra em 19978. O desenvolvimento da aplicação na plataforma Windows prossegue sob o nome Generations Family Tree, sendo adquirida, em 2002, pela TGN9. A Leister Productions, Inc. prossegue o desenvolvimento da versão para MacOS, actualmente na versão 9, apresentada em Março de 2007. A Leister comercializa ainda uma versão para iPhone/iPod Touch, disponível na App Store da Apple por 11,99 euros. Esta versão permite sincronizar os dados com a aplicação para Mac.

Para além da oferta comercial, é possível encontrar aplicações gratuitas, que podem ser descarregadas livremente da Internet. Podem ser aplicações desenvolvidas por entidades privadas, como o Personal Ancestral File, o Family Tree Builder ou a versão Standard Edition do Legacy, mas também programas desenvolvidos em comunidade. O Gramps é um dos melhores exemplos deste tipo de aplicação, cujo desenvolvimento depende inteiramente de um esforço comunitário e cujo código-fonte é distribuído livremente na Internet. Uma característica habitual neste tipo de programas é a capacidade de execução em múltiplas plataformas. De facto, o Gramps está disponível não só para GNU/Linux, mas também para sistemas MS Windows, Mac OS, BSD ou Solaris.

O PhpGedView é um caso especial dentro conjunto aqui analisado. Trata-se de uma aplicação Web escrita em PHP, cujo principal objectivo é a gestão de websites inteiramente dedicados à genealogia, integrando recursos de importação de GEDCOM e de apresentação de árvores e outros recursos específicos ao tema. É desenvolvida também comunitariamente, segundo as regras para o software livre.

Os exemplos abordados com maior detalhe representam algumas das aplicações mais populares para as três principais plataformas: Windows, MacOS e Linux. Mas outros nomes são igualmente reconhecidos e referidos nas preferências dos utilizadores. A consulta do principal portal e fórum português de genealogia, o Geneall10, revela referências frequentes a aplicações gratuitas como o Personal Ancestral File ou o Family Tree Builder, mas também aplicações comerciais como o The Master Genealogist e o Legacy Family Tree. A utilização de múltiplas ferramentas é igualmente uma opção referida, sobretudo ferramentas gratuitas, que são utilizadas no desempenho de tarefas específicas: desenho de árvores genealógicas, exportação de relatórios ou páginas Web, etc.

Na verdade, a avaliação da popularidade de uma aplicação é difícil de concretizar, quer pelo volume de aplicações disponível, quer pela ausência de estatísticas relativamente a vendas e transferências da Internet. Um inquérito realizado aos internautas do sítio World Vital Records, referido no fórum Geneall, cuja versão original não foi possível recuperar, revelou o The Master Genealogist como o programa mais utilizado (34%), seguido de aplicações como o Family Tree Maker (19%) e o Legacy Family Tree (17%). Na lista surgem ainda o Roots Magic (6%), o Personal Ancestral File (5%), Reunion (4%) e o GenBox Family History (3%).

---

<sup>8</sup> <http://www.eogn.com/archives/news9730.htm>

<sup>9</sup> <http://www.ancestry.com/learn/library/article.aspx?article=6104>.

<sup>10</sup> [http://www.geneall.net/P/forum\\_msg.php?id=186769&fview=c](http://www.geneall.net/P/forum_msg.php?id=186769&fview=c)

Uma outra via na avaliação da popularidade, depende da análise das vendas da loja online da Amazon, dentro da categoria Família & Genealogia. Na verdade, o cruzamento dos dados referidos no parágrafo anterior, com as vendas da Amazon (lojas .com e .co.uk), permite confirmar a popularidade de alguns dos programas. O Family Tree Maker aparece no topo das vendas em várias lojas da Amazon (Estados Unidos, Canadá e Reino Unido) nas suas diferentes versões: Family Tree Maker Essentials, Family Tree Maker Deluxe ou, Family Tree Maker Platinum. No caso norte-americano, as diferentes edições ocupam os três primeiros lugares na tabela de bestsellers. Outras aplicações que surgem no conjunto das 10 mais vendidas são Reunion 9 (aplicação para Mac) em 4º, RootsMagic Family Tree em 5º, e as edições Deluxe 2008 e Essentials 2008 do FTM novamente em 6º e 7º, Personal Historian Software em 9º. Family Tree Heritage Deluxe V7.0 em 10º. No Reino Unido a situação é ligeiramente diferente, embora o Family Tree seja igualmente popular nas vendas da Amazon, ocupando os 2 primeiros lugares com as versões 2009 Platinum e Deluxe.

É possível constatar que a popularidade das aplicações varia entre países. Recorrendo à estratégia anteriormente utilizada, a análise das vendas na Amazon francesa e alemã revela escolhas diferentes por parte dos utilizadores. No caso francês, os utilizadores a preferem aplicações como o Heredis ou o Geneatique, enquanto o caso alemão revela preferência por nomes como Stammbaum e, também, o Family Tree Maker.

Apesar do volume de aplicações actualmente existente, a generalidade apresenta funcionalidades semelhantes, distinguindo-se sobretudo pela presença de opções avançadas destinadas a um público mais exigente. Os utilizadores podem criar árvores familiares, vistas de costados, adicionar fotografias, áudio e vídeo, cruzar fontes e notas, importar e exportar GEDCOM. A introdução do formato GEDCOM representou um passo importante na adopção de software para tratamento de informação genealógica, permitindo a exportação e importação de informação entre diferentes programas. Algumas aplicações apresentam funcionalidades especificamente destinadas à investigação genealógica e visam auxiliar nesse processo, através da gestão de listas de tarefas, apresentando sugestões de progressão na pesquisa e detectando inconsistências nos dados.

Um conjunto tão amplo de aplicações revela, desde logo, a existência de interesse da parte de empresas de desenvolvimento que deverá traduzir-se na existência de um mercado e no interesse dos próprios utilizadores. Avaliar a popularidade de cada nome é uma tarefa complexa, embora as aplicações mais referidas e indicadas como preferidas sejam as comerciais, com valores que variam entre os 30 e 100 dólares e que apresentam funções avançadas de análise, tratamento e exportação da informação.

Estes programas de aplicação foram - e continuam a sê-lo, junto de um público mais dedicado às questões da genealogia e da história familiar - as principais ferramentas informáticas utilizadas em genealogia até ao aparecimento de aplicações baseadas na Web, que são, de alguma forma, auxiliares no crescente volume de informação da World Wide Web.

## 1.2. Os sítios na Web

O aparecimento da World Wide Web, na década de 90, veio revolucionar o acesso à informação e, como num primeiro momento com os computadores pessoais, cedo os aficionados da genealogia descobriram um meio poderoso de pesquisa e partilha de recursos.

Na verdade, a utilização da Internet, como meio de partilha de informação genealógica, deverá ser anterior ao surgimento da Web, ainda que numa escala certamente bem mais reduzida e difícil de comprovar, recorrendo a mailing lists, newsgroups ou discussion groups. A origem das mensagens de correio electrónico remonta ao ano de 1961, com a criação de terminais de acesso remoto Compatible Time-Sharing System (CTSS) no MIT. As primeiras mailing lists<sup>11</sup> electrónicas surgem pouco depois, e a grande rede mundial começa a insinuar-se com a entrada em funcionamento da ARPANET<sup>12</sup> em 1969. A primeira mensagem de uma das mais antigas mailing lists dedicadas à genealogia, a ROOTS-L<sup>13</sup>, actualmente associada ao Ancestry, remonta a 12 de Dezembro de 1987. O conteúdo desta mailing list é pesquisável desde a sua primeira mensagem (apesar do vazio de cerca de 6 anos que se segue ao ano de 1990) e permite evidenciar um crescimento constante até ao final da década de 90, passando da centena de mensagens mensais em 1990, para números na ordem de vários milhares a partir de 1996. O final da década de 90 assinala o início de um período de queda na utilização da *mailing list* que será indício da emergência de uma nova realidade.

Ano	Número de mensagens
1987*	6
1988	361
1989	519
1990*	1500
1996*	22458
1997	45872
1998	32809
1999	23144
2000	16925
2001	15417
2002	13902

\* Faltam alguns meses

**Tabela 1. Volume de mensagens da *mailing list* ROOTS-L.**

Na ainda curta história da Internet, foi indubitavelmente o advento da Web que veio revolucionar o acesso à informação, atraindo um público cada vez mais abrangente. O volume de informação genealógica disponível online continua a crescer: *mailings lists*,

<sup>11</sup> Uma mailing list electrónica, ou lista de correio electrónico, é uma listagem de endereços de correio electrónico criada com o objectivo de facilitar a distribuição de informação subordinada a um determinado tema.

<sup>12</sup> A “Advanced Research Projects Agency” (ARPA) foi criada em 1958 pelo presidente Dwight D. Eisenhower, como uma dependência do Departamento de Defesa Norte-Americano, em plena Guerra Fria, no rescaldo do lançamento do satélite russo Sputnik.

<sup>13</sup> <http://archiver.rootsweb.ancestry.com/th/index/ROOTS-L>

fóruns, sítios pessoais, bases de dados alimentadas pelo Estado e por privados, são apenas alguns exemplos.

A popularidade do tema pode ser avaliada pelo volume de sítios na Web dedicados ao tema. Uma simples pesquisa no Google por “Genealogy” revela 68 milhões de entradas dedicadas ao tema. Os números continuam a impressionar mesmo com pesquisas por “genealogia”, “genealogía” ou “genealogie”. Se considerarmos apenas as páginas escritas em língua portuguesa, é possível encontrar cerca de 1 milhão e 200 mil entradas para o tema.

Portal	Número de entradas global (por idioma)	Número de entradas local (por país)
google.com	68.000.000	-
google.pt	1.200.000	147.000
google.com.br	1.200.000	562.000
google.de	2.040.000	1.730.000
google.es	1.230.000	354.000
google.co.uk	64,100,000	3,790,000
google.fr	18 300 000	6 420 000

Tabela 2. Número de entradas para uma pesquisa por genealogia no Google (2009).

Importa analisar alguns sítios na Web, pelo volume de informação que acumulam e pelo número de visitantes que recebem. Os maiores encontram-se, sem dúvida, nos E.U.A., onde as questões da origem familiar assumem particular relevo, mas também na Europa, em países como Reino Unido, França ou Alemanha. Será ainda tida em consideração a realidade portuguesa.

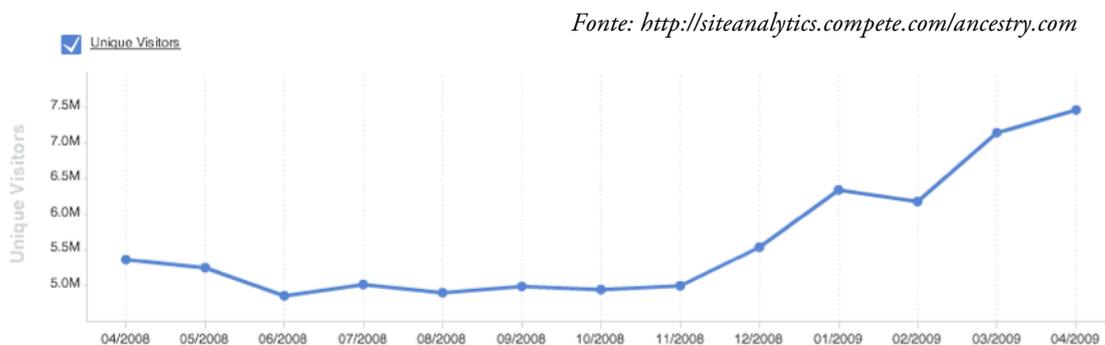
De acordo com o arquivo da Internet Wayback Machine, a primeira versão de um dos mais populares sítios dedicados à genealogia – o ancestry.com - remonta ao ano de 1996 (uma pesquisa no serviço whois da domaintools.com revela o registo do domínio a 16 de Maio de 2005), seis anos após a idealização e criação da World Wide Web (WWW) no CERN<sup>14</sup>. Nessa primeira versão podia aceder-se a várias bases de dados, como a Social Security Index, a American Marriage Records, uma biblioteca geográfica com mais de um milhão de locais (30 mil em 1854), documentação sobre a imigração para os EUA e bibliografia genealógica (dados recolhidos no próprio sítio Web).

O Ancestry.com reúne actualmente um extenso volume de informação, desde censos, registos militares, registos de imigração. Disponibiliza acesso, mediante assinatura mensal ou anual, a um amplo conjunto de fontes<sup>15</sup> dos EUA e também do exterior, possuindo versões do website dedicadas ao Canadá, Reino Unido, Austrália, Alemanha, Itália, França, Suécia e China.

<sup>14</sup> <http://info.cern.ch>;

<sup>15</sup> <http://search.ancestry.com/search/CardCatalog.aspx>;

A este volume maciço de informação está associado um conjunto de utilizadores em constante crescimento. Recorrendo ao serviço da análise de tráfego da Compete.com<sup>16</sup>, verifica-se que o sítio dedicado à genealogia situa-se na posição n.º 156 entre os sítios com maior popularidade, alcançando um volume de visitas mensais superior a sete milhões (dados de Abril de 2009). As estatísticas revelam ainda que, no espaço de um ano, o domínio duplicou o número de visitas, apresentando, em Dezembro de 2007, uma média de 2,5 milhões de visitas por mês. Independentemente da imprecisão no número de acessos, o Ancestry aparece nos diferentes serviços de análise e estatística como o mais popular e requisitado na sua área<sup>17</sup>.



**Figura 1. Visitantes únicos no portal Ancestry.com, entre Abril de 2008 e Abril de 2009.**

O Ancestry é apenas um dos mais antigos e populares sítios na Web dedicados em exclusivo à informação e à pesquisa genealógica. A The Generations Network, empresa que detém o Ancestry, é proprietária de um conjunto de domínios que ultrapassa o milhar<sup>18</sup>, dedicados directamente ou indirectamente à temática da genealogia e história familiar, e alguns dos mais populares na Web. Para além das versões internacionais do Ancestry, detém domínios como o Myfamily.com, um sítio da Web 2.0 analisado mais à frente neste capítulo, e o Rootsweb.com, dedicado mais directamente às necessidades da investigação em genealogia e à enumeração dos recursos disponíveis sobre o tema. De referir ainda que a mesma empresa detém o Familytreemaker.com, sob o qual podemos encontrar a popular aplicação desktop para tratamento de informação genealógica.

No campo dos investimentos estritamente privados é possível encontrar outros exemplos, como o Genealogy Today, da responsabilidade da Genealogy Today LLC.

O genealogytoday.com, lançado em 1999, disponibiliza acesso gratuito a um pacote básico e, mediante assinatura, a serviços mais avançados. Segundo a Compete.com regista cerca de 250 mil acessos mensais.

Uma outra empresa, a World Vital Records, Inc., opera o portal homónimo, onde possibilita o acesso a bases de dados genealógicas mediante uma subscrição mensal de 39,95 dólares, ou uma subscrição anual no valor de 119,95 dólares. O arquivo Wayback

<sup>16</sup> Empresa de análise web: “Compete’s clickstream data are collected from a 2,000,000 member panel of US Internet users”;

<sup>17</sup> 232º com 5,538,622 visitantes por mês (<http://siteanalytics.compete.com/ancestry.com/?src=dt100>);

957º de acordo com o Alexa (<http://www.alexa.com>); 1º de acordo com o top 50 da Progenealogists (<http://www.progenealogists.com/top50genealogy2008.htm>)

<sup>18</sup> <http://whois.domaintools.com/ancestry.com>

Machine revela registos para o [worldvitalrecords.com](http://worldvitalrecords.com) que remontam a 2002 (o domínio foi registado em Junho de 2001) e de acordo com a Compete o sítio acumula 200 mil visitas mensais. As dados de acesso da [domaintools.com](http://domaintools.com) revelam um serviço utilizado essencialmente pelo público norte-americano e, em menor escala, por outros países de língua inglesa: E.U.A. (68,3%), Austrália (10,9%), UK (4,6%), África do Sul (4,4%) e Canadá (3,3%)

O interesse pela genealogia nos EUA é particularmente visível na comunidade Mórmon (Igreja dos Santos dos Últimos Dias) e no empenho com que tratam a história familiar. A instituição desenvolveu um trabalho de recolha de registos, que ultrapassou largamente as fronteiras dos Estados Unidos, e utilizou essa informação para alimentar os múltiplos centros de história da família que tem espalhados pelo mundo, e também sob a forma de bases de dados online acessíveis através do serviço [FamilySearch.org](http://FamilySearch.org). O portal aproxima-se do milhão mensal de visitantes provenientes essencialmente dos Estados Unidos (61,5%), mas também utilizadores do Reino Unido (10,7%), Alemanha (5,6%), Austrália (5,4%) e Canadá (3,6%).



Figura 1. Visitantes únicos no portal [FamilySearch.org](http://FamilySearch.org), entre Maio de 2008 e Maio de 2009.

Fonte: <http://siteanalytics.compete.com/familysearch.org/>

Na Europa o esforço de digitalização da informação genealógica tem sido realizado a dois grandes níveis, o estatal e o privado, ainda que com especial propensão para o primeiro. A França e o Reino Unido representam dois casos onde a digitalização e disponibilização online da informação tem sido mais abordada.

A Inglaterra disponibilizou recentemente o acesso online e integral a uma base de dados pesquisável, com os censos de 1911<sup>19</sup>, acessível em [1911census.co.uk](http://1911census.co.uk). Este trabalho resultou de uma parceria entre a entidade pública The National Archives e a Brightsolid, empresa privada que detém actualmente o portal [Findmypast.com](http://Findmypast.com). A mesma empresa disponibiliza acesso a outros censos, registos paroquiais, registos militares, entre outros. O [Findmypast.com](http://Findmypast.com) é, de resto, apresentado como um dos mais completos portais de informação genealógica no Reino Unido: “is a family history and genealogy website based in London, England and holds over 650 million family history records”.

A origem do [Findmypast.com](http://Findmypast.com) remonta a 2002, sob a designação [1837online.com](http://1837online.com). O portal foi desenvolvido pela Title Research Group que desde Abril de 2003 disponibiliza online todos os registos paroquiais da Inglaterra e Gales, desde 1837 até ao presente: “Following the transcription, scanning and indexing of over two million images, the company launched the first website to allow the public easy and fast access to the complete indexes, which

<sup>19</sup> <http://www.brightsolid.com/news/recent-news/brightsolid-launches-1911-census-online1>

until then had only been available on microfiche film in specialist archives and libraries”<sup>20</sup>. Em Novembro de 2006 o 1837online.com passa a ser conhecido como Findmypast.com e em Dezembro de 2007 a Scotland Online (actual Brighsolid), uma empresa fornecedora de acessos e soluções Internet, compra o portal Findmypast.com<sup>21</sup>. O volume de registos digitalizados acessíveis online continuou a crescer, com o portal a anunciar presentemente a disponibilização de registos que remontam a 1664. O acesso à informação é efectuado mediante assinatura ou pagamentos pontuais.

A Brightsolid ltd. é ainda responsável pelo desenvolvimento do portal ScotlandsPeople.gov.uk, resultante de uma parceria entre General Register Office for Scotland, National Archives of Scotland e Court of the Lord Lyon, onde é possível consultar registos paroquiais e civis, dados de censos e outros registos.

A versão britânica do Ancestry – ancestry.co.uk – constitui também uma importante fonte de informação para o Reino Unido, permitindo o acesso a registos paroquiais, dados dos censos, informações militares, registos de emigração e imigração, listas telefónicas, entre muitos outros. O modelo de acesso à informação é o mesmo do seu congénere norte-americano. De acordo com a Alexa, a maioria dos utilizadores do website são provenientes do Reino Unido (74,2%) e dos E.U.A. (11,0%), ocupando a posição número 525 no Reino Unido, entre os websites mais visitados.

A França é também um exemplo revelador do interesse pelas questões genealógicas, onde é possível encontrar projectos desenvolvidos por privados e pelo Estado. O portal culture.fr, da responsabilidade do Ministério da Cultura francês, integra um motor de pesquisa que permite questionar múltiplas bases de dados genealógicas, e é um exemplo de acção estatal no sentido de levar a genealogia a um público mais abrangente.

A Fédération Française de Généalogie, uma associação que agrupa associações de genealogistas amadores, disponibiliza um portal nacional de pesquisa, o bigenet.org. Neste caso, o acesso às fontes é possível apenas através de pagamento de um determinado valor.

Na Web francesa podemos encontrar ainda outros projectos de carácter privado, como o caso do www.genealogie.fr da responsabilidade da Notrefamille.com SA, entidade existente desde 1994 e responsável pela disponibilização de serviços online como o Cadeaux.com ou o Notrefamille.com. No genealogie.fr a estratégia é idêntica à de outros portais privados, cobrando o acesso às fontes, através de um pagamento pontual ou de uma mensalidade (14,99 por um mês; 9,99 durante 3 meses; 8,99 durante 6 meses). O portal anuncia o acesso a um amplo conjunto de informação, proveniente de arquivos ou da responsabilidade de outros utilizadores.

Em Portugal, tem sido desenvolvido algum trabalho esporádico e o pontual de digitalização de fontes de informação genealógica. O Arquivo da Universidade de Coimbra procedeu, recentemente, à digitalização e colocação online dos registos paroquiais que constam no seu acervo. O Arquivo de Braga disponibilizou, durante algum tempo, o acesso a uma base dados com informação relativa às Inquirições de Genere, entretanto suspensa devido à “incapacidade na manutenção dos serviços associados”<sup>22</sup>. Nas Universidades, e resultado de investigações académicas, tem sido feito trabalho de recolha e organização de

---

<sup>20</sup> <http://www.findmypast.com/media/news/national-burial-index-goes-live.jsp>

<sup>21</sup> <http://www.findmypast.com/media/news/findmypast-scotland-online-family-history.jsp>

<sup>22</sup> <http://www.adb.pt/Default.aspx?tabid=9&lang=pt-PT>

informação, resultando, em alguns casos, em bases de dados online como a do NEPS23, ou o protocolo desenvolvido entre a câmara de Óbidos e a UC, que permitiu a transcrição (não apenas a digitalização dos documentos) e introdução da informação dos registos paroquiais numa base de dados.

O principal portal português dedicado à genealogia é o Geneall.net, anteriormente conhecido por Genea Portugal, tem por objectivo a disponibilização online de informação relativa às famílias portuguesas – inclui outras famílias europeias –, estabelecendo a ligação com os descendentes de D. Afonso Henriques, Carlos Magno, Fernando I o Magno, Hugo Capeto ou Guilherme I. O Genea Portugal foi criado em Abril de 2000, em resultado de uma parceria com a Portugal, resultando na sua integração enquanto canal do portal Sapo, como é possível esclarecer no artigo “Genea Portugal e o Futuro” disponível no Geneall.net<sup>24</sup>. O acordo com a PT Comunicações termina seis anos depois, apesar do aparente sucesso que o canal parece ter conseguido: de acordo com o mesmo artigo o canal encontrava-se entre os 5 mais visitados do portal e aproximava-se das 5 milhões de páginas vistas por mês. Após a cisão com o Sapo, o projecto prosseguiu sob a actual designação, permitindo o acesso à informação mediante uma assinatura. De acordo com a Alexa, 32% dos utilizadores do portal são provenientes de Portugal, seguidos pelos provenientes do Brasil, 15%.

Não existe, porém, qualquer iniciativa centralizada, de grande escala, privada ou pública, no sentido de valorizar o acesso a este património sob a forma de bases de dados online.

### 1.3. As redes sociais online

A World Wide Web sofreu ao longo dos primeiros anos do novo milénio, um conjunto de transformações que se reflectiram também no mundo da genealogia online. Assiste-se à constituição de um novo modelo de Web (Forrest, 2006) – a Web 2.0 (O'Reilly, 2005) – com um grande poder de captação de utilizadores, em alguns casos na ordem das centenas de milhão (Eldon, 2008), modelo que é caracterizado pelo entendimento da Web enquanto plataforma e pelo uso ferramentas novas e mais poderosas (wikis, blogs, partilha de fotografias, etc), que apresentam um enfoque especial nos aspectos cooperativos e na constituição de redes sociais. Assistiu-se ao aparecimento de redes temáticas, onde a relação entre os indivíduos é dominada por preferências pessoais (musicais, bibliográficas, etc), pelas suas actividades profissionais ou mesmo pelos laços familiares.

As redes sociais firmadas na Web conseguiram verdadeira projecção e popularidade nestes primeiros anos do século XXI, muito embora seja possível encontrar algumas experiências de efeito limitado ainda na segunda metade da década de 90 como a “Six Degrees” ou a “SocialNet” (Rivlin, 2006). Um dos primeiros serviços desta vaga de rede sociais dá pelo nome de “Friendster” e surge em 2002, alcançando em poucos meses cerca de 3 milhões de registos, tendo sido alvo de uma tentativa de aquisição pela Google, ainda em 2003 (Rivlin, 2006). Estava aberto o caminho para o aparecimento de outros serviços como o MySpace apresentado em 2003, o Orkut criado na Google em 2004, o Facebook que surge igualmente em 2004, Flickr fundado no mesmo ano, entre muitos outros<sup>25</sup>.

---

<sup>23</sup> <http://www.neps.ics.uminho.pt/>

<sup>24</sup> <http://www.geneall.net/P/article.php?id=112>

<sup>25</sup> Wikipedia contributors, “List of social networking websites”, Wikipedia, The Free Encyclopedia.

[http://en.wikipedia.org/wiki/List\\_of\\_social\\_networking\\_websites](http://en.wikipedia.org/wiki/List_of_social_networking_websites) (28.05.2009)

Surgiram também novas redes sociais temáticas, dedicadas à genealogia e à exploração dos laços familiares e assistiu-se à integração de funções características do novo conceito de Web em projectos anteriores.

O Myfamily, detido pela The Generations Network, anteriormente referida como responsável pelo Ancestry, é um dos mais antigos e populares sítios dedicados à partilha de informação genealógica entre utilizadores. Disponível no ciberespaço desde 1998, sofreu recentemente uma remodelação significativa, procurando aproximá-lo das características da Web 2.0. É importante notar que, embora não disponibilize informação de forma directa, os utilizadores podem ter acesso aos serviços do Ancestry, que funciona como repositório de informação.

Avançando pelo campo das redes sociais, podemos encontrar na Web mais exemplos dedicados exclusivamente à genealogia. O Geni.com é considerado um caso de sucesso, tendo sido eleito um dos 50 websites do ano pela revista Time. Com níveis de crescimento elevados desde a sua criação em Janeiro de 2007, possuía 2 milhões de utilizadores e 30 milhões de perfis em Setembro de 2007 (Kuo, 2008): "Geni is creating a family tree of the whole world", segundo o director-geral David Sacks. O geni.com é mais um serviço inteiramente dependente dos utilizadores, que são incitados a construir as suas árvores genealógicas e a convidar novos utilizadores. Da mesma forma, é possível o acesso aos habituais recursos, como a partilha de fotografias e vídeos, a troca de comentários, os alertas de eventos, etc. Ainda, segundo David Sacks, o modelo de negócio do website passa pelo cruzamento entre modelos de assinatura (Geni Pro por 4.95 dólares mensais), comercialização de presentes para a família e em última análise a publicidade (Kuo, 2008).

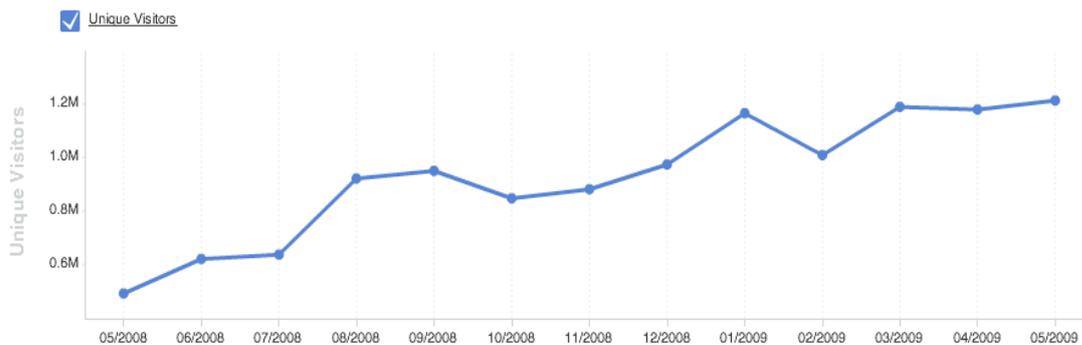
O serviço conhecido em Portugal como Meusparentes.com.pt integra uma rede de sítios com origem no alemão Verwandt.de. A dinâmica do serviço é muito semelhante à que caracteriza o Geni, partindo do mesmo pressuposto, a criação de árvore genealógica, e tendo como base de o convite de familiares que podem auxiliar na recriação dessa árvore familiar. Para além dos habituais perfis de utilizador, da partilha de fotografias e comentários, destacam-se algumas funcionalidades adicionais, como a possibilidade de análise da distribuição geográfica dos apelidos, a consulta da distribuição no espaço dos indivíduos de determinada família e uma secção de introdução à pesquisa genealógica, onde são apresentados alguns conceitos e conselhos de pesquisa.

O Geni e os meusparentes são *websites* inteiramente dependentes do esforço colaborativo e permitem apenas pesquisas sobre informação introduzida pelos outros utilizadores. Não possuem fontes de informação próprias decorrentes de trabalho de arquivo, nem integram motores de busca cruzada. Neste campo, um outro portal, o Myheritage.com, representa um caso peculiar.

O Myheritage.com reúne as características de um motor de pesquisa genealógica, permitindo realizar pesquisas em múltiplas bases de dados e associando ferramentas habitualmente presentes em programas de aplicação - criação de árvores familiares, armazenamento de fotos, vídeos e documentos e calendário. Sem assumir-se explicitamente como um sítio Web 2.0, o MyHeritage permite organizar a informação genealógica através de árvores familiares, agregar fotos, vídeos e documentos em torno do conceito de página familiar. O utilizador pode então convidar outros membros da família,

que auxiliem com a introdução de novos dados, e partilhar a informação com outros utilizadores. A consulta do serviço Analytics da Compete revela um número próximo do milhão de visitantes mensais, com uma clara tendência de crescimento ao longo do último ano. No sítio oficial é apresentado um número superior a 31 milhões de membros e 333 milhões de nomes. O portal está disponível nos principais idiomas mundiais: Inglês, Chinês, Espanhol, Português, Francês, Japonês, Russo, entre muitos outros. Uma característica particular do MyHeritage.com é a dispersão global dos utilizadores do serviço: E.U.A. ( 29,7%), Federação Russa (5.8%), Reino Unido (5.6%), Itália (4.4%), Índia (3.7%), França (3.4%) e outros (47,4%).

*Fonte: <http://siteanalytics.compete.com/>*



**Figura 2. Visitantes únicos no portal MyHeritage.com, entre Maio de 2008 e Maio de 2009.**

## 2. MHK: um sistema para a reconstituição de comunidades históricas

A segunda metade do século XX assistiu à fundação de uma nova corrente de investigação histórica alicerçada na análise nominal e que ficou conhecida como Micro-história. Esta nova corrente, em ruptura com uma História Social francesa de carácter macroscópico, adopta uma tradição antropológica que privilegia a comunidade e defende que "a informação nominal dispersa em muitos e variados documentos pode ser usada para reconstituir a história de vida de muitas pessoas" (Carvalho, 2001).

Neste novo tipo de abordagem o historiador é colocado perante um duplo desafio: o volume de informação, que pode ser arreatador, e o reconhecimento de que, mesmo assim, uma parte, mais ou menos significativa, da informação poderá não ter chegado aos nossos dias. Uma comunidade da época moderna que envolva cerca de 3 000 pessoas produz um volume de informação na ordem das 10 mil referências nominais por ano, que se eleva para um milhão de referências, considerado o período de um século (Carvalho 2001). No caso português esta informação está dispersa pelas mais diversas fontes - registos paroquiais, registos notariais, processos de inquirição, inquirições de género, entre outros - e arquivos – Arquivo Nacional da Torre do Tombo, arquivos distritais sob a tutela da Direcção Geral de Arquivos (DGARQ), das Universidades ou mesmo de instituições privadas. O próximo capítulo apresenta este aspecto com maior detalhe.

Perante um volume de informação tão volumoso, parece natural e inevitável o recurso a meios informáticos que permitam o tratamento de informação em grande escala. A utilização de meios informáticos representa, por si só, um desafio, pois exige a "formalização, representação dos dados, modelização de processos de tratamento da informação e manuseamento inteligível de resultados enormemente complexos" (Carvalho, 1997).

É neste contexto particular que surge o MHK enquanto ferramenta de tratamento documental e análise histórica. De resto, o desdobramento da sigla permite revelar, à partida, um pouco da natureza da aplicação: a sigla MHK resulta da expressão "Micro-History with Kleio" e, numa visão muito sintética, é um sistema informático, da autoria de Joaquim Carvalho da Universidade de Coimbra, destinado à análise micro-histórica, recorrendo a uma notação de transcrição de documentos, da autoria de Manfred Thaller da Universidade de Colónia, conhecida por Kleio.

De acordo com o enunciado no capítulo terceiro – Métodos – da dissertação "Comportamentos morais e estruturas sociais numa paróquia de Antigo Regime" (Carvalho, 1997), o MHK enquanto sistema dimana de um conjunto de princípios que implicam a primazia da fonte, a flexibilidade na representação dos dados, a identificação reversível e *a posteriori* das ocorrências e ainda mecanismos avançados de visualização de dados. A presença destes requisitos permite compreender mais convenientemente o sistema que se clarifica ao longo dos próximos pontos, nos seus principais componentes: a base de dados, a inserção de dados e a interface exploratória.

### 2.1. Base de dados (SGBD)

Identificado o grande desafio – tratar informaticamente comunidades históricas, a partir de um conjunto diversificado de documentação histórica – seria necessário definir "uma estrutura flexível de base de dados para suportar a informação recolhida durante uma investigação nominal" (Carvalho, 1997).

Antes de avançar na estrutura formal da base de dados propriamente dita, e para melhor poder compreendê-la, é importante recordar alguns dos princípios subjacentes ao processo de formalização: a primazia à fonte, a flexibilidade da representação dos dados, identificação reversível e automática.

A orientação às fontes, como oposição à orientação aos resultados, é preferida de forma a permitir reversibilidade e a reconstrução do percurso efectuado, tendo como condicionante exigir mais "tempo, requisitos técnicos e sofisticação formal" (Carvalho, 1997).

A identificação reversível é um aspecto fundamental de toda a arquitectura. Não só torna possível a reconstituição das histórias de vida dos indivíduos em estudo como permite anular ou reverter qualquer decisão entretanto tomada.

Na modelação de dados seguiu-se um método conhecido por "entidade-relacionamento estendido" ("Extended Entity Relationship Model ou EER"), uma representação abstracta e conceptual utilizada na formalização das estruturas de dados, que estende o modelo entidade-relacionamento original.

O modelo entidade-relacionamento convencional é baseado em entidades<sup>26</sup> que podem ser descritas por determinados atributos e estabelecer relações entre si.

Na prática, uma determinada pessoa é uma entidade, caracterizada por certos atributos, como o nome, o local de nascimento ou a residência, que se relaciona com outras entidades (pessoas ou instituições) através de relações familiares, económicas ou de sociabilidade, por exemplo.

O modelo estendido baseia-se nos mesmos princípios do modelo entidade-relacionamento tradicional, com a adição do conceito de especialização. Desta forma, uma entidade pode ser um caso particular de uma superior, estabelecendo-se uma relação de especialização. "A relação de especialização é uma relação abstracta, que estrutura a representação da informação mas que não corresponde a nenhum facto real" (Carvalho, 1997). Uma entidade especializada "mantém os mesmos atributos que a entidade de nível mais geral e acrescenta alguns específicos do seu sub-tipo" (Carvalho, 1997). No MHK, um aforamento ou um emprazamento são casos particulares ou sub-tipos de escritura que é a entidade mais geral.

Apresentada a estrutura conceptual do modelo de dados é necessário passar à sua implementação prática. A implementação seguida no MHK segue os princípios do modelo relacional que utiliza tabelas para representar a informação. Por exemplo, "uma entidade do tipo pessoa, para a qual se registam os atributos: nome, data de nascimento e data de morte (...) seria representada por uma tabela denominada "pessoas" com três colunas: "nome", "data de nascimento", "data de morte". Cada linha da tabela representa uma pessoa concreta. Cada coluna representa um atributo." (Carvalho, 1997).

Uma entidade, ou pessoa neste caso, pode ser descrita por dois tipos de atributos: aqueles que são fixos no tempo e espaço e aqueles que são susceptíveis de alteração. Este segundo tipo, que representa a maioria dos atributos registados no MHK, exige um tratamento especial. Na verdade, apenas o "id" e "sexo" são tratados como atributos fixos e representados directamente na tabela de pessoas. Os restantes são representados numa tabela independente, a tabela de atributos.

---

<sup>26</sup> Neste contexto particular, uma entidade é "algo que existe ou existiu e pode ser representado num computador" (Carvalho, 1997)

A estrutura da base de dados utilizada pelo MHK é composta por três grandes zonas. Uma primeira zona (A) onde consta a informação sobre entidades, atributos e relações, e que é relativamente estável, uma zona de especialização (B) que vai evoluindo ao longo do tempo e, finalmente, a zona de registo da meta-informação (C) que contém a informação que descreve a própria base de dados<sup>27</sup>.

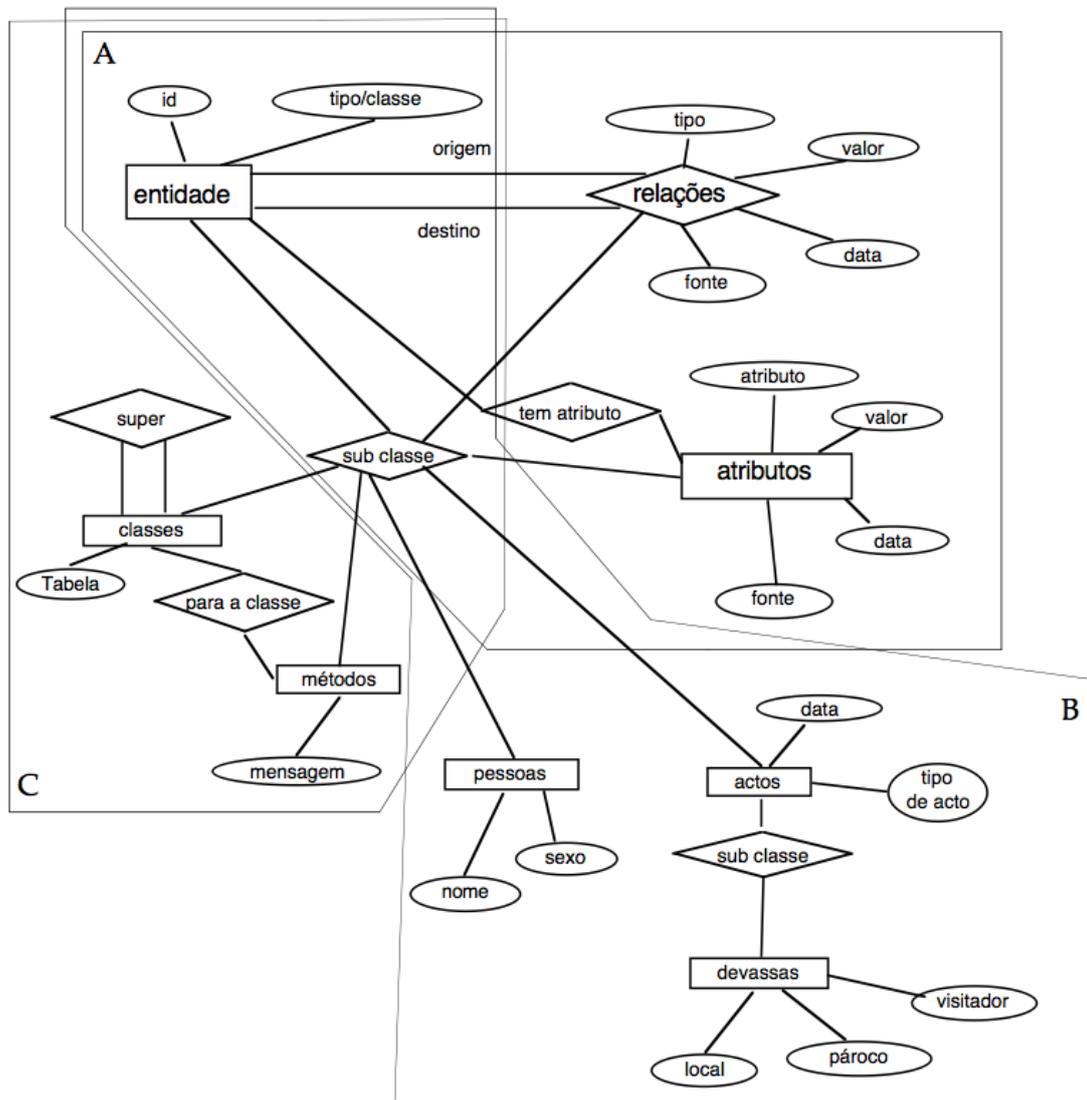


Figura 3. Estrutura de dados (Carvalho, 1997)

Para concluir, importa referir que a estrutura de dados apresentada é concretizada com recurso ao servidor de bases de dados MySQL, embora este modelo possa ser concretizado com recurso a outros servidores.

<sup>27</sup> A área de meta-informação regista as entidades existentes na base e como se relacionam entre si. Foi, segundo Joaquim Carvalho, a forma encontrada para assegurar a evolução da base de dados sem grandes esforços de reprogramação, simplificando o processo de adição de novas entidades e fontes.

## 2.2. Inserção de dados (Kleio)

Apresentada a estrutura da base de dados, é útil proceder à explicação dos mecanismos de inserção de informação adoptados no MHK. A estrutura definida, com os seus propósitos de flexibilidade e simplicidade, implicou um nível de formalização que a afastou inevitavelmente das fontes históricas. Sabe-se, ainda, que o investigador em História interessa-se e tem à sua disposição uma multiplicidade de documentos cuja estrutura pode ser muito variável.

A solução para esse duplo desafio – o carácter formal da base de dados e a complexidade das fontes - implicou a adopção de uma estratégia que permitisse, por um lado, dar primazia à fonte e, por outro, automatizar, tanto quanto possível, o processo de "normalização". Prescindiu-se dos habituais mecanismos de inserção de dados, como os formulários, e optou-se pela transcrição dos documentos históricos em simples documentos de texto, respeitando uma sintaxe e uma notação específicas. Deste modo, tornou-se possível a transcrição relativamente próxima da fonte, procurando o mínimo de acção do operador e remetendo as decisões mais importantes para a fase posterior de consulta e extracção da informação.

A linguagem de transcrição utilizada é conhecida por Kleio e foi desenvolvida a partir do final da década de 70, sob a orientação de Manfred Thaller, no Max-Planck Intitute für Geschichte em Göttingen, na Alemanha. Esta linguagem, de aprendizagem relativamente simples e de leitura fácil, era parte integrante de um sistema mais amplo de análise histórica desenvolvido na mesma instituição e apresentado sob a mesma designação.

A nível conceptual, a linguagem Kleio apresenta pontos de contacto com a estrutura de dados apresentada no ponto anterior e, inevitavelmente, com os modelos de entidade-relação. A linguagem assenta em "grupos" e "elementos" que, de algum modo, podem ser correlacionados com as "entidades" e "atributos". Um grupo é a "unidade principal" que pode ser descrita por múltiplos elementos. À semelhança das entidades que especializam outras entidades, também um grupo pode conter outros grupos, até ao nível superior ou de topo designado por documento.

```
kleio$gacto.str/prefix=stmaria
  fonte$baptismos1562-1563/localizacao-a.d.l./freguesia de santa maria de
  obidos

  bap$b1562-1/20/12/1562/fl.1/igreja de santa maria/manuel
  frois(cura)/obs= assina manuel frois (cura)
  n$luzia/f/id=b1562-1-per1
  pn$jordao fernandes/id=b1562-1-per2
  mn$brites fernandes/id=b1562-1-per3
  pad1$afonso fernandes/id=b1562-1-per4
  mad1$catarina andre/obs=ha mais informacao, no entanto nao se le por
  estar rasgado/id=b1562-1-per5
```

Figura 4. Exemplo de uma transcrição de registo paroquial (acto de baptismo).

Na prática, esta hierarquia de grupos (entidades), compostos por elementos (atributos) e valores, permite a criação de formatos de transcrição que se adaptam a um conjunto alargado de fontes históricas, provenientes dos mais variados períodos históricos: registos

paroquiais, escrituras, escambos e devassas, entre outros. A criação de formatos de transcrição para estas fontes é realizada através de um ficheiro de texto (gacto2.str) que inclui as instruções necessárias ao tradutor genérico existente no MHK.

Após a transcrição dos documentos, com recurso a um certo nível de codificação e marcação, é necessário invocar a ferramenta que irá realizar a normalização da informação, de acordo com os requisitos da base de dados. Esta tarefa é desempenhada por um tradutor, no caso um tradutor desenvolvido numa linguagem conhecida por Prolog<sup>28</sup>, que é invocado a partir da interface gráfica. Do processo de tradução resulta informação normalizada em formato XML que poderá ser importada na base de dados. Este processo de tradução e importação da informação, com os respectivos relatórios, é controlado através da interface gráfica que será analisada no próximo ponto.

O tradutor é uma peça fundamental na arquitectura do MHK e é resposta encontrada para o duplo desafio enunciado anteriormente, tornando possível "uma metodologia de transcrição que produz textos muito legíveis e próximos da fonte original e uma estrutura muito formalizada de dados" (Carvalho, 1997).

### 2.3. A interface gráfica de utilizador (GUI)

Uma vez desenhada a estrutura da base de dados e introduzidos os documentos históricos transcritos é necessário extrair a informação, idealmente de forma simples mas flexível. Nos actuais sistemas informáticos este trabalho fica, habitualmente, a cargo da Interface Gráfica do Utilizador (GUI na sigla inglesa) que faz a ponte entre o servidor de bases de dados e o utilizador final. A interacção com o MHK processa-se essencialmente através de um *browser* Web alimentado a partir de um servidor que pode ser instalado localmente, desta forma apenas visível ao computador do utilizador ou, em alternativa, instalado num servidor e acessível a partir da Internet.

O MHK está estruturado graficamente em torno de secções para consulta, importação e administração. A partir de uma página inicial o utilizador pode seleccionar uma secção a que terá acesso após o início de sessão com um utilizador e senha.

---

<sup>28</sup> Prolog é uma linguagem de programação de alto nível (bastante afastada do código binário ou código de máquina), cujo desenvolvimento remonta aos anos 70, e que implica a aplicação dos princípios da lógica matemática (PROgramação em LÓGica). É uma linguagem essencialmente declarativa, em que a solução de um dado problema depende da declaração dos factos e relações conhecidos a priori, por oposição às linguagens procedimentais, que procuram fornecer os passos necessários para a resolução do problema. Nas instalações do MHK é utilizada uma implementação desta linguagem conhecida por SWI-Prolog, da responsabilidade de Jan Wielemaker, investigador na Universidade de Amesterdão no Human Computer-Studies Laboratory (HCS), anteriormente conhecido por Sociaal-Wetenschappelijke Informatica (SWI).

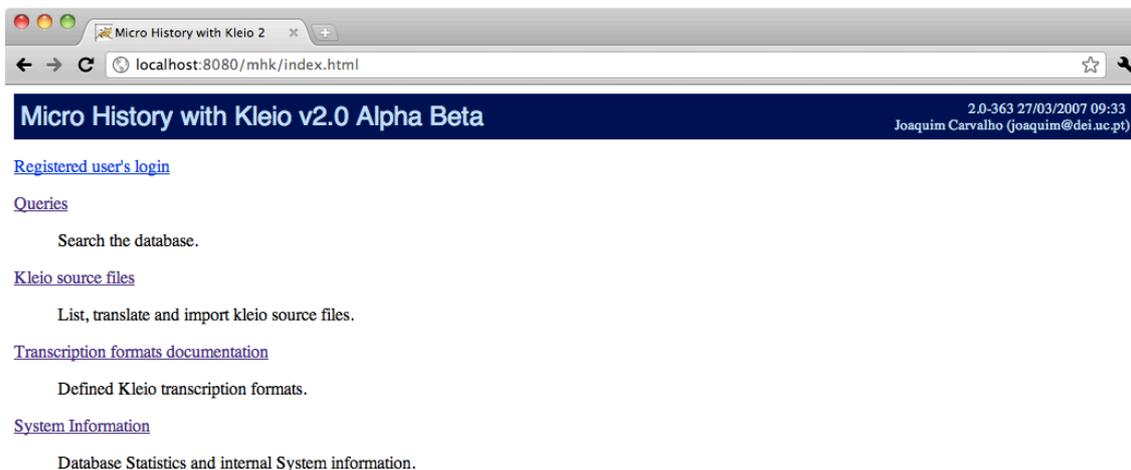


Figura 5. Página de entrada

A pesquisa na base de dados faz-se a partir da ligação Search. Nesta secção o utilizador pode explorar a base de acordo com os atributos (residência, título, etc) e os tipos de relações (familiar, económica, etc) introduzidos e efectuar consultas SQL directamente sobre as tabelas.

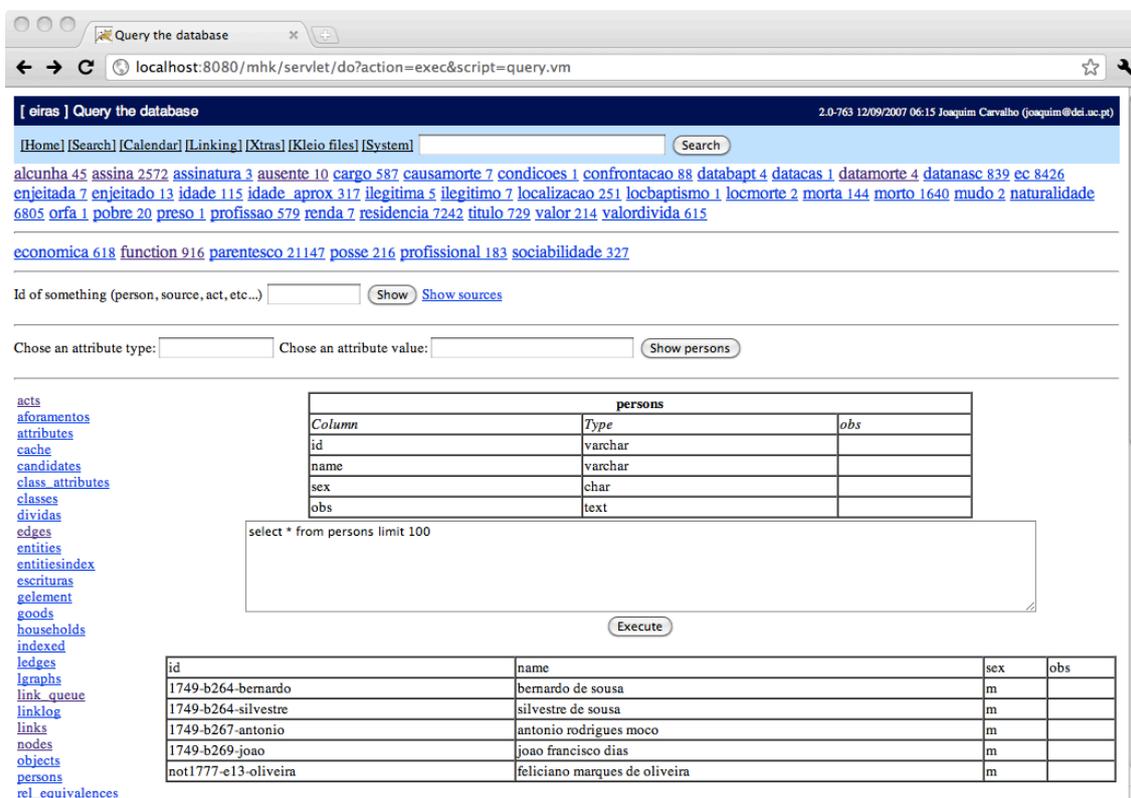


Figura 6. Página de Pesquisa

A vista de Calendário oferece uma outra possibilidade de exploração dos dados. Seleccionando um determinado dia do mês e ano podem consultar-se todos os actos para essa data, isto é, os acontecimentos que ficaram registados nas fontes históricas.

A partir de Kleio Files (Ficheiros Kleio) realizam-se todas as operações de tradução e importação das fontes transcritas, de acordo com os métodos descritos no ponto anterior. Aqui o investigador controla o processo de tradução e importação para a base de dados

através da consulta de relatórios onde são discriminados possíveis erros. Cada linha corresponde a um documento Kleio (\*.cli) e apresenta as operações possíveis: traduzir e consultar relatório de tradução, importar e consultar relatório de tradução. Permite ainda carregar ficheiros Kleio a partir de uma qualquer localização no disco do utilizador.

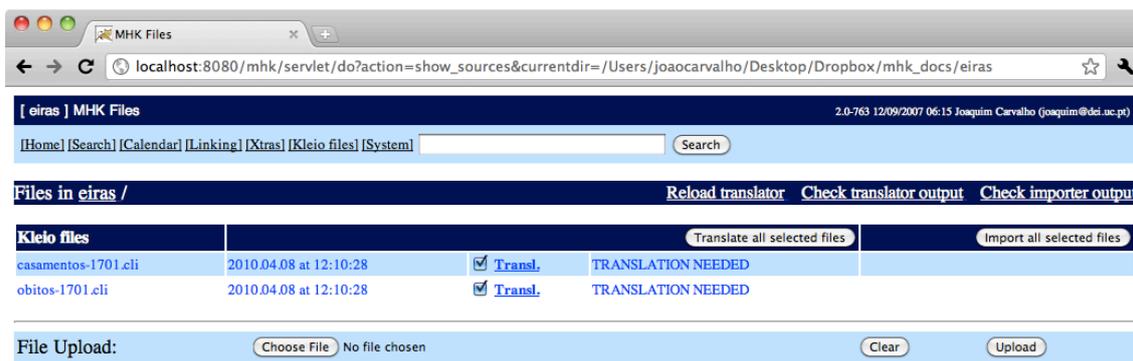


Figura 7. Ficheiros Kleio

Finalmente, uma secção de administração que permite visualizar estatísticas da base de dados e efectuar algumas operações de manutenção.

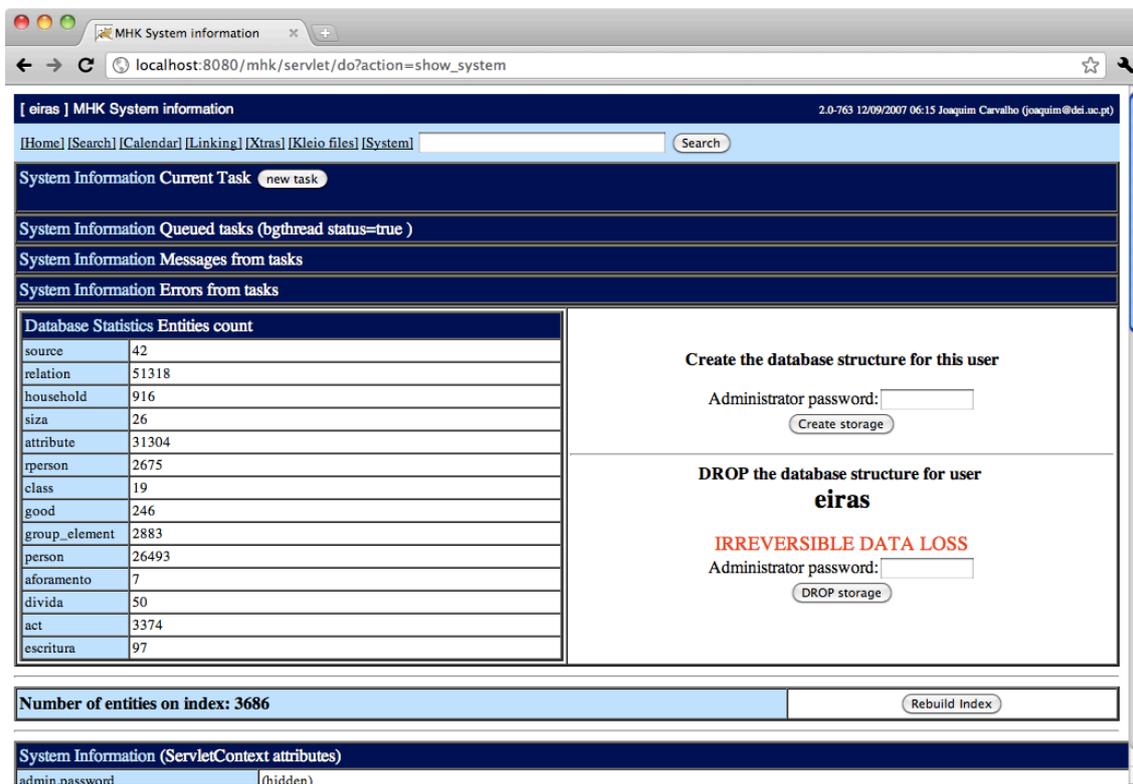


Figura 8. Página de Sistema

## 2.4. Algumas bases de dados: exemplos concretos

Até ao momento em que se iniciou o processo de transformação descrito no presente trabalho, o MHK havia sido utilizado, essencialmente, em trabalhos de investigação na Faculdade de Letras da Universidade de Coimbra. Destes trabalhos de investigação resultaram várias bases de dados que agregam diferentes tipos de informação documental. Os registos paroquiais são o tipo de documento tratado com frequência, mas não constituem o único exemplo de fontes abordadas, onde se inclui também documentação

medieval – cartas régias, emprazamentos, testamentos, escambos, entre outros –, e registos notariais da Época Moderna.

A paróquia de Soure foi a primeira comunidade estudada com recurso ao MHK e é o exemplo prático utilizado por Joaquim Ramos de Carvalho na conceptualização e desenvolvimento do sistema (Carvalho, 1997).

Outro estudo de doutoramento (Campos, 2003) foi levado a cabo por Rosário Campos, na vila da Lousã, e constitui um exemplo particularmente interessante e representativo do tipo de informação que a abordagem micro-histórica permite recuperar (Carvalho 2003). Neste estudo procurou avaliar-se o impacto social do estabelecimento de uma fábrica de papel na Lousã do século XVIII.

A vila de Eiras, actualmente uma freguesia de Coimbra, foi alvo de um estudo de mestrado, desta vez conduzido por Ana Isabel Ribeiro (Ribeiro, 2005), tendo por base não só registos paroquiais, mas também registos notariais e róis de confessados.

Um projecto de colaboração entre a Câmara Municipal de Óbidos e a Universidade de Coimbra tornou possível a constituição de uma das bases de dados mais completas realizadas até ao momento. O trabalho, desenvolvido ao longo de vários anos, envolveu a transcrição de todos os registos paroquiais de Óbidos, desde que há registos até finais do século XVIII. É, por ora, o exemplo mais revelador do volume de informação que uma base de dados e aplicação deste tipo pode armazenar. Esta base de dados ultrapassa as 100 mil ocorrências isoladas, entre baptismos, casamentos e óbitos, onde estão identificadas mais de 6 mil pessoas que reúnem 58.897 ocorrências.

Estes projectos, acompanhados pelas respectivas bases de dados, são um importante testemunho do sucesso da MHK em trabalhos de investigação e dos relevantes conjuntos de informação histórica que permite tratar. Para além dos números, há ainda a diversidade de fontes, como se procurará explorar no próximo capítulo.

### 3. A informação genealógica

Numa perspectiva de abertura da aplicação a um público mais alargado, onde as questões da reconstituição das famílias são fundamentais, importa agora abordar algumas questões referentes à genealogia, em particular as fontes históricas de maior relevância para o tema. "Em princípio só interessam à genealogia fontes que mencionam pessoas concretas, com seus nomes e outras formas de identificação (...) que refiram o parentesco destas pessoas com as outras" (Amaral, 2000).

Para períodos mais recuados, as fontes disponíveis reportam-se sobretudo a famílias nobres ou com alguma notoriedade, de que são exemplo os Livros de Linhagens. A partir do século XVI, a informação disponível alarga-se às restantes famílias, mesmo sem notoriedade. Ao longo da época Moderna, com o importante impacto das normas do Concílio de Trento, o volume de documentação adensa-se com registos paroquiais e também registos notariais, processos de inquisição, inquirições de género, entre outros.

O Arquivo Nacional da Torre do Tombo possui na sua página online<sup>29</sup> uma secção dedicada à genealogia onde identifica os principais núcleos e colecções documentais com interesse para um trabalho de pesquisa em genealogia.

Há, ainda, pelo menos um exemplo de uma biblioteca dedicada exclusivamente ao tema, a Biblioteca Genealógica de Lisboa, aberta ao público desde 2006. Além de reunir um acervo próprio, acessível ao público nas suas instalações na Calçada do Marquês de Abrantes, procura constituir uma base de dados, acessível online<sup>30</sup>, com os títulos disponíveis nas restantes bibliotecas e arquivos nacionais.

Ao longo dos próximos pontos apresentam-se com maior detalhe, para o território português, aquelas que são consideradas as fontes privilegiadas de informação genealógica. Não é de todo um trabalho de levantamento exaustivo, mas antes uma breve passagem pelos principais tipos de fontes e pelas condicionantes que estiveram na sua origem e preservação.

Quanto à sua localização, estas fontes distribuem-se por arquivos públicos, organizados sob a tutela da DGARQ ou de universidades, e arquivos privados, como os arquivos das igrejas paroquiais ou das misericórdias, por vezes instalados "em sítios nem sempre apropriados para a conservação do seu acervo" (Araújo, 2007).

#### 3.1. Registos paroquiais

Os registos paroquiais conhecem-se desde as primeiras décadas do século XVI e foram tornados obrigatórios, para a generalidade dos países católicos, a partir de 1563, por uma norma da 24<sup>a</sup> sessão do Concílio de Trento. Desta forma, embora de origem um pouco mais recuada, os registos de casamento e baptismo assumem carácter obrigatório apenas em 1563, obrigação que se estende aos registos de óbito a partir de 1614. São indiscutivelmente uma das fontes mais abundantes em informação genealógica, para Portugal e para países que foram abrangidos e cumpriram as disposições do Concílio de Trento.

---

<sup>29</sup> <http://antt.dgarq.gov.pt/pesquisar-na-torre-do-tombo/genealogia-ou-historia-local/genealogia/>

<sup>30</sup> <http://www.biblioteca-genealogica-lisboa.org/>

O Registo Civil, que substituirá os registos paroquiais, foi criado pelo Decreto de 16 de Maio de 1832, numa primeira tentativa de secularizar este tipo de registos. No entanto, só em 1911, com a instituição do Registo Civil obrigatório, os registos paroquiais perdem a sua força legal. Os arquivos anteriores a 1911 são, assim, transferidos para as Conservatórias do Registo Civil.

Actualmente estes registos encontram-se depositados, na sua maioria, nos 16 arquivos distritais sob a dependência da Direcção-Geral de Arquivos (DGARQ) e nos arquivos sob a dependência da Universidade de Coimbra (Arquivo da Universidade de Coimbra) e da Universidade do Minho (Arquivo Distrital de Braga).

O grau de preservação dos arquivos varia de distrito para distrito, existindo casos onde persistiram até actualidade mais de 90% dos livros, enquanto noutros casos esse número fica abaixo dos 50% (Daupias D'Alcochete, 1984).

<b>Arquivo</b>	<b>Registos preservados</b>
Viana, Braga, Porto e Coimbra	> 90%
Aveiro, Guarda, Castelo Branco, Lisboa, Setúbal, Évora, Beja e Faro	70-90%
Santarém	> 60%
Leiria, Portalegre, Vila Real e Viseu	< 50%

*Tabela 3. Percentagem de registos preservados por distrito (Daupias D'Alcochete, 1984).*

Há ainda a registar vários esforços de digitalização e disponibilização pública destes registos paroquiais, levados a cabo por vários arquivos como o Arquivo Distrital do Porto<sup>31</sup> ou o Arquivo da Universidade de Coimbra<sup>32</sup>. Neste processo tem sido utilizada uma plataforma conhecida por DIGITARQ<sup>33</sup>, que foi desenvolvida pelo Arquivo Distrital do Porto em colaboração com a Direcção-Geral de Arquivos e a Universidade do Minho.

O Arquivo Regional da Madeira disponibiliza uma base de dados que permite a pesquisa e consulta online de registos de baptismo e casamento, num volume que ultrapassa os 200 mil registos<sup>34</sup>.

### **3.2. Registos notariais**

Os registos notariais englobam um conjunto de documentação mais diversificada, como escrituras, doações, testamentos, entre outros. Este tipo de documentação é resultado do trabalho de tabeliães, em Portugal substituídos pelos notários no século XX. A origem do tabelionado remonta à Baixa Idade Média - encontram-se os primeiros registos da sua existência ainda no século XII -, criados por “determinação régia, impondo-se, gradualmente, a nível local no governo de D. Afonso III” (Coelho, 1997), no século XIII.

Os tabeliães e escrivães produzirão ao longo dos últimos séculos da Idade Média, e sobretudo ao longo da Época Moderna, importantes volumes de documentação. Um

<sup>31</sup> <http://pesquisa.adporto.pt/cravfrontoffice/>

<sup>32</sup> <http://www.uc.pt/auc/projectos/>

<sup>33</sup> <http://digitarq.pt/>

<sup>34</sup> [http://www.arquivo-madeira.org/item1.php?lang=0&id\\_channel=19&id\\_page=242](http://www.arquivo-madeira.org/item1.php?lang=0&id_channel=19&id_page=242)

tabelião de Lamego, do século XV, produziu quase 300 actos ao longo de um único ano, com uma média de 23,7 documentos por mês (Coelho, 1997).

Existem fundos de registos notariais nos arquivos distritais e no Arquivo Nacional da Torre do Tombo.

### 3.3. Nobiliários

Para o período medieval conhecem-se obras - ou fragmentos de obras - que são de especial interesse para a genealogia, como os Livro Velho de Linhagens ou o Livro de Linhagens ou Livro Terceiro, organizado pelo conde D. Pedro. Estas obras descrevem a genealogia das principais famílias nobres portuguesas na Idade Média. Alexandre Herculano disse, a respeito do Livro de Linhagens do conde D. Pedro, tratar-se, “provavelmente do melhor testemunho da literatura genealogica peninsular (...) Constitui a única fonte que descreve vários factos da história portuguesa dos séculos XIII e XIV do ponto de vista da nobreza” (Herculano *apud* Pizarro, 1982).

Na época Moderna um dos exemplos mais conhecidos é a obra de Felgueiras Gaio, o Nobiliário de famílias de Portugal, obra publicada em 17 volumes a partir dos originais manuscritos disponíveis na Misericórdia de Barcelos, já em meados do século XX. Existem outros exemplos, igualmente notáveis, como a História genealógica da casa real portuguesa de António Caetano de Sousa, recentemente reeditada, a Pedatura Lusitana: nobiliário de famílias de Portugal de Cristóvão Alão de Morais ou o Nobiliário de famílias ilustres de Portugal de Rodrigo da Cunha.

### 3.4. Outras fontes

As fontes históricas com especial interesse para os estudos de genealogia não se confinam aos registos paroquiais, notariais e nobiliários. Na obra *Fontes da genealogia em Portugal*, da autoria de Augusto Ferreira do Amaral, identifica-se um conjunto diversificado de fontes que inclui Sumários matrimoniais, Habilitações de genere, Processos judiciais, Justificações de nobreza, Registo de Vínculos, Livros de chancelaria régia, Livros do registo geral de mercês, entre outros.

Os Róis de Confessados e de Crismas são outra importante fonte de informação genealógica, também com origem nos registos produzidos nas paróquias. Contrariamente dos registos paroquiais, depositados nos arquivos públicos, muitos destes documentos permanecem sob controlo da Igreja Católica.

Há ainda registos de passaporte de que o Arquivo Regional da Madeira tem online dados do período entre 1872 e 1915, num total de 35686 registos.

Além das fontes históricas, há ainda a ter conta edições especializadas na genealogia, como monografias familiares, a reedição de obras clássicas, a publicação de fontes ou mesmo obras de referência como o “Dicionário de Bibliografia Para Genealogistas” em 2 volumes, da autoria de Maria João Vieira Craigie (Craigie).

#### 4. Os modelos de negócio

Numa perspectiva de evolução do MHK em direcção a uma maior acessibilidade e facilidade de acesso considerou-se pertinente abordar, mesmo de uma forma sucinta, os modelos de negócio adoptados por plataformas comerciais com afinidades temáticas.

Assim, importa fazer uma análise dos modelos de negócio adoptados por entidades ligadas à produção de conteúdos e disponibilização de serviços *online*, seja numa perspectiva generalista, ou mais focada na genealogia. Um esforço que implica a observação de um período de tempo relativamente curto, mas marcado por profundas transformações tecnológicas. As primeiras incursões comerciais no campo da genealogia, numa abordagem baseada na utilização de novas tecnologias, ocorrem ainda antes a popularização da Internet e sem a revolução da World Wide Web.

Num primeiro momento os serviços e produtos explorados comercialmente são essencialmente programas para computador, dedicados à construção de árvores genealógicas e à organização da investigações pessoais, e também os primeiros conteúdos em formato digital. Nesta perspectiva, é interessante referir, a título de exemplo, a empresa norte-americana Infobases, co-fundada por Paul Allen em 1990 (Allen), que foi criada com o objectivo de distribuir publicações em formato digital - em disquetes e posteriormente em CD-ROM's. Ao longo dos primeiros anos de existência distribuiu essencialmente publicações de carácter religioso e documentos históricos, acabando por especializar-se em informação genealógica. A Infobases compra em 1997 a totalidade da editora Ancestry, aquisição que está na origem da Ancestry, Inc., empresa responsável pela criação de uma das maiores bases de dados genealógicas na Internet. De acordo com Paul Allen, a empresa conseguiu alcançar um fluxo de tesouraria positivo ainda em 1998 (Allen). A mesma empresa entra na era das dot-com enquanto Myfamily.com, que é também o nome do serviço homónimo lançado em Dezembro de 1998 e que alcança o milhão de utilizadores registados nos primeiros 140 dias (Allen). Com o romper da bolha das dot-com, a empresa atravessa um período de reestruturação e contenção de despesas, conseguindo alcançar de novo resultados positivos em 2002. Esta empresa está na origem de um dos maiores grupos dedicados à exploração comercial de serviços e de informação genealógica, a The Generations Network (TGN).

Definitivamente, será o aparecimento da Word Wide Web e a vulgarização do acesso à Internet que irão marcar o crescimento das empresas dedicadas à genealogia e conduzirá aos principais modelos de negócio adoptados na actualidade. A Web, entendida inicialmente como veículo de partilha de informação, não tarda a ser explorada como plataforma comercial. Os modelos de negócio sofreram, inevitavelmente, a transição de um mercado tradicional, baseado nos meios de distribuição convencionais, para novas formas e modelos de comércio, assentes essencialmente no comércio electrónico.

Na realidade actual, uma parte significativa das entidades dedicadas à exploração comercial da genealogia na Internet enfrenta desafios idênticos a outros serviços e plataforma Web e, conseqüentemente, adoptam modelos semelhantes. É possível encontrar, por um lado, o modelo mais tradicional com assinaturas mensais, que é explorado isoladamente ou de forma complementar por uma parte significativa das entidades, mas também outros modelos, baseados em publicidade, em mecanismos de afiliação e mesmo tipos completamente novos como o *freemium* - resultado da conjugação as palavras inglesas *free* e *premium*. Cada um destes modelos será analisado com maior detalhe ao longo das próximas

páginas, tendo sempre em perspectiva a forma como têm sido aplicados nos serviços genealógicos online.

### A. Assinaturas

O modelo baseado em assinaturas, mensais ou com outra periodicidade, é frequente dentro do conjunto de empresas que exploram a genealogia online enquanto negócio. É especialmente comum em entidades que disponibilizam acesso a informação genealógica sob a forma de bases de dados.

A The Generations Network (TGN), nome actual da Myfamily.com e cuja história está intimamente ligada à Infobases e à Ancestry, Inc. anteriormente referidas, utiliza o modelo de assinatura desde o primeiro momento com o lançamento do portal Ancestry.com, em 1996. Presentemente é possível aceder à informação mediante uma assinatura. São disponibilizados dois grandes serviços, um de abrangência mundial e outro que permite apenas acesso a informação dos E.U.A. O custo varia consoante o pagamento seja feito anualmente, trimestralmente ou de forma mensal.

	U.S. DELUXE (valor mensal)	WORLD DELUXE (valor mensal)
Anual	\$ 12.95	\$ 24.95
Trimestral	\$ 16.95	\$ 27.95
Mensal	\$ 19.95	\$ 29.95

Tabela 4. Assinaturas disponibilizadas pelo portal Ancestry.

Outras entidades aplicam este modelo, mesmo que de forma não exclusiva, como a FindMyPast, a WorldVitalRecords, a Geni Inc. ou a MyHeritage Ltd, entre outras.

A WorldVitalRecords segue um modelo de acesso semelhante ao do Ancestry, com dois tipos de acesso (US Collection - 39,95 dólares por ano - e World Collection - 119,95 dólares). A World Vital Records Inc. é um empreendimento da responsabilidade de Paul Allen, o fundador da Infobases e ligado ao portal Ancestry.

O FindMyPast, portal da empresa escocesa Brighsolid, permite igualmente um acesso mediante assinatura, embora este não seja o único modelo de acesso. O sistema de assinatura coexiste com um sistema de créditos adquiridos em pagamentos pontuais, conhecidos por *Pay As You Go*. As assinaturas convencionais podem ser efectuadas por períodos de 6 meses (54,95 libras) ou um ano completo (89,95 libras). A Brighsolid é também um exemplo interessante pelas parcerias que tem estabelecido com os Arquivos Nacionais Britânicos para a digitalização e gestão de importantes fontes de informação genealógica, como os registos paroquiais e os censos.

Este modelo baseado em assinaturas é comumente acompanhado por períodos experimentais (*trial* em inglês), permitindo ao utilizador testar o serviço e decidir se pretende prosseguir com a assinatura.

### B. Freemium

O modelo conhecido por *Freemium* resulta da coexistência de um serviço apresentado como isento de custos, para o utilizador, e um serviço mais avançado, disponível mediante assinatura. O principal pressuposto deste modelo é a co-existência de dois grupos de utilizadores: uma base alargada de utilizadores que usufrui do serviço mais básico e um grupo mais restrito que está disposto a pagar por um serviço mais completo e cujas receitas

permitem suportar os custos globais da plataforma. É um modelo abordado por muitas empresas que possuem serviços integrados na denominada Web 2.0.

Na área da genealogia são várias as empresas que procuram rentabilizar os seus projectos com recurso a este modelo. A The Generations Network utiliza-o no seu website que possui uma vertente mais Web Social, o MyFamily.com. Este disponibiliza um serviço básico, sem quaisquer custos para o utilizador, e um serviço denominado Essentials que é comercializado por 29,95 dólares anuais. É anunciado ainda o lançamento de um serviço Premium, cujos detalhes não estão presentemente disponíveis. Neste caso, o serviço Essentials diferencia-se do serviço básico gratuito pela ausência de publicidade, pela possibilidade de criar temas próprios e pela atribuição de um endereço personalizado.

Um dos mais recentes actores no mundo da genealogia online, a Geni, Inc., empresa responsável pelo portal homónimo, lançou em Janeiro de 2007 um serviço completamente gratuito, em busca de uma massa de utilizadores significativa que pudesse ser rentabilizada posteriormente. É uma estratégia comum a muitas das empresas e serviços online que surgiram nos últimos anos, onde a preocupação em angariar utilizadores se sobrepõe à rentabilidade. A introdução de um modelo *premium* ocorreu num momento posterior, com o lançamento do Geni Pro. A empresa permite pagamentos por períodos de um ano (\$59.40), dois anos (\$95), cinco anos (\$148), ou mesmo uma assinatura vitalícia por 299 dólares. O custo mensal do serviço baixa de forma progressiva consoante o período de assinatura escolhido.

	Custo Mensal	Custo Anual
1 ano	\$4.95	\$59.40
2 anos	\$3.96	(-20%) \$95
5 anos	\$2.48	(-50%) \$148
Vitalício	-	\$299

Tabela 5. Custos de assinatura do portal Geni.com.

O serviço Geni Pro distingue-se do pacote básico gratuito pela presença de funcionalidade adicionais ("Hot Matches", "Collaborators", "Tree Statistics"), pela ausência de publicidade e pela prestação de apoio prioritário aos utilizadores.

### C. Publicidade

A publicidade na Web tem sido explorada praticamente desde os seus primeiros tempos, inicialmente sob a forma de *webbanners*, o tipo de publicidade que será a forma dominante até ao final da década de 1990, e que está também na origem de um conjunto de empresas no período das dot-com.

O modelo actualmente entendido como de maior sucesso foi desenvolvido pela Google e lançado em 2000, sob a forma de pequenos anúncios textuais cujos links permitem acesso a outros *sites* com produtos e serviços. Os anunciantes pagavam de acordo com conjuntos de mil impressões, método conhecido por CPM (*Cost-per-thousand impressions*). Aquando do seu lançamento o sucesso do serviço foi limitado, e a DoubleClick, entretanto adquirida pela própria Google Inc., continuou líder em publicidade online até ao *crash* das dot-com.

As transformações entretanto introduzidas permitiram, finalmente, obter o sucesso que caracteriza actualmente este modelo. O modelo *pay-per-click* foi introduzido pela Google em 2002 e seguia um modelo já existente baseado em cliques e em lanços, mas

aperfeiçoando-o com a introdução de uma característica inovadora conhecida por relevância - esta abordagem inovadora permitia que lanços mais baixos, mas com mais cliques gerassem maiores receitas (Karp, 2008).

Mais recentemente, com a explosão de sites conhecidos por Web 2.0, ou a denominada Web Social, surgiu a necessidade de rentabilizar grandes massas de utilizadores que se tornaram habituais neste tipo de serviços. Sem modelos de negócio claros a possibilidade mais imediata foi a publicidade. A exploração de publicidade nas redes sociais, e na Web social em geral, continua a constituir um importante desafio para as empresas que exploram estes portais - facebook, flickr, delicious – e cujos resultados não é possível ainda avaliar convenientemente.

No caso particular dos serviços *online* dedicados à genealogia, são também comuns os exemplos de modelos baseados em publicidade, frequentemente associados a assinaturas que permitem eliminar os anúncios. Embora em alguns casos, como no Findmypast, a publicidade presente no website seja, no momento em que se escreve este texto, essencialmente para promoção interna, existem outros serviços onde é possível encontrar o modelo publicitário aplicado. O Ancestry possui uma secção dedicada à componente publicitária e a eventuais anunciantes que poderão publicitar no seu portal<sup>35</sup>. Ainda dentro do grupo TGN, a publicidade é particularmente visível no serviço Myfamily.com.

O portal Os Meus Parentes utiliza publicidade no seu modelo de acesso gratuito, com recurso aos serviços da Google AdWords.

#### **D. A comunidade e doações**

A Web 2.0 permitiu, definitivamente, que os utilizadores assumissem um papel central na partilha e mesmo produção de conteúdos. Os utilizadores partilham opiniões, fotografias, vídeos, na generalidade dos casos sem qualquer expectativa relativamente a qualquer remuneração.

A utilização da comunidade como uma mais valia que permite gerar receitas, apesar da sua popularização nos últimos anos, não é um fenómeno inteiramente novo. Antes da denominada Web 2.0 já existiam exemplos de empresas bem sucedidas que dependiam do esforço comunitário. O mundo do desenvolvimento de software *open source* é ilustrativo deste fenómeno. Empresas como a Red Hat desenvolveram produtos e modelos de negócio em torno de uma tecnologia desenvolvida em comunidade e em código aberto - o Linux *kernel* e todas as bibliotecas GNU que compõem o sistema.

As novas empresas da Web 2.0 trouxeram a necessidade de obter proveitos a partir das suas grandes massas de utilizadores. Redes sociais como o Facebook, o Flickr ou o Twitter possuem milhões, mesmo centenas de milhões de utilizadores, que utilizam os seus serviços de forma gratuita. Um dos principais desafios para os novos projectos tem sido conseguir monetizar esta imensidão de utilizadores.

É frequente a utilização do esforço comunitário. Muitas entidades esperam alcançar uma massa crítica suficientemente alargada que permita posteriormente explorar o serviço comercialmente através de serviços *premium*, da introdução de publicidade ou mesmo a venda do serviço a terceiros que encontrem um modelo de negócio satisfatório.

Estes tipo de serviços estão extremamente dependentes dos utilizadores que contribuem com a introdução de informação. No caso dos serviços genealógicos existem o geni.com e

---

<sup>35</sup> <http://www.ancestry.com/advertising/default.aspx>

osmeusparentes.com.pt que foram lançados como serviço inteiramente gratuitos onde os utilizadores podem criar as suas árvores genealógicas.

No entanto, mesmo empresas que disponibilizam informação de âmbito próprio, como o portal Ancestry, enveredam por estratégias que permitam envolver os utilizadores num espírito mais comunitário. O Ancestry possui um serviço denominado World Archives Project, que permite aos utilizadores consultar arquivos digitalizados e transcrever para um formulário a informação principal, como nomes, datas, locais, transformando-os em informação pesquisável. O resultado, sob a forma de índice, poderá ser pesquisado livremente por toda a comunidade, mas o acesso aos originais só é possível mediante assinatura. Como contrapartida os membros activos podem beneficiar de um desconto de 10% numa assinatura anual “U.S. Deluxe” ou 15% numa assinatura “World Deluxe”. A empresa considera ainda a possibilidade de dar acesso às imagens originais aos contribuintes activos como alternativa a este modelo actual.

As doações podem ser consideradas uma forma de dependência de uma comunidade. Este modelo baseia-se no princípio de que os utilizadores que beneficiem e fiquem satisfeitos com determinado produto ou serviço irão retribuir de alguma forma. Não é um modelo particularmente popular em sites dedicados à genealogia. Na Web em geral, a Wikipedia é possivelmente o caso mais paradigmático de dependência da comunidade e de doações: os custos da plataforma têm sido suportados por doações recolhidas em algumas campanhas anuais.

### **E. Merchandising**

É, sobretudo, um modelo de negócio complementar e no universo da genealogia nunca é utilizado como modelo principal de rentabilização dos projectos. Há, porém, alguns exemplos de projectos dedicados à genealogia que utilizam a comercialização de *merchandising* - presentes e lembranças - como mecanismo de obter receitas suplementares.

Neste campo merece especial referência o portal francês genealogie.com, parte integrante do grupo Notrefamille.com SA criado em 1994 e que é detentor de um conjunto de serviços online que ultrapassam a genealogia. Através de um conjunto de domínios distintos, mas perfeitamente interligados, o grupo explora a disponibilização de serviços em torno da família, onde se inclui a genealogia, mas também, beleza, culinária, a comercialização de presentes, entre outros. Na verdade, são exploradas diferentes vertentes de negócio, que incluem a publicidade, a assinatura e a comercialização de produtos associados às diferentes temáticas. No caso particular da genealogia são comercializados produtos como vídeos com sínteses dos principais acontecimentos de um determinado ano de nascimento, árvores familiares impressas, livros dedicados à genealogia, postais com reproduções de fotografias antigas e mesmo colectâneas com música lançada no ano de nascimento da pessoa.

O portal Ancestry possui também uma loja online onde comercializa alguns produtos complementares, como livros, *software*, álbuns de fotografias, postais e mapas antigos, entre outros.

### **F. Afiliação**

O modelo de afiliação implica uma parceria entre um portal principal e um segundo *website*. O segundo envia utilizadores para o primeiro que, no caso de realização de alguma

mais valia, partilha uma parte do rendimento conseguido. É essencialmente uma forma de fazer crescer a base de utilizadores e, conseqüentemente, também as receitas, não obstante a partilha imediata com o afiliado de uma parte dos proveitos.

A afiliação é especialmente popular em determinados sectores da Web, como o comércio de artigos (Amazon) ou na prestação de serviços de alojamento Web. Na vertente genealógica existem também alguns exemplos de portais que utilizam este modelo para atrair novos utilizadores.

A TGN através do Ancestry e a FindMyPast exploram este modelo e permitem a criação de contas de afiliados. Os afiliados recebem um *banner* Web que deverão colocar nos respectivos *websites*, servindo assim de pólos de atracção para os portais genealógicos. No caso particular do Ancestry, os membros afiliados recebem até 60 % de comissão por cada utilizador que enviem e que faça uma assinatura.

O Notrefamille.com também oferece um programa de afiliação dedicado à componente genealógica, retribuindo 10% de cada venda.

### **G. Códigos promocionais e cupões de desconto (Promotional Codes and Vouchers)**

Mais do que um modelo de negócio só por si, os cupões de desconto e códigos promocionais representam uma estratégia de promoção utilizada para atrair novos utilizadores e assim gerar receitas pela via da publicidade ou de assinaturas. Podem assumir essencialmente estas duas formas: como códigos promocionais que oferecem um determinado desconto na aquisição de serviços ou como *vouchers* que podem ser adquiridos ou oferecidos por outras entidades.

O portal Ancestry, em algumas das suas versões internacionais, permite a aquisição de *vouchers* a partir de terceiros como alternativa de aquisição dos seus serviços, recomendando-os a quem não possui cartão de crédito ou como presente para oferecer a alguém.

O britânico FindMyPast.com segue uma estratégia idêntica, permitindo a aquisição de *vouchers* que podem ser oferecidos a terceiros. Para usufruir de um *voucher* é necessário o registo de utilizador no portal. Os *vouchers* podem ser adquiridos em diversos locais, como bibliotecas e arquivos e mesmo através dos The National Archives.

A oferta deste tipo de promoções são frequentes noutros serviços *online*.

Pretendeu-se neste capítulo realizar um levantamento dos principais tipos de modelos de negócio utilizados na rentabilização comercial de serviços *online*. Estes modelos são particularmente interessantes para uma perspectiva de constituição de um serviço comercial que tenha por base o Time Link.

## 5. Actores e casos de uso (requisitos)

Como anteriormente referido, o MHK surge como uma aplicação de uso restrito e confinado a um grupo relativamente fechado de investigadores que o utilizaram no âmbito de actividades de investigação científica na Faculdade de Letras da Universidade de Coimbra. A sua evolução, de acordo com o apresentado no segundo capítulo, foi condicionada pelas necessidades deste grupo de utilizadores, sendo a preocupação fundamental, aquando da sua concepção, a satisfação de objectivos relacionados com a investigação e com a reconstituição de comunidades históricas.

O objectivo de levar a aplicação a um outro nível, libertando-a do mundo restrito da investigação, obriga a um conjunto de alterações significativas, através de um processo de adaptação e adequação a novos públicos, tendo por base os actuais trâmites de funcionalidade, usabilidade e acessibilidade.

Ao longo das próximas páginas definem-se os principais requisitos para esta nova etapa e identificam-se os principais utilizadores. Utiliza-se, para este fim, e como alternativa à apresentação tradicional dos requisitos, uma metodologia conhecida por casos de uso (*use cases*, em inglês).

"A use case is a mechanism for recording the behavior of some system. It consists of a set of stakeholders with goals and shows the scenarios that may play out in completing those goals, the Main Success Scenario and Extensions--alternate success scenarios and error cases. It provides a well defined method for producing requirements that fulfills several important criteria when properly written" (Battista, 2009).

### 5.1. Actores e objectivos (actors and goals)

Quais os actores e quais os seus objectivos, são duas premissas determinantes na descrição de um caso de uso. Os actores e os seus objectivos irão determinar o funcionamento da aplicação e as formas de interacção entre o sistema e o utilizador. Sistema e utilizador são dois actores de um caso de uso, tendo presente que um actor não tem de corresponder forçosamente a uma pessoa, pois pode assumir a forma de um outro sistema que interage com o primeiro.

No caso particular do novo sistema, o Time Link, para além do investigador em História e Micro-História, que são os actores habituais, será possível identificar novos actores, como o investigador em genealogia, o cidadão comum que se interessa pela sua genealogia, o profissional que presta serviços de investigação, o professor e o aluno que utilizam o MHK como ferramenta pedagógica e a comunidade em geral.

A pessoa que está a reconstruir a sua genealogia é caso mais imediato e é representado pelo cidadão comum que, por qualquer razão, desperta para a temática da genealogia e decide reconstruir a sua história familiar. Interessa-se sobretudo pela pesquisa de nomes, auxiliada por intervalos de tempo, pela pesquisa de locais ou profissões e pelas relações familiares existentes.

O genealogista profissional e o amador são outros actores possíveis. Quando o interessado na reconstituição da sua história familiar constata que, por razões diversas, não tem disponibilidade para realizar a investigação, é frequente recorrer a serviços de investigação profissional.

Há ainda o investigador em genealogia que se distingue do anterior por um interesse intrínseco e que não é movido necessariamente pela prestação de serviços comerciais de investigação.

Temos o investigador em Micro-história, interessado na reconstituição de biografias e na compreensão das comunidades históricas.

Os professores e alunos poderão ser também actores importantes, numa perspectiva de utilização pedagógica, útil na aprendizagem da História, de questões da Demografia, etc.

Finalmente, a comunidade, representada pelas câmaras municipais. As câmaras municipais poderão desempenhar um papel particular. Representam por um lado a comunidade de munícipes, que serão os utilizadores, e são os *stakeholders* que podem financiar a inserção de informação e acompanhar o desenvolvimento.

Todos estes actores desempenham um papel activo enquanto utilizadores do sistema (pesquisando ou inserindo informação), excepção feita aos municípios que poderão desempenhar um papel importante na disponibilização pública de maiores volumes de informação.

## 5.2. Casos de uso (Use cases)

"Use cases are popular largely because they tell coherent stories about how the system will behave in use" (Cockburn, 2001).

Os casos de uso podem assumir diversas formas, com diferentes níveis de formalização. De acordo com Cockburn (2001) é possível identificar vários tipos, desde o forma casual mais descritiva à forma mais completa (*fully dressed form*) e mais formalizada.

Ao longo das próximas linhas apresentam-se casos de uso na forma casual, por entender adaptar-se melhor ao formato narrativo do presente documento. A sua forma descritiva simplifica os casos e torna mais apelativa a leitura.

Estes primeiros casos de uso, de seguida descritos, são relativamente transversais aos actores identificados no ponto anterior, correspondendo de forma geral a um actor ou tipo utilizador genérico. No entanto, o actor e utilizador poderá especializar-se em investigador, genealogista, professor ou cidadão comum, facto que poderá implicar casos de uso particulares à sua actividade. Estes actores especializados poderão ter acesso a funções particulares que não estão acessíveis aos restantes utilizadores.

Segundo John M. Battista (Battista, 2009) o principal objectivo dos casos de uso deverá ser capturar a intenção do actor e não os detalhes da interface de utilizador.

### **Entrar no Portal Web**

É o caso de uso mais básico e a partir do qual são possíveis os seguintes. Pressupondo uma ligação efectiva à Internet e um navegador (*browser*) aberto, o utilizador introduz o endereço do portal Web e acede ao Time Link. O sistema dá as boas-vindas ao utilizador, mostra as últimas interacções dos restantes utilizadores (identificações, comentários, etc.) e as novidades do sistema (actualizações ao sistema como novas funcionalidades, novos dados), e dá a possibilidade do utilizador se identificar através das suas credenciais de acesso. Pode ser comunicado ao utilizador qual a informação a que tem acesso enquanto utilizador anónimo e enquanto utilizador registado, caso exista essa distinção. É ainda dada a possibilidade de criação de uma nova conta de utilizador.

### **Criar uma conta de utilizador**

O utilizador introduz os seus dados pessoais num formulário de registo. O formulário de registo é processado pelo Time Link, que procede a uma verificação prévia da informação e, no caso de existirem inconsistências ou incompatibilidades, remete de novo para o utilizador, solicitando a correcção dos campos em questão. Caso contrário, regista-se o novo utilizador na base de dados e é enviado um e-mail solicitando a confirmação da inscrição. Quando o utilizador clica na ligação constante nessa mensagem de *email*, a conta fica activa e passa a ser possível o acesso enquanto utilizador registado.

### **Consultar área pessoal**

O utilizador registado, através da interacção com a base de dados, vai produzindo e acumulando informação que acaba por ficar dispersa pelas várias ocorrências da base. São biografias, comentários, notas pessoais, etc. A partir da sua área pessoal tem acesso ao resultado da sua interacção com o sistema.

### **Aceder a múltiplas bases de dados**

O utilizador identifica-se com um nome de utilizador e palavra-chave. É recebido por uma página do tipo “*dashboard*” onde pode verificar quais as bases de dados que são acessíveis e a sua última actividade no sistema. Poderá ter conhecimento de outras bases de dados a que poderá aceder, mas que não estão incluídas na sua assinatura, no caso de uma utilização comercial. Em qualquer momento pode regressar à sua página pessoal e seleccionar outra base de dados ou, em alternativa, simplesmente utilizar algum mecanismo de selecção rápida. Caso o utilizador possua acesso a uma única base de dados esta deverá ser automaticamente seleccionada.

### **Pesquisar**

O utilizador poderá efectuar vários tipos de pesquisa: a partir do nome de um antepassado, de um intervalo de tempo, uma residência e profissão, etc. O sistema deverá devolver uma listagem com ocorrências que incluem o nome ou expressão pesquisados, e de acordo com o intervalo de tempo escolhido. No caso de uma pesquisa realizada sobre múltiplas bases de dados, o sistema poderá devolver uma listagem geral das bases de dados que incluem resultados para a pesquisa efectuada, com indicação do número de ocorrências disponíveis em cada uma, e acessíveis através de um *link*.

### **Explorar**

Nos casos em que o utilizador tem apenas uma vaga ideia do que procura, ou simplesmente pretende obter uma ideia do tipo de informação que a base de dados alberga, poderá recorrer a uma vertente exploratória. Explora, assim, a base de dados através de listagens com atributos como nome, residência, profissão, cargo, etc. Para cada um dos atributos obtém listagens que permitem aceder às fichas biográficas.

### **Consultar fichas biográficas e de ocorrência**

A partir das listagens obtidas na exploração ou nas pesquisas da base de dados, o utilizador poderá aceder às fichas biográficas dos indivíduos através das respectivas matrículas de identificação ou através dos nomes. Cada ficha biográfica reúne a informação identificada como pertencendo à mesma pessoa.

### **Comentar**

O utilizador encontrou uma ocorrência na base de dados que já conhecia de outras fontes. Verifica que a informação está incorrecta ou incompleta. O utilizador previamente identificado no sistema deixa um comentário textual com informação adicional sobre a ocorrência ou simplesmente responde a um comentário anteriormente colocado por outro utilizador.

### **Identificar e construir biografias**

A construção de biografias é um aspecto fundamental em bases de dados que lidam com ocorrências isoladas de pessoas. O utilizador previamente identificado no sistema pretende reconstituir biografias de pessoas que está a estudar. No caso de ainda não ter sido feita nenhuma identificação para uma determinada pessoa, as ocorrências permanecem dispersas pela base de dados. A partir de uma listagem de ocorrências identificam-se aquelas que pertencem a um mesmo indivíduo e realiza-se a sua aglutinação numa ficha biográfica. Progressivamente, explorando a base de dados, o utilizador vai encontrando novas ocorrências que vai acrescentando à ficha. Numa base de dados acessível publicamente há um número indeterminado de utilizadores que poderão reunir a sua própria colecção de biografias, que poderão ficar acessíveis a consulta de terceiros.

### **Consultar ficha de família**

A partir da ficha biográfica ou de uma ocorrência isolada o utilizador decide consultar a ficha de família de determinada pessoa. Nessa nova página é-lhe apresentada a ascendência e descendência, bem como a possibilidade de aceder a uma visualização gráfica da informação, sob a forma de árvore genealógica, vista de costado ou *pedigree*.

### **Explorar árvore genealógica**

O utilizador encontra-se na ficha de família e pretende uma visualização gráfica da informação que está acessível nessa ficha. A partir de uma lista de opções escolhe o tipo de visualização que pretende aceder (vista de costado, árvore genealógica ou *pedigree*, por exemplo). Quando entra na visualização, devido à extensão da informação genealógica, poderá não conseguir visualizar toda a informação de uma única vez. Assim, será necessário recuar ou avançar no intervalo que está a ser exibido.

### **Exportar e importar GEDCOM**

O utilizador encontra-se na ficha de família e pretende exportar a informação que está a visualizar. Poderá descarregar um ficheiro do tipo gedcom que guarda no seu computador e pode importar noutras aplicações de tratamento de informação genealógica.

O caso inverso acontece quando o utilizador tem dados que registou com um programa de tratamento de informação genealógica e pretende transferir essa informação para o novo sistema.

Estes casos de uso serão fundamentais na prossecução dos trabalhos descritos no capítulo seguinte onde se aborda finalmente o novo sistema, com especial atenção à Interface Gráfica de Utilizador.

## 6. Time Link: o desenho de um novo sistema

Identificados os requisitos passemos à especificação do novo sistema, o Time Link. Ao longo do próximos pontos serão analisadas questões de tecnologia e numa segunda parte especificados cada um dos módulos que compõem o Time Link. Cada módulo representa uma secção de implementação relativamente autónoma, que engloba em si um conjunto de funcionalidades significativas.

### 6.1. A plataforma

O Time Link é na verdade uma aplicação Web que requer a instalação de um conjunto de software adicional, nomeadamente um servidor Web (no caso o Apache Tomcat desenvolvido pela *The Apache Software Foundation*) e um servidor de bases de dados (MySQL da *Sun Microsystems* - adquirida pela *Oracle* em 2010). O Time Link e os elementos de suporte são desenvolvidos em tecnologia de código aberto<sup>36</sup>: O Apache Tomcat é lançado sob uma licença Apache License version 2 e o MySQL usa uma licença GPL. O Time Link é desenvolvido sob a New BSD License.

O Apache Tomcat e o MySQL são, no conjunto das aplicações não comerciais, dois dos produtos mais utilizados na criação de aplicações Web complexas e na gestão de bases de dados. Ao longo das próximas linhas apresentam-se alguns detalhes de cada uma das tecnologias por entender-se útil à visão geral da plataforma de suporte ao Time Link.

#### O Servidor de Base de Dados (MySQL)

A base de dados é, possivelmente, o principal componente do Time Link, ainda que tal facto nem sempre seja perceptível para o utilizador final. A estrutura de dados apresentada no segundo capítulo é concretizada com recurso ao servidor de bases de dados conhecido por MySQL.

O MySQL foi desenvolvido por Michael Widenius (também conhecido por Monty) na empresa sueca TcX DataKonsult AB a partir de componentes ISAM (Indexed Sequential Access Method)<sup>37</sup>, que haviam sido produzidos no final da década de 70, e inspirando-se e assegurando a compatibilidade com outro servidor de bases de dados já existente, conhecido por miniSQL ou mSQL<sup>38</sup>. A primeira versão pública do servidor (MySQL 3.11.1) foi lançada em Outubro de 1996, primeiro para sistemas Solaris e um mês depois para sistemas Linux, acompanhada pela disponibilização pública do código-fonte.

O servidor MySQL ganhará popularidade com o crescimento da Web, surgindo frequentemente associado à linguagem PHP na produção de *websites* dinâmicos. A companhia MySQL AB é adquirida pela Sun Microsystems, em 2008, e esta pela Oracle em 2010, a actual detentora da tecnologia.

---

<sup>36</sup> <http://www.gnu.org/philosophy/free-sw.html>

<sup>37</sup> ISAM (Indexed Sequential Access Method), originally developed by IBM for mainframe computers, is a method for storing large amounts of data for fast retrieval. ISAM forms the basic data store for almost all databases. <http://apachepersonal.miun.se/~vivasp/IRIS31/IRIS31-060.pdf>

<sup>38</sup> From Visions to Reality - an interview with David Axmark, Co-Founder of MySQL AB  
<http://dev.mysql.com/tech-resources/interviews/david-axmark.html>

O MySQL é um sistema de gestão de base de dados relacional que utiliza a linguagem SQL (Structured Query Language)<sup>39</sup> como interface de consulta.

Para o utilizador final avançado, a linguagem SQL é particularmente importante no momento de consulta e manipulação da informação, especialmente em operações mais complexas de consulta que implicam, em muitos casos, o recurso a expressões SQL.

### **O Servidor Web (*Tomcat servlet container*)**

O Time Link é composto ainda, para além da base de dados, por uma aplicação Web que corre sobre o servidor Apache Tomcat, um servidor de aplicações Web. O desenvolvimento da aplicação é realizado em linguagem Java, que mais que uma simples linguagem de programação, é entendida como uma plataforma composta por dois grandes componentes: a API Java, que fornece as classes de programação e a máquina virtual Java (JVM) que fica encarregue de executar o código (Greanier, 2004).

De um ponto de vista formal, Java é uma linguagem de alto nível, com uma sintaxe derivada de outra linguagem popular, o C++. Segue o paradigma de programação orientado a objectos, tendo como principais características ser interpretada<sup>40</sup>, independente da arquitectura e *multithreaded*<sup>41</sup>. Actualmente é utilizada num conjunto amplo de aplicações, desde a Web, a programas de computador Desktop, aplicações para telemóveis, entre outros.

A sua origem remonta ao início da década de 90, então ainda sob a designação Oak, e foi desenvolvida no âmbito de um projecto secreto da Sun Microsystems denominado "Green Project" (Projecto Verde). Como era objectivo do projecto a produção de uma linguagem para uma nova geração de equipamentos electrónicos inteligentes, procurou-se torná-la independente da plataforma (Greanier, 2004), facto que se revelará determinante para a sua posterior adopção generalizada.

O Projecto Verde não alcançou grande receptividade junto da indústria e apenas com a explosão da World Wide Web a linguagem encontra aquele que será o seu ecossistema de eleição. A linguagem Oak é renomeada para Java, devido à existência de uma outra linguagem com o mesmo nome e, em Maio de 1995, é apresentada oficialmente pela Sun (Greanier, 2004).

O lançamento da Java Servlet API<sup>42</sup> para desenvolvimento de aplicações em Java que correm num servidor acessível na Internet, em 1997, veio revolucionar o desenvolvimento para a Web (Haines, 2002), e contribuir para a afirmação da linguagem e da plataforma.

O **Apache Tomcat**<sup>43</sup>, embora seja também um servidor de páginas Web estáticas, destaca-se pelo suporte às tecnologias Java Servlet e JavaServer Pages, tecnologias utilizadas na criação de websites dinâmicos e que são a base da interface gráfica do Time Link.

---

<sup>39</sup> Linguagem de consulta estruturada desenvolvida nos anos 70 e utilizada em consultas em bases de dados relacionais.

<sup>40</sup> O código não é executado directamente pelo processador, mas antes por um programa intermédio, o interpretador. Interessante notar que há linguagens interpretadas que podem ser compiladas.

<sup>41</sup> Múltiplas linhas de execução que permitem a execução de várias tarefas em simultâneo.

<sup>42</sup> Classe Java construída de acordo com a Java Servlet API (Application Programming Interface) e que gera respostas a pedidos HTTP.

<sup>43</sup> O Tomcat começou como uma nova versão do container que estava integrado no Java Servlet Development Kit (JSDK), resultado do trabalho de James Duncan Davidson, empregado da Sun. Com o objectivo de

O Time Link é constituído por classes do tipo Servlet que respondem a pedidos HTTP. Estas classes são responsáveis pela ponte entre o modelo de dados e a visualização da informação. A formatação da informação utiliza a linguagem de templates Velocity<sup>44</sup>, templates apoiados por CSS e Javascript.

Ao longo deste ponto apresentaram-se os principais componentes do Time Link: a base de dados e a interface gráfica de utilizador. No ponto seguinte deste capítulo serão apresentados esclarecimentos adicionais sobre a interface do Time Link.

## 6.2. A interface de utilizador

A interface de utilizador, ou Interface Gráfica de Utilizador/GUI (Graphic User Interface), é, para o utilizador final, a única face visível de todo o sistema. É através dessa interface que acede à informação e também através dela constrói uma imagem, positiva ou negativa, de todo o sistema.

No desenvolvimento do Time Link adoptou-se um padrão de implementação conhecido por MVC (Model-View-Controller), cujo principal objectivo é separar a lógica interna do programa da componente de visualização da informação. A formatação da apresentação fica a cargo de templates Velocity (“View”) que lidando directamente com Servlets<sup>45</sup> (“Controller”) obtêm a informação do Modelo de Dados (“Model”). Este modelo tem um impacto importante ao nível da implementação, permitindo produzir código mais limpo e escalável, mas também no processo de desenvolvimento, tornando possível que equipas de programação se possam focar na lógica interna da aplicação, enquanto equipas de *design* se dedicam à concepção da interface.

O acesso à interface gráfica de utilizador do Time Link é efectuado através de um *browser* Web, quer seja uma instalação local num computador pessoal ou uma instalação num servidor acessível na Internet. A apresentação e organização da informação segue as convenções mais comuns na World Wide Web. O *website* é acessível a partir de um URL - <http://localhost:8080/mhk> para as instalações locais e <http://timelink.fl.uc.pt/mhk> para as instalações remotas - e os conteúdos organizam-se em páginas acessíveis através de hiperligações.

Uma parte significativa do trabalho de desenvolvimento do Time Link tem sido realizado com recurso a design centrado no utilizador (*user-centered design*). Neste modelo de desenvolvimento os utilizadores finais desempenham um papel preponderante e acabam por integrar-se, de alguma forma, no processo de desenvolvimento, influenciando decisivamente a direcção que a interface deve tomar. É um modelo que se caracteriza por uma avaliação e melhoria constante do produto sempre tendo por objectivo uma resposta às necessidades do utilizador.

O facto de se tratar, na verdade, de um *website* implica que as questões de usabilidade e acessibilidade são semelhantes às da World Wide Web, sendo necessário ter em conta as directivas existentes para a construção de *websites*, tanto ao nível da usabilidade como da acessibilidade. São, de resto, questões transversais a todos os módulos descritos neste

---

disseminar a tecnologia JSP a Sun abre o código fonte da aplicação e entrega a sua manutenção à Apache Software Foundation. (Brittain, 2008)

<sup>44</sup> Apache Velocity é uma linguagem para criação de templates e também um motor que permite a execução desses mesmos templates.

capítulo, e que se revelam fundamentais para assegurar a igualdade de acesso à informação, independentemente do browser e sistema operativo ou das suas condições físicas de acesso de cada utilizador.

“In a broad sense, accessibility simply means ensuring that a given page on the Web is able to be accessed. Accessibility is not about disability, rather, it’s about people getting to the shared information that the vision of the Web has made manifest.” (Thatcher, Burks, Heilmann, 2006).

As alterações produzidas no Time Link visaram não só a criação de novas formas de visualização da informação, mas também, possibilitar uma experiência global mais apelativa para o utilizador final. É importante que estas decisões se reflectam também no código HTML produzido que, extirpado de formatação e outros elementos adicionais, permite melhorar a experiência não só em termos de usabilidade mas também na acessibilidade. A adopção de tecnologias Web mais actuais como o CSS (que permite separar as opções de formatação do conteúdo propriamente dito) e Javascript (permite novas formas de interacção mais dinâmicas) foram fundamentais neste campo.

Ainda no campo da acessibilidade<sup>46</sup>, é importante ter presente o trabalho desenvolvido pelo World Wide Web Consortium (W3) que tem desempenhado um papel importante na promoção de orientações abrangentes que se dirigem não só aos criadores de conteúdos, mas que se aplicam também nas ferramentas produção, visualização e reprodução de conteúdos, bem como nas tecnologias assistivas.

### 6.3. Os módulos

Cada módulo representa uma secção relativamente autónoma e engloba um conjunto de funcionalidades importantes para a experiência global do utilizador. Esta divisão tem como objectivo simplificar o desenvolvimento e permitir a sua implementação faseada. Em cada módulo é enunciada uma breve contextualização que pretende transmitir a ideia base subjacente, seguida por uma descrição textual e gráfica da Interface Gráfica de Utilizador (GUI), e eventualmente por alguns detalhes de implementação, sempre que se entenda necessário e pertinente.

#### 6.3.1. Famílias e genealogias

##### Contextualização

A reconstituição de famílias e respectivas árvores genealógicas é uma funcionalidade fundamental, sobretudo se tivermos em consideração que o utilizador mais frequente procura essencialmente informação relativa às suas raízes ancestrais, e tem como grande objectivo a reconstituição da sua genealogia, procurando recuar geração após geração até ao limiar dos dados disponíveis.

É igualmente importante considerar a existência de diferentes públicos, com múltiplas necessidades e ambições e, assim, disponibilizar formas diferenciadas de apresentação da informação. A apresentação gráfica da informação deve adaptar-se a esta condicionante e,

---

<sup>46</sup> <http://www.w3.org/WAI/intro/accessibility.php>

além das árvores genealógicas comuns, deve permitir a existência outras formas de exibição como as árvores de costados<sup>47</sup> ou os diagramas de *pedigree*<sup>48</sup>.

A partir dos casos de uso apresentados no capítulo precedente é possível enunciar os principais objectivos a concretizar neste módulo:

- \* reconstrução de genealogias;
- \* suporte a fichas de família;
- \* representação gráfica da informação (árvores genealógicas tradicionais, árvores de costados ou diagramas de *pedigree*);
- \* importação e exportação da informação em formato GEDCOM<sup>49</sup>.

Ao longo das próximas páginas descreve-se a interface gráfica que será apresentada ao utilizador final e que corresponde às especificações funcionais da aplicação. A sua concretização implica não só essa nova interface gráfica, mas também novas tabelas na base de dados e novos procedimentos que ficarão responsáveis pelo seu preenchimento. Estas especificações técnicas são descritas mais à frente em “Notas de Implementação” e “Estrutura de Dados”.

### **A Interface Gráfica de Utilizador (GUI)**

O acesso à ficha de família é possível através da ficha de pessoa real seleccionando o separador "Família".

#### ***Ficha de família***

A ficha de família permite consultar a ascendência e descendência imediata do indivíduo seleccionado. Acessível a partir do separador família, pode ser consultada para qualquer pessoa real que tenha descendência.

---

<sup>47</sup> Exibem toda a ascendência materna e paterna, recorrendo à numeração de Sosa-Stradonitz..

<sup>48</sup> Mostram a ascendência e descendência de um indivíduo com recurso a símbolos especiais: um quadrado para os homens, um círculo para as mulheres, uma linha para ligar pessoas.

<sup>49</sup> GEDCOM é um formato genérico de exportação de informação genealógica, criado para permitir a troca de dados entre diferentes aplicações.

Time Link

http://www.timelink.pt/show/id/rp-132

OBIDOS Time Link

Search Explore Calendar

Início > Pessoa Real > António Silva

**António Silva**  
b. 14.04.1851

Biografia Identificacoes Pessoas relacionadas Familia

Vista de Costados Árvore Genealógica Exportar GEDCOM

João Silva Bárbara Dias José Santos Josefa de Carva

António Silva b. 14.04.1851 c. 14.08.1876 Maria dos Santos b. 14.04.1852 m. 24.11.1890

Francisco dos Santos Silva b. 12.03.1878

Mariana dos Santos Silva b. 19.10.1879 c. 12.08.1898 Mário Mendes f. João f. Mariana

António Silva b. 14.04.1851 c. 03.09.1899 Anabela Simões b. 14.04.1852

Figura 9. Ficha de família

Neste ecrã exemplificativo podemos ver os dois casamentos de António Silva, a respectiva data de casamento e os filhos de cada casamento (neste caso existiu um segundo casamento por morte da primeira mulher). Podemos ver ainda os pais de António Silva e da sua esposa Maria dos Santos. Relativamente aos filhos do casal podemos ver a data de baptismo e, caso disponíveis, as datas de casamento e óbito, bem como os primeiros nomes de eventuais descendentes. A partir de qualquer nome é possível aceder às respectivas fichas biográficas.

### *Árvore Genealógica*

O acesso à árvore genealógica de António Silva é possível a partir da ficha de família em "Visualizar Árvore Genealógica".

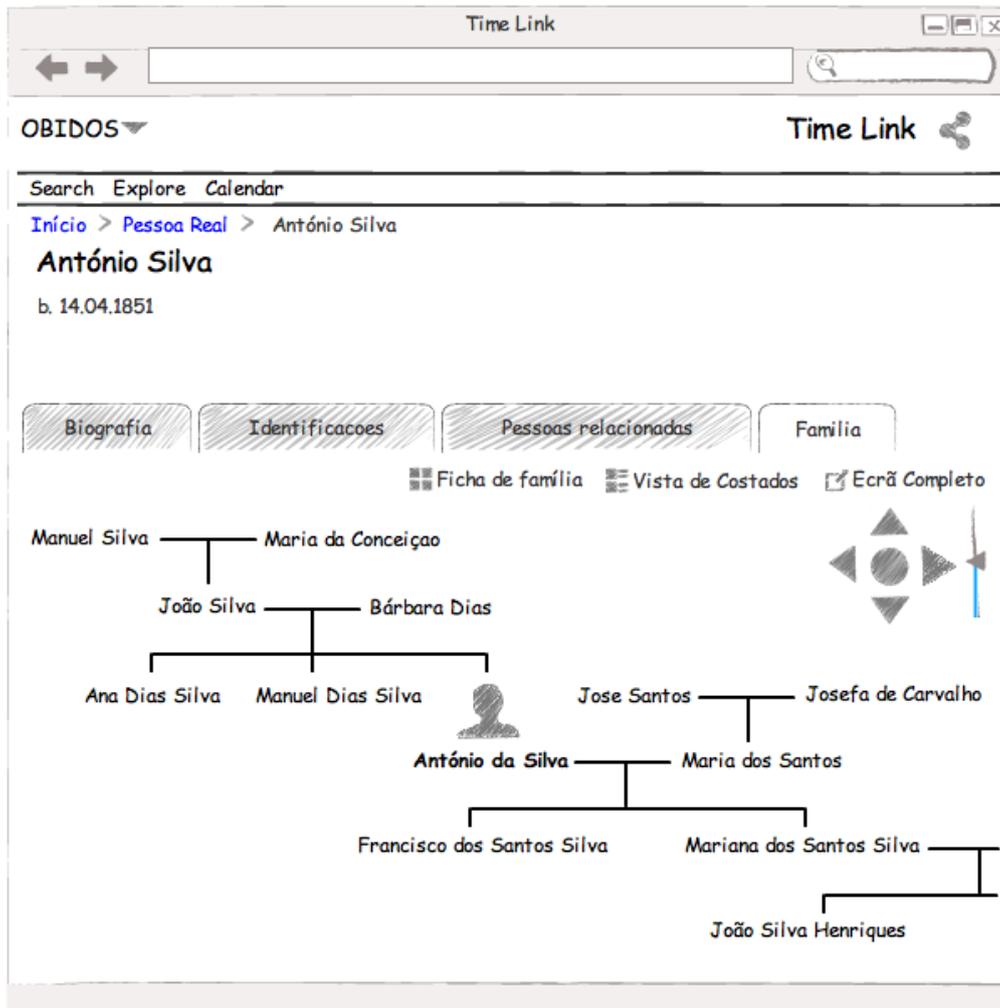


Figura 10. Árvore genealógica

A árvore genealógica toma por foco o indivíduo cuja ficha biográfica estava a ser consultada. Por constrangimentos de espaço no ecrã, esta poderá ser apresentada apenas parcialmente, disponibilizando comandos de para deslocar, aproximar ou afastar.

No exemplo aqui apresentado a árvore surge numa nova janela, cuja dimensão poderá ser ajustada pelo utilizador. Apresenta o indivíduo António da Silva no centro com a sua ascendência, a sua esposa, os seus irmãos, filhos e netos.

### ***Outras formas de visualização gráfica***

Existem outros modelos de visualização gráfica, para além das 'tradicionais' árvores genealógicas, como as vistas de costados e os diagramas de *pedigree*. Cada um dos modelos reúne as preferências de um determinado público. Poderá ser dada a possibilidade de selecção do tipo de visualização gráfica a usar na apresentação da informação. A título de exemplo apresenta-se aqui uma vista de costados a partir do mesmo indivíduo anteriormente referido.

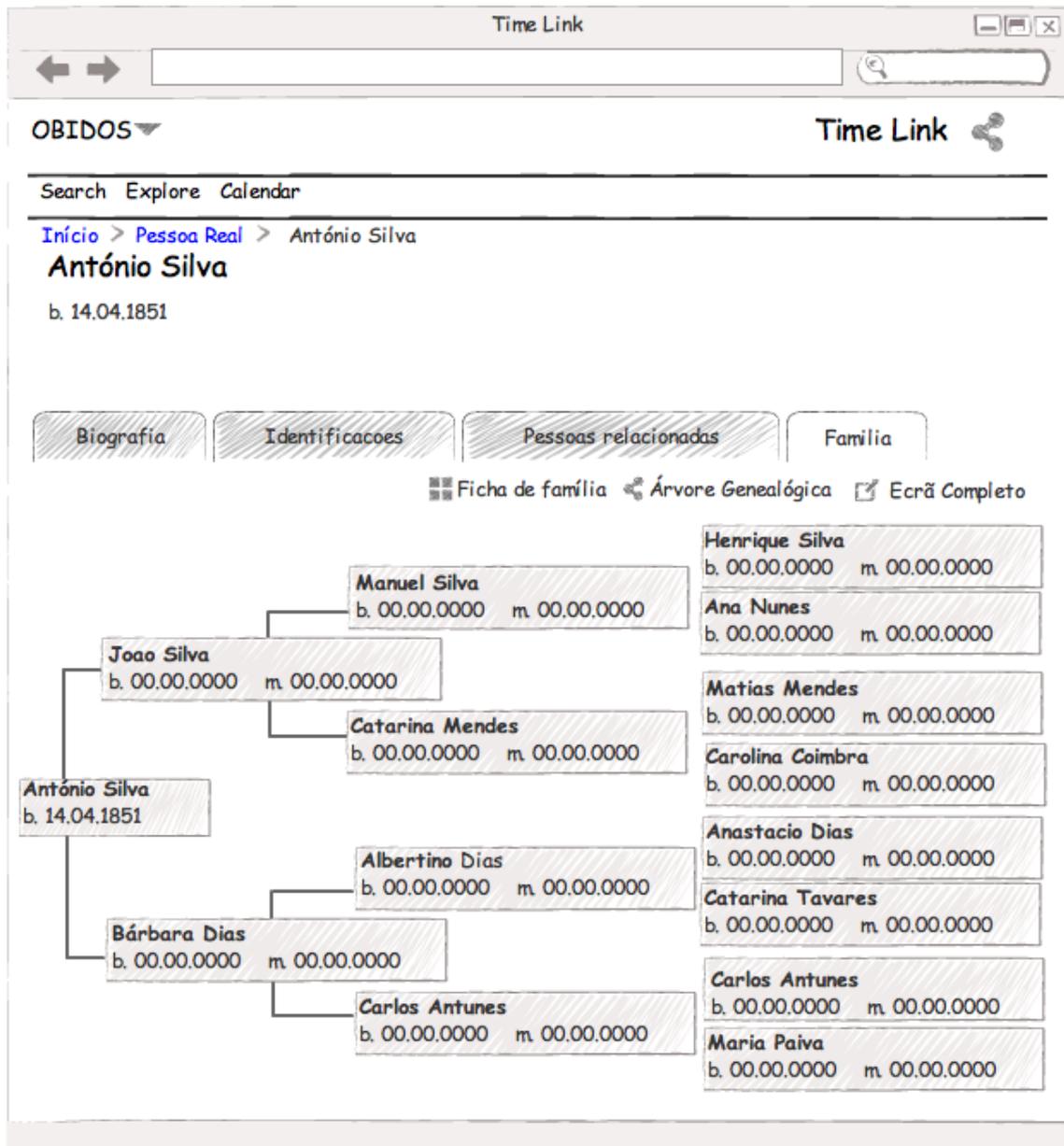


Figura 11. Vista de costados

## Exportação e Importação

A importação e exportação de informação tem como objectivo permitir a troca de informação entre diferentes aplicações. O utilizador poderá importar informação que tenha recolhido com recurso a outra aplicação informática ou mesmo exportar informação para análise exterior ao Time Link.

A exportação de informação de família processa-se através da hiperligação "Exportar GEDCOM". Desta forma será criado um ficheiro que respeita a norma GEDCOM, que inclui toda a informação da família consultada, e que poderá ser lido pela generalidade dos programas de genealogia.

A importação de informação na norma GEDCOM poderá ser realizada a partir da secção Importar da aplicação, como uma alternativa aos processos habituais de tradução e importação de ficheiros Kleio.

## Notas de Implementação

A obtenção da informação que preencherá as tabelas descritas no ponto seguinte depende da definição de conjuntos designados por tercetos. Estes tercetos fornecem informação genealógica e serão distintos em cada tipo de acto. Por exemplo, para um acto de baptismo o terceto será constituído por n (a criança baptizada), pn (o pai da criança) e mn (a mãe da criança). E, opcionalmente, pelos pais/avós e por alguns tercetos incompletos como os padrinhos/pai/mãe do padrinho. Num acto de casamento existem pelo menos dois tercetos, um para cada noivo, que serão constituídos pelo noivo/a, pnoivo/a (pai do noivo/a) e mnoivo/a (mãe do noivo/a). Neste tipo de acto poderão existir ainda outros tercetos relativos aos avós dos noivos ou mesmo decorrentes de casamentos anteriores.

Como as funções variam de acto para acto é necessário que sejam parametrizadas em local próprio, envolvendo a criação de uma nova tabela. Esta tabela denominada triples ou tercetos poderá ser modificada e aperfeiçoada constantemente, sempre que sejam adicionados novos tipos de acto ou que se identifiquem novos tercetos. A informação aí constante será usada pela componente da aplicação responsável pelo preenchimento das tabelas a seguir descritas.

fn	fpn	fmn	act
n	pai	mae	b
p	ppai	mpai	b
m	pmae	mmae	b

Tabela 6. Exemplo da tabela de tercetos para os actos de baptismo

## Estrutura de Dados

Descrevem-se agora as principais tabelas que armazenarão a informação relativa a famílias, decorrentes das considerações enumeradas no ponto anterior, e que tornarão possível a existência das funcionalidades acima descritas.

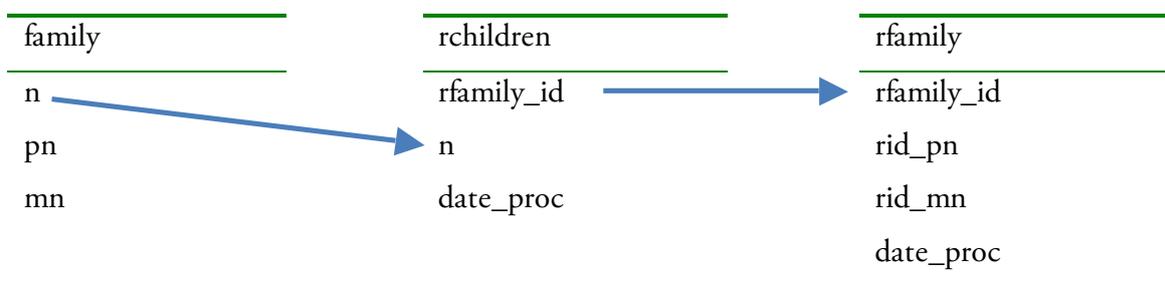


Tabela 7. Novas tabelas para armazenamento da informação genealógica.

A tabela family armazena informação relativa a famílias, que tem como ponto de partida os conjuntos de tercetos, será utilizada como auxiliar no preenchimento das tabelas rfamily e rchildren.

A tabela *rfamily* regista as matrículas correspondentes às biografias consolidadas dos pais (*rid\_pn* e *rid\_mn*). No caso de um dos pais não possuir biografia consolidada poderá ser armazenada a respectiva matrícula de ocorrência.

Por fim, a tabela *rchildren* faz a ligação entre *family* e *rfamily*, registando a matrícula de cada criança baptizada (*n*) da tabela *families* e a matrícula de família (*rfamily\_id*) da tabela *rfamilies*.

Nota técnica: é necessário alterar os procedimentos que gerem as identificações (consolidação de biografias) para garantir que as tabelas de famílias são actualizadas quando é efectuada uma nova identificação.

### 6.3.2. Identificação colectiva/cooperativa de pessoas

#### Contextualização

No Time Link, uma base de dados típica pode ser constituída por milhares de referências a pessoas, reunidas a partir de fontes diversas como são os registos paroquiais, os registos notariais, ou mesmo letras de câmbio, entre muitas outras. Interligar estas referências - operação conhecida por *record linkage*<sup>50</sup>, na expressão inglesa - e recriar as histórias de vida dessas pessoas revela-se crucial, seja em estudos de reconstituição de comunidades históricas - o objectivo que está na génese do MHK - ou noutro tipo de estudos como a genealogia. O objectivo do *record linkage* é produzir uma espécie de índice que reúne todas as ocorrências de uma pessoa (Carvalho, 88).

Desta forma, a reconstrução de biografias foi sempre uma operação nuclear do Time Link, cuja importância se torna mais clara quando se atende a um dos seus principais fundamentos, que defende a tomada de decisões em fases posteriores à recolha e inserção da informação. Estas biografias individuais são também frequentemente referidas como **pessoas reais**, por oposição às **pessoas** que são na verdade apenas ocorrências isoladas.

Neste módulo vão introduzir-se alterações ao mecanismo de identificação de pessoas que visa tornar o processo mais flexível e simples.

Na versão original do MHK a identificação de pessoas recolhidas numa base de dados ficava restrita a um único utilizador. Compreende-se esta opção pela utilização dada habitualmente à aplicação, que permanecia isolada no computador pessoal do investigador que era o único operador sobre esses dados.

Com a abertura da base de dados a um público mais abrangente, e com a possibilidade de diferentes utilizadores acederem à mesma base de dados, constatou-se a necessidade de poderem ser tomadas diferentes posições sobre esses dados e, conseqüentemente, diferentes identificações. As operações de *record linkage* podem revelar-se divergentes, pelo carácter ambíguo de muitos dados e pela maior ou menor tolerância do utilizador face à incerteza desses registos.

Ao longo das próximas linhas descrevem-se as alterações produzidas na base de dados e na

---

<sup>50</sup> A designação Record Linkage (RL) surge em primeiro lugar associada às necessidades de saúde pública dos EUA, em particular, à necessidade de reunir os registos médicos dos cidadãos, dispersos por várias fichas sem chave de identificação. A origem da expressão é atribuída a Halbert L. Dunn em 1946 (Dunn, 1946), tendo sido realizados importantes avanços na década de 50 e 60 do século XX, incluindo a teoria probabilística que é a base de muitas implementações informáticas actuais.

interface gráfica de utilizador, que permitiram a coexistência de diferentes opiniões sobre as ocorrências da vida de uma pessoa, isto é, diferentes identificações ou diferentes biografias individuais.

Em suma, os objectivos principais a concretizar neste módulo são:

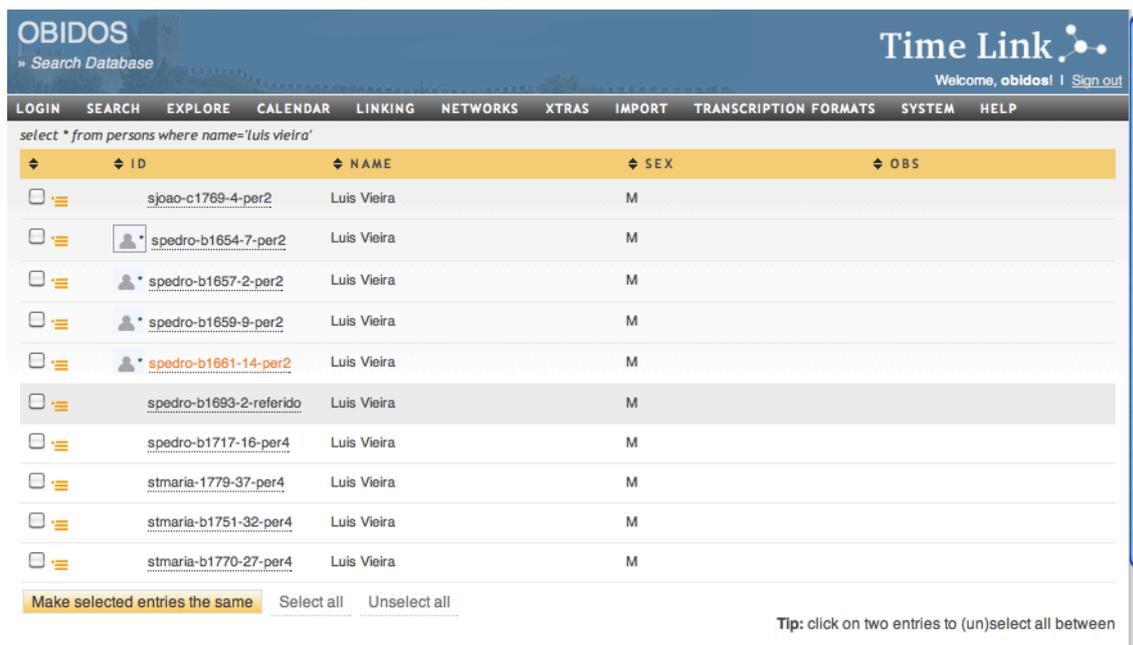
- Possibilitar a coexistência de identificações produzidas por diferentes utilizadores;
- Permitir a comparação dos resultados produzidos por cada utilizador;
- Tornar os processos de identificação mais flexíveis;
- Distinguir as ocorrências por identificar das já identificadas.

### Notas de Implementação

As alterações na estrutura da base de dados limitam-se, neste módulo, à adição de um campo suplementar na tabela *rentities*, que permita registar o utilizador responsável por determinada pessoa real (*rperson*). Esta tabela poderá também guardar, para cada entidade, a data da primeira identificação e da última alteração.

### A Interface Gráfica de Utilizador

Por outro lado, a interface gráfica será alvo de alterações bem mais significativas. A existência de caixas de verificação ao longo das várias listagens produzidas continuará a ser o principal mecanismo de interligação de ocorrências. Para simplificar também as operações de selecção poderão ser seleccionadas duas linhas distintas e todas as ocorrências intermédias serão seleccionadas.



The screenshot shows the OBIDOS Time Link interface. At the top, there is a navigation bar with 'OBIDOS' and 'Time Link' logos, and a search bar. Below the navigation bar, there is a menu with options: LOGIN, SEARCH, EXPLORE, CALENDAR, LINKING, NETWORKS, XTRAS, IMPORT, TRANSCRIPTION FORMATS, SYSTEM, HELP. The main content area displays a search query: 'select \* from persons where name='luis vieira''. Below the query, there is a table with columns: ID, NAME, SEX, and OBS. The table contains 10 rows of data, each with a checkbox in the ID column. The first row is highlighted in yellow. Below the table, there are buttons: 'Make selected entries the same', 'Select all', and 'Unselect all'. A tip at the bottom right says: 'Tip: click on two entries to (un)select all between'.

ID	NAME	SEX	OBS
<input checked="" type="checkbox"/> sjao-c1769-4-per2	Luis Vieira	M	
<input type="checkbox"/> spedro-b1654-7-per2	Luis Vieira	M	
<input type="checkbox"/> spedro-b1657-2-per2	Luis Vieira	M	
<input type="checkbox"/> spedro-b1659-9-per2	Luis Vieira	M	
<input type="checkbox"/> spedro-b1661-14-per2	Luis Vieira	M	
<input type="checkbox"/> spedro-b1693-2-referido	Luis Vieira	M	
<input type="checkbox"/> spedro-b1717-16-per4	Luis Vieira	M	
<input type="checkbox"/> stmaria-1779-37-per4	Luis Vieira	M	
<input type="checkbox"/> stmaria-b1751-32-per4	Luis Vieira	M	
<input type="checkbox"/> stmaria-b1770-27-per4	Luis Vieira	M	

Figura 12. Listagem com caixas de verificação.

Constatou-se que a ligação de duas ocorrências, que não apareçam na mesma listagem, seria demasiado complexa e implicava aceder a diferentes páginas para copiar as respectivas matrículas de identificação (ID). Uma possível solução é a recolha de referências enquanto se explora a base de dados que guardadas num 'saco' conjunto poderão ser interligadas posteriormente.

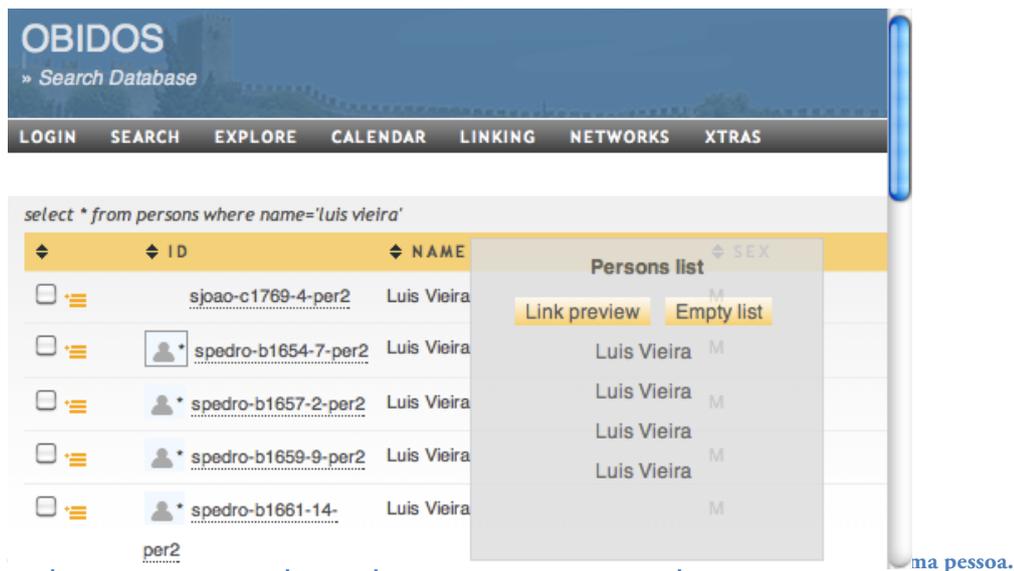


Figura 13. ... uma pessoa.

À medida que os utilizadores procedem à identificação de ocorrências coexistem ocorrências identificadas e outras que aguardam uma identificação. A distinção entre os dois é importante não só para perceber quais as ocorrências que estão por identificar, mas também quais foram registadas como pertencendo ao mesmo indivíduo. Essa distinção das ocorrências pertencentes a um mesmo indivíduo pode ser feita com recurso a elementos gráficos que marcam as ocorrências já identificadas com cores diferenciadas para cada um. As listagens poderão continuar a incluir todas as ocorrências ou, eventualmente, poderá ser dada a possibilidade de mostrar apenas a primeira ocorrência de um indivíduo, omitindo-se as seguintes para simplificar a leitura.

<input type="checkbox"/>		<a href="#">stiago-c1764-22-per3</a>	Maria Martins	F
<input type="checkbox"/>		<a href="#">stiago-c1769-44-per3</a>	Maria Martins	F
<input type="checkbox"/>		<a href="#">stiago-c1770-46-per3</a>	Maria Martins	F
<input type="checkbox"/>		<a href="#">stiago-o1756-11-per1-per3</a>	Maria Martins	F
<input type="checkbox"/>		<a href="#">stiago-o1770-6-per1-per2</a>	Maria Martins	F
<input type="checkbox"/>		<a href="#">stiago-o1776-15-per1</a>	Maria Martins	F
<input type="checkbox"/>		<a href="#">stiago-o1780-12-per1</a>	Maria Martins	F
<input type="checkbox"/>		<a href="#">stiago-o1784-16-per1-per2</a>	Maria Martins	F
<input type="checkbox"/>		<a href="#">stmaria-b1623-5-per3</a>	Maria Martins	F
<input type="checkbox"/>		<a href="#">stmaria-b1625-15-per3</a>	Maria Martins	F
<input type="checkbox"/>		<a href="#">stmaria-b1627-11-per3</a>	Maria Martins	F

Figura 14. Listagem com ocorrências identificadas e com a distinção de indivíduos através das cores.

Tornada possível a coexistência de diferentes identificações é importante comparar as decisões tomadas por outros utilizadores. O utilizador responsável por uma identificação aparece na ficha biográfica de um determinado indivíduo, identificando-se também a existência de outras opiniões sobre esse indivíduo.

OBIDOS » Real Person Info (maria martins) Time Link  
Welcome, obidos! | Help | Sign out

LOGIN SEARCH EXPLORE CALENDAR LINKING NETWORKS XTRAS IMPORT TRANSCRIPTION FORMATS SYSTEM HELP

### Maria Martins by LUISFERNADES

Other Identifications:: rp-sjoao-b1672-5-mn by obidos

Id: rp-10406 Status: N Sex: F Seq: 10406 Obs:

REAL PERSON INFO	ALL THE ACTS	RELATED PERSONS	IDENTIFICATIONS	FAMILY
YEAR	FUNCTION	ATTRIBUTES	RELATIONS	NAME
1.  1669-02-03	bap: mn	residencia : Sobral IMAPI	mae of Maria has as marido: Joao Henriques	Maria Martins
2.  1672-02-29	bap: mn	residencia : Sobral IMAPI	mae of Antonia has as marido: Joao Henriques	Maria Martins
3.  1672-02-29	bap: mn	residencia : Sobral IMAPI	mae of Maria has as marido: Joao Henriques	Maria Martins
4.  1677-08-17	obito: mulher	residencia : Sobral IMAPI	has as marido: Joao Henriques	Maria Martins
5.  1677-09-28	bap: mn	residencia : Sobral IMAPI	mae of Manuel has as marido: Joao Henriques	Maria Martins

A comparação das identificações faz-se a partir do separador ‘Identificações’ acessível a partir da mesma ficha biográfica. Nesse separador são apresentadas as diferentes identificações produzidas e marcam-se as ocorrências que foram tidas em consideração por cada utilizador. No caso de posições divergentes identificam-se também com cores diferentes as ocorrências que foram ligadas a indivíduos diferentes. Para simplificar o processo de apresentação da informação poder-se-á considerar mostrar uma única vez, numa única coluna, todas as posições concordantes e identificar os vários utilizadores responsáveis.

OBIDOS » Identifications (maria martins) Time Link  
Welcome, obidos! | Help | Sign out

LOGIN SEARCH EXPLORE CALENDAR LINKING NETWORKS XTRAS IMPORT TRANSCRIPTION FORMATS SYSTEM HELP

### Maria Martins by LUISFERNADES

Other Identifications:: rp-sjoao-b1672-5-mn by obidos

Id: rp-10406 Status: N Sex: F Seq: 10406 Obs:

REAL PERSON INFO	ALL THE ACTS	RELATED PERSONS	IDENTIFICATIONS	FAMILY			
YEAR	ID	LUISFERNADES	OBIDOS	FUNCTION	ATTRIBUTES	RELATIONS	NAME
1. <input checked="" type="checkbox"/> 1669-02-03	sjoao-b1669-9-per3			bap: mn	residencia : Sobral IMAPI	mae de Maria tem como marido: Joao Henriques	Maria Martins
2. <input checked="" type="checkbox"/> 1672-02-29	sjoao-b1672-5-per3			bap: mn	residencia : Sobral IMAPI	mae de Antonia tem como marido: Joao Henriques	Maria Martins
3. <input checked="" type="checkbox"/> 1672-02-29	sjoao-b1672-5a-mn			bap: mn	residencia : Sobral IMAPI	mae de Maria tem como marido: Joao Henriques	Maria Martins
4. <input checked="" type="checkbox"/> 1677-08-17	sjoao-o1677-11-per2			obito: mulher	residencia : Sobral IMAPI	tem como marido: Joao Henriques	Maria Martins
5. <input checked="" type="checkbox"/> 1677-09-28	sjoao-b1677-25-per3			bap: mn	residencia : Sobral IMAPI	mae de Manuel tem como marido: Joao Henriques	Maria Martins

\* Users that identified this person=[luisfernaDES, obidos]

Figura 16. Separador de ‘Identificações’ da Pessoa Real.

### 6.3.3. Integração de mapas (Google Maps)

O tratamento de informação geográfica no Time Link encontra-se ainda em fase de consolidação, resultado de trabalhos de investigação em curso que vieram dar um novo impulso a este tipo de informação. A própria estrutura de dados já prevê a existência de geoentidades, com atributos próprios, mas ainda sem visualização gráfica. Permitir a visualização de informação geográfica sobre um mapa é fundamental, não só para investigadores, mas também para a generalidade dos utilizadores, pois permite descobrir detalhes de outra forma invisíveis.

A apresentação da informação no espaço será concretizada com recurso ao Google Maps e à respectiva API (Application Programming Interface)<sup>51</sup>. Esta API disponibiliza as ferramentas necessárias para a integração de mapas em aplicações Web, incluindo a possibilidade de adicionar marcadores, linhas, linhas poligonais e polígonos, entre outros elementos.

Será também analisada a utilização dos mecanismos de *geocoding* disponibilizados pela Google Maps API. A utilização desta funcionalidade é particularmente interessante como mecanismo auxiliar de georreferenciação de atributos e entidades geográficas, tendo a capacidade de acelerar significativamente o tratamento da informação geográfica.

#### A Interface Gráfica

Identificam-se, desde logo, cinco situações onde a visualização em mapa pode ser particularmente relevante.

1. Ficha de local: agrega actos e entidades para um determinado local.
2. Ficha de pessoa real: mostrando todos os atributos georreferenciáveis da pessoa: residência, naturalidade, loccasamento, etc...
3. Ficha de pessoa real - separador de pessoas relacionadas: apresentar no mapa um ponto por cada atributo georreferenciado de cada pessoa relacionada, e desenhar uma linha em relação à pessoal real em questão.
4. Redes no espaço: um ponto por cada nó e uma linha por cada aresta. Quando um nó não tem localização, determina-se a localização a partir da localização dos conhecidos e usando um algoritmo de 'layout'.
5. Pesquisa geográfica: permitir pesquisar pessoas que vivem a uma distância determinada ou visualizar outros locais que fiquem próximos do actual.

---

<sup>51</sup> <http://code.google.com/apis/maps/index.html>

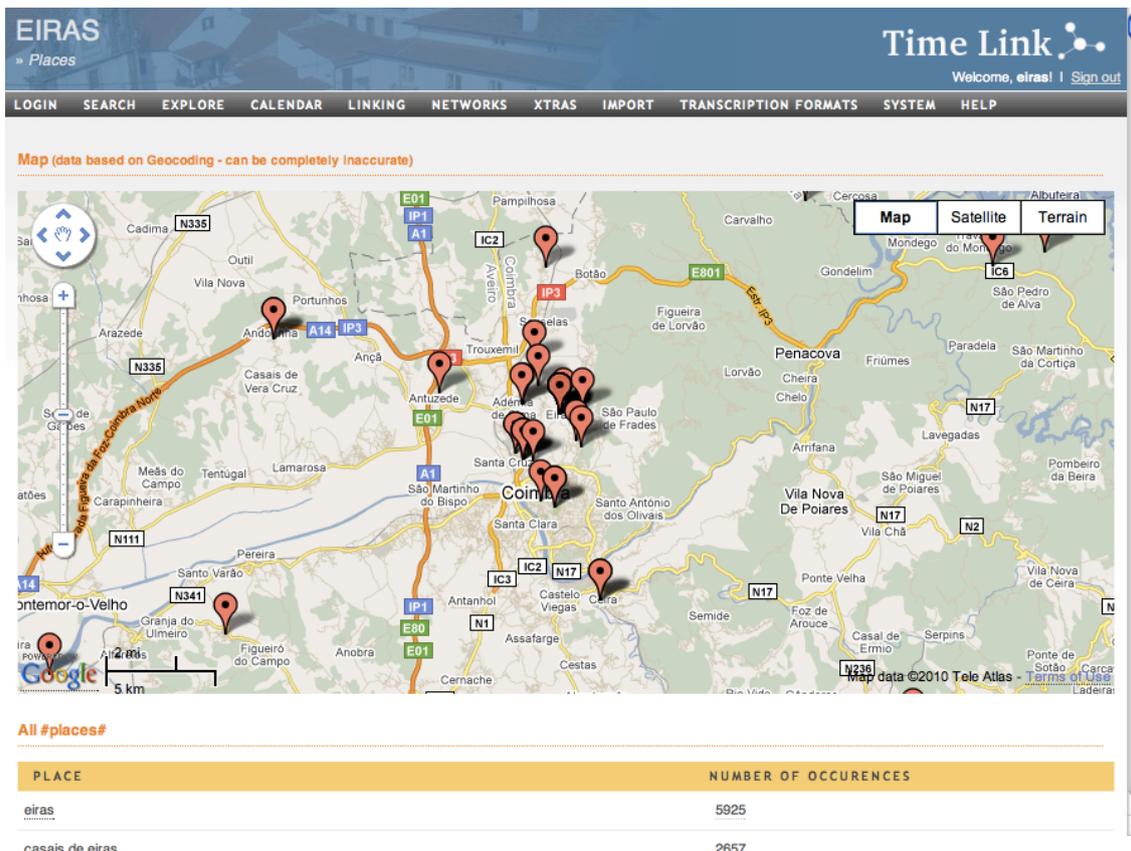


Figura 17. Visualização com as referências espaciais da base de Eiras-Coimbra, recorrendo à Google Maps API.

## Notas de Implementação

A informação geográfica, como referido no início deste ponto, está disponível na base de dados do Time Link e é constituída por entidades geográficas, ou geoentidades, e atributos de carácter geográfico, como locais de nascimento e óbito, residência, entre outros. Esta informação, no seu estado actual, não possui quaisquer coordenadas geográficas que permitam apresentá-la sobre um mapa. Desta forma, é fundamental proceder a alterações na base de dados que permitam o armazenamento de informação georreferenciada. As novas tabelas deverão guardar as coordenadas e designação de cada entidade geográfica e relacioná-las com os atributos geográficos.

Para fazer corresponder a cada referência geográfica as respectivas coordenadas propõem-se duas soluções, numa visão de simplificação do processo e aceleração da inserção de informação. Numa solução seriam usadas tabelas de toponímia, com coordenadas geográficas que permitiriam propor para cada referência geográfica existente na base de dados do Time Link uma ou mais correspondências. Neste processo é necessário lidar com variações de ortografia, com a existência de localidades com nome iguais ou semelhantes e mesmo e a possibilidade de diferentes designações (seja num mesmo período ou em diferentes períodos). Neste caso uma implementação passa pela utilização de uma tabela de equivalências de atributos georreferenciáveis, onde se regista que para o atributo X, os valores Y e Z são equivalentes.

A API do Google Maps permite geocoding<sup>52</sup>, isto é, a obtenção de coordenadas geográficas a partir de outra informação descritiva, como um endereço completo, ou mesmo um

<sup>52</sup> <http://code.google.com/apis/maps/documentation/javascript/services.html#Geocoding>

topónimo simples. Assim, remetendo um topónimo ao Google Maps consegue obter-se uma lista de locais prováveis, com as respectivas coordenadas geográficas que tornam possível a exibição em mapa.

A aplicação do Geocoding ao tipo de informação geográfica disponível no Time Link levanta também algumas questões. É não só a ambiguidade da informação, que frequentemente refere apenas o nome de uma localidade, sem informação administrativa adicional (sabemos que podem existir várias localidades com o mesmo nome), ou mesmo o carácter histórico dessa informação (muitos topónimos deixaram de existir ou mudaram de nome, o que faz com que não apareçam nas bases de dados actuais como a do Google Maps).

As duas soluções brevemente apresentadas - a tabela de toponímia e os servidores de *geocoding* - não são mutuamente exclusivas e poderão mesmo assumir um papel complementar, dependendo do tipo de dados que estão a ser tratados em determinado momento. Em qualquer dos casos, a validação da informação pelos utilizadores, e o seu registo na base interna do Time Link, é um aspecto fundamental.

#### 6.3.4. Comentários

##### Contextualização

Os comentários são umas das formas mais antigas de conteúdo gerado por utilizadores (*user-generated content*) na Internet. Os *newsgroups* e as *mailing lists* electrónicas tornaram possível a partilha de informação e a troca de opiniões através da Internet, muito antes do advento da Web 2.0 e mesmo da World Wide Web. Os comentários são na verdade uma transposição desse mesmo princípio para os *websites*, sejam blogues, *sites* noticiosos, redes sociais *online* ou mesmo locais de comércio electrónico.

O introdução de um sistema de comentários no Time Link segue os mesmos princípios: incentivo à participação da comunidade, troca de opiniões e partilha de informação suplementar. Além da troca de mensagens textuais simples seria igualmente interessante considerar a possibilidade de carregamento de imagens que sirvam de ilustração à mensagem do utilizador.

Porém, para além de tornar possível essa participação activa dos utilizadores, é necessário assegurar a qualidade da informação e o bom funcionamento de todo o sistema. Um sistema de comentários, mesmo restrito a utilizadores registados, está sujeito a agressões externas como são as mensagens de publicidade abusiva, linguagem imprópria ou mesmo de teor ilegal. Com o crescimento do número de utilizadores registados e participativos será cada vez mais importante a existência de um sistema de moderação que poderá também ele ser gerido pela própria comunidade. É um modelo usado com frequência por websites com uma participação comunitária significativa<sup>53</sup>.

Em síntese, os principais objectivos deste módulo são:

- Permitir a colocação de comentários associados às diferentes entidades da base de dados (actos, ocorrências e biografias/fichas de pessoa real);

---

<sup>53</sup> A título de exemplo, o Slashdot, que é um dos mais antigos sites com uma forte participação comunitária - reunia aproximadamente 17 milhões de comentários em Novembro de 2009 (<http://slashdot.org/articles/06/11/09/1534204.shtml>) -, utiliza uma política de moderação (<http://slashdot.org/faq/com-mod.shtml>) baseada nos seus utilizadores: "Today any regular Slashdot reader is probably eligible to become a moderator".

- Assegurar a moderação dos comentários através da participação dos utilizadores regulares da plataforma.

### Notas de Implementação

Uma nova tabela 'comments' ficará responsável por guardar todos os comentários produzidos pelos utilizadores. Nessa tabela ficarão registados a data de colocação do comentário, o comentário propriamente dito, o nome do utilizador e a entidade a que refere esse mesmo comentário.

ID	comment	created	user	score	entity_id
(varchar)	(text)	(timestamp)	(varchar)	(int)	(varchar)

Figura 18. Tabela de comentários de utilizador

Cada comentário poderá ser classificado por utilizadores autorizados ou que reúnam suficiente reputação - esta reputação no sistema poderá depender da antiguidade do registo do utilizador, da sua participação, da classificação média obtida, entre outros factores - e essa classificação ficará disponível em conjunto com o comentário. A este sistema de moderação dependente dos utilizadores, poderá ser associado um mecanismo de detecção automática de SPAM<sup>54</sup> que permitirá criar um primeiro filtro.

### Interface Gráfica

O sistema de comentários do Time Link permite o carregamento de informação textual identificada por um título e pelo nome do utilizador responsável, e que ficará associada a actos, a ocorrências isoladas e a fichas de pessoa real/biografias.

A partir da página de entrada será possível consultar os últimos comentários dos utilizadores e aceder directamente à entidade a que se reportam. Será também interessante tornar possível ao utilizador a consulta de todos os comentários por si colocados. Estes comentários poderão ainda desempenhar um papel importante na criação de reputação associada a cada utilizador.



Figura 19. Ecrã com lista de comentários na página de entrada

<sup>54</sup> Termo usado habitualmente para designar mensagens não solicitadas ou inadequadas em determinado contexto.

The screenshot shows a web browser window with the URL `timelink.fl.uc.pt/mhk/obidos/id/sjoao-b1685-7-per1`. The page header includes the 'OBIDOS' logo and 'Time Link' branding. A navigation menu contains links for LOGIN, SEARCH, EXPLORE, CALENDAR, LINKING, NETWORKS, XTRAS, IMPORT, TRANSCRIPTION FORMATS, SYSTEM, and HELP. The main content area displays the profile for 'Maria' with the following details: **id:** sjoao-b1685-7-per1, **Groupname:** N, **Sex:** F, **Line:** 1734. Below this is a 'PERSON INFO' section with a table:

YEAR	FUNCTION	ATTRIBUTES	RELATIONS	NAME
1685-03-04	bap: n		tem como pai:  domingas luis tem como mae:  maria francisca	maria

Below the table is a 'User comments' section containing the following text:

Maria, com o Id:sjoao-b1685-7-per-1  
por Luisfernandes, 2009-08-19 23:08:39.0  
Maria, com o Id:sjoao-b1685-7-per-1, Tem pai Domingas luis e mae Maria Francisca. Algo está errado

Figura 20. Ecrã com um comentário agregado a uma entidade da base de dados

A partilha de outro tipo de informações, para além de texto, é igualmente interessante, e poderá ser considerado o carregamento de outro tipo de conteúdos, sobretudo imagens ou ilustrações, assim como digitalizações de registos paroquiais, por exemplo.

### 6.3.5. Introdução de dados por utilizadores

#### Contextualização

O processo habitual de inserção de dados no Time Link implica a tradução de documentos de texto com as transcrições das fontes históricas (ver capítulo 2). Pelas razões enunciadas nesse segundo capítulo, continuará a ser o mecanismo preferencial para as instalações locais. No entanto, nas bases de dados de acesso público na Internet, a inserção de novos dados pelos utilizadores levanta alguns desafios adicionais. Por lado, é expectável que a maioria dos utilizadores não tenha interesse, nem disponibilidade, para aprender a notação Kleio, mas ainda assim seria interessante que pudessem carregar algum tipo de informação. Por outro lado, um grupo mais restrito de utilizadores poderia interessar-se pelo carregamento de maiores volumes de informação e nesse caso seria interessante a utilização do método de tradução e importação.

Para atender a esta eventual dualidade, os utilizadores poderão ter à sua disposição duas formas distintas de inserção de dados. A primeira forma, mais simples, permitiria inserir dados através de formulários tradicionais - ou métodos mais graficamente apelativos sob a forma de árvore genealógica - que permitem a inserção de informação relativamente padronizada, como a informação que se obtém a partir de registos paroquiais. Em alternativa seria sempre possível recorrer a ficheiros Kleio que carregados no sistema, ou mesmo editados directamente no *browser*, seriam posteriormente traduzidos e importados.

Além dos desafios de inserção da informação há também questões relacionadas com a validação e a qualidade da mesma. Uma possível solução será utilizar mecanismos de regulação semelhantes aos apresentados no sistema de comentários. De qualquer forma, a informação que é carregada pelos utilizadores deverá estar claramente identificada na base

de dados, e na interface. Os restantes utilizadores poderão então atribuir pontuação e mesmo avaliar a credibilidade e reputação no sistema do utilizador responsável por determinados dados.

### Notas de Implementação

Para permitir a inserção de informação pelos utilizadores são necessárias alterações na estrutura de dados existente, tendo em vista o registo da autoria da informação e garantir a estanquidade da informação atrás referida.

O processo de inserção tradicional da informação povoa as diferentes tabelas com a informação proveniente das transcrições em formato Kleio. As tabelas afectadas pelo processo são diversas, podendo mesmo, em determinados casos, ser criadas novas tabelas. O modelo de dados descrito no segundo capítulo permite perceber que a informação entra nas tabelas obedecendo a uma hierarquia de especialização. Desta forma, sempre que é registada informação na tabela 'persons', esta é igualmente registada na tabela 'entities'. Assim, o registo das alterações produzidas pelos utilizadores é possível com a criação de uma nova coluna 'user' na tabela 'entities', onde fica registado o utilizador responsável pela inserção da informação, acompanhado pelas datas de inserção ou alteração.

Id	Class	Inside	The_order	The_level	The_line	Groupname	Username	Created	Updated	indexed
(varchar)	(varchar)	(varchar)	(decimal)	(decimal)	(decimal)	(varchar)	(varchar)	(timestamp)	(timestamp)	(timestamp)

Tabela 8. Tabela de entities

### 6.3.6. Suporte a múltiplos idiomas

A interface do Time Link permite a apresentação em vários idiomas, estando disponível na versão actual em português e inglês. A adição de novas traduções faz-se com recurso a ficheiros simples de propriedades, isto é, ficheiros de texto que em cada linha apresentam uma chave acompanhada pela tradução na língua respectiva.

Os ficheiros de tradução estão guardados na pasta *resources* do código-fonte. Cada língua suportada deverá incluir um ficheiro com o nome MHKStrings, seguido pelo código ISO<sup>55</sup> que identifica a língua internacionalmente. Por exemplo, o ficheiro de tradução para português denomina-se MHKStrings\_pt.properties. Existe ainda um ficheiro MHKStrings.properties, sem indicação do idioma, que inclui a tradução inglesa e é o ficheiro de base utilizado sempre que não existe uma tradução para a língua do utilizador.

Todos os *browsers* Web actuais permitem a definição de um idioma padrão a ser utilizado na apresentação das páginas, permitindo mesmo a existência de diferentes níveis de prioridade (se um *website* não disponibiliza determinado idioma deverá utilizar-se o seguinte, e assim por diante). Na sua implementação actual o Time Link utiliza estas definições do *browser* para decidir a versão a apresentar ao utilizador.

A tradução de elementos relativos aos dados é mais complexa e não é ainda assegurada no modelo actual. Existem várias soluções possíveis, desde a sua tradução através de uma tabela de equivalências existente na base ou mesmo a possibilidade de colocar a tradução directamente nos ficheiros onde se realiza a definição das fontes (gacto2.str). Neste

<sup>55</sup> ISO 639-1:2002, Codes for the representation of names of languages — Part 1: Alpha-2 code

segundo caso poderia existir um campo adicional que armazenava a definição de um campo noutras línguas, por exemplo, pai (en:father,fr:père).

```
# Portuguese Localization File
# Version 0.1
language.code=pt

#Generic
time.link=Time Link
please.wait=Por favor, aguarde...
logout=Terminar sessão
year=Ano
success=Operação efectuada com sucesso
function=Função
attributes=Atributos
relations=Relações
```

Figura 21. Exemplo de ficheiro de tradução.

### 6.3.7. URL's únicos

Na sua versão original, antes do presente trabalho, o MHK estava muito focado na existência de uma única estação de trabalho, isto é, o computador pessoal de um único utilizador. Mesmo permitindo a existência de múltiplas bases de dados a implementação não permitia distingui-las a partir do respectivo endereço (URL)<sup>56</sup>. Numa aplicação disponível publicamente na Internet, que deverá agregar diferentes bases de dados, é fundamental assegurar que um endereço é único, e que inclui informação relativa à base de dados seleccionada.

Para além deste carácter inequívoco, é importante assegurar a legibilidade desse mesmo endereço, simplificando-o na medida do possível.

Utilizou-se na implementação uma biblioteca de código aberto conhecida por `UrlRewriteFilter`<sup>57</sup>, que permite lidar com as questões de mapeamento entre dois endereços Web equivalentes: um que pode ser utilizado internamente e outro que é apresentado ao utilizador final. A configuração faz-se através de um ficheiro de configuração (`urlrewrite.xml`) que define as regras de mapeamento entre os endereços internos e os endereços apresentados.

---

<sup>56</sup> URL, Uniform Resource Locator, é o endereço de um recurso acessível em rede (numa rede do tipo intranet ou na Internet).

<sup>57</sup> A `UrlRewriteFilter`, baseada no `mod_rewrite` utilizado pelo servidor Web Apache, é uma biblioteca de código aberto, disponibilizada sob uma licença BSD compatível com a licença do Time Link.

<http://code.google.com/p/urlrewritefilter/>

```

<?xml version="1.0" encoding="iso-8859-1"?>
<!DOCTYPE urlrewrite PUBLIC "-//tuckey.org//DTD UrlRewrite 3.2//EN"
    "http://tuckey.org/res/dtds/urlrewrite3.2.dtd">
<urlrewrite>
  <rule>
    <from>/login/</from>
    <to>/servlet/do?action=login.vm</to>
  </rule>
  <rule>
    <from>/signup/</from>
    <to>/servlet/do?action=signup.vm</to>
  </rule>
  <rule>
    <from>/([a-z]+)/id/([a-z0-9]+)/</from>
    <to>/servlet/do?dbname=$1&action=show&id=$2</to>
  </rule>
</urlrewrite>

```

Figura 22. Exemplo do ficheiro de configuração

O resultado final são endereços únicos, que incluem todos os elementos necessários à identificação da base de dados, e também uma maior legibilidade.

<b>Original</b>	http://timelink.fl.uc.pt/mhk/servlet/do?action=exec&script=search.vm
<b>Final</b>	http://timelink.fl.uc.pt/mhk/ <b>obidos</b> /show/search.vm

<b>Original</b>	http://timelink.fl.uc.pt/mhk/servlet/do?action=show&id= rp-not1704-e1-per4
<b>Final</b>	http://timelink.fl.uc.pt/mhk/ <b>obidos</b> /id/rp-not1704-e1-per4

### 6.3.8. Interligação entre diferentes bases de dados

O Time Link, à semelhança do que acontecia na versão original do MHK, permite a coexistência de bases de dados distintas, por norma dedicadas a uma paróquia ou a um determinado arquivo histórico.

No MHK, a transição entre bases de dados efectuava-se através de um *login* constituído por utilizador e palavra-chave. Existia uma correspondência directa entre o utilizador e a base de dados, relação que não seria de todo viável numa aplicação disponível na Internet, que pretende albergar múltiplas bases de dados e diferentes utilizadores.

Com o Time Link e o objectivo de disponibilização *online*, foi necessário romper com esta ligação directa entre o utilizador e a base de dados, por forma a permitir o acesso de um utilizador a múltiplas bases.

## Notas de Implementação

Criou-se o conceito de um super-utilizador, do qual estão dependentes os restantes utilizadores registados. Este utilizador possui não só todos os privilégios de administração do sistema, mas serve também de referência para os restantes, por exemplo, nas operações de reconstrução de biografias. Poderá vir a ser também o utilizador responsável pelas identificações produzidas no futuro por um sistema de *record linkage* automático.

Após uma autenticação bem sucedida, o utilizador pode consultar e transitar entre bases de dados, sem necessidade de nova autenticação. O conjunto de bases de dados disponíveis e acessíveis poderá ser diferente para cada utilizador, dependendo do respectivo perfil de identificação.

Na implementação actual, as informações relativas ao registo do utilizador continuam a ser obtidas a partir de um ficheiro de propriedades, mas poderão, no futuro, ser facilmente obtidas a partir de uma base de dados de utilizadores. O registo de utilizadores é abordado com maior detalhe no ponto seguinte.

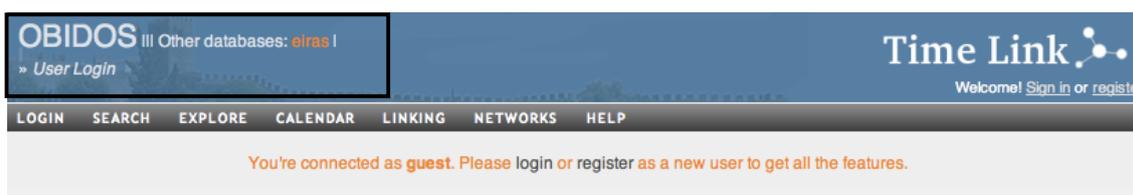


Figura 23. Indicação da base de dados seleccionada actualmente e das restantes bases de dados acessíveis para o utilizador.

### 6.3.9. Factores de infra-estrutura: instalação e registo de utilizadores

O Time Link pode ser instalado de forma isolada ou em servidores Web acessíveis na Internet. O primeiro tipo de instalação é usada sobretudo por investigadores profissionais ou amadores da prosopografia e da genealogia, e permite ter no computador pessoal todos os elementos necessários à execução do Time Link e ao tratamento de informação em modo *off-line*. A instalação em servidores Web acessíveis através da Internet é especialmente interessante para bases de dados de acesso público ou mesmo para bases de acesso reservado mas que implicam trabalho de equipa.

A instalação é possível em diferentes plataformas, assim existam implementações das tecnologias que servem de suporte e já anteriormente identificadas: MySQL, Tomcat e Java. Foram testadas com sucesso instalações em sistemas MS Windows, MacOS X e GNU/Linux. As instalações em ambiente Linux são particularmente úteis e interessantes em bases de dados acessíveis na Internet, dado o excelente desempenho que estes sistemas permitem obter.

O registo de utilizadores nas instalações *online* é efectuado através de um formulário electrónico acessível a partir da página de entrada. A partir deste formulário recolhem-se as informações do utilizador, como o nome ou senha pretendida. O utilizador recebe posteriormente uma mensagem de confirmação no endereço de correio electrónico que indicou. Após essa confirmação, a conta do utilizador deverá ficar activa e pronta a ser usada.

Figura 24. Formulário de registo de utilizadores online.

### MHK/Time Link Portable

A construção do Time Link como uma aplicação Web trouxe muitas vantagens, muitas referidas nos pontos anteriores, como a fácil portabilidade entre diferentes plataformas ou a possibilidade ser usada em servidores acessíveis publicamente por um conjunto bastante alargado de utilizadores. No entanto, esse carácter baseado nas tecnologias Web acarreta também algumas desvantagens, especialmente notórias no momento da primeira instalação de todo o sistema. Numa instalação tradicional do Time Link é necessário tratar primeiro de todos os pacotes de *software* identificados como a base tecnológica – onde se incluem o servidor de bases de dados (MySQL), o servidor de aplicações Web (Apache Tomcat), os binários Java e ainda o Prolog, utilizado pelo tradutor. Acresce a necessidade de instalar manualmente a aplicação Web propriamente dita e configurar uma primeira base de dados em branco.

É neste contexto que surge um pacote de software conhecido por MHK Portable ou Time Link Portable. O objectivo principal deste pacote é contornar as dificuldades enunciadas no parágrafo anterior e permitir uma instalação simplificada.

A primeira versão deste pacote auto-contido foi criada para ambientes MS Windows e permite correr uma versão completa do Time Link a partir de uma pasta que é simplesmente descompactada no disco do utilizador. Este pacote contém um executável que permite lançar de uma só vez todos os componentes necessários e inclui, ainda, uma função de auto-actualização. Sempre que o utilizador inicia a aplicação, esta tem a capacidade de comunicar com um servidor central e verificar a existência de novas versões

da aplicação Web e, caso positivo, realizar o *download* e concluir todo o processo de actualização.

Lançados todos os componentes necessários é aberta uma janela do *browser* Web com a página de entrada do Time Link e colocado um ícone na barra de tarefas do Windows, a partir do qual é possível fechar ou reiniciar os serviços do Time Link.

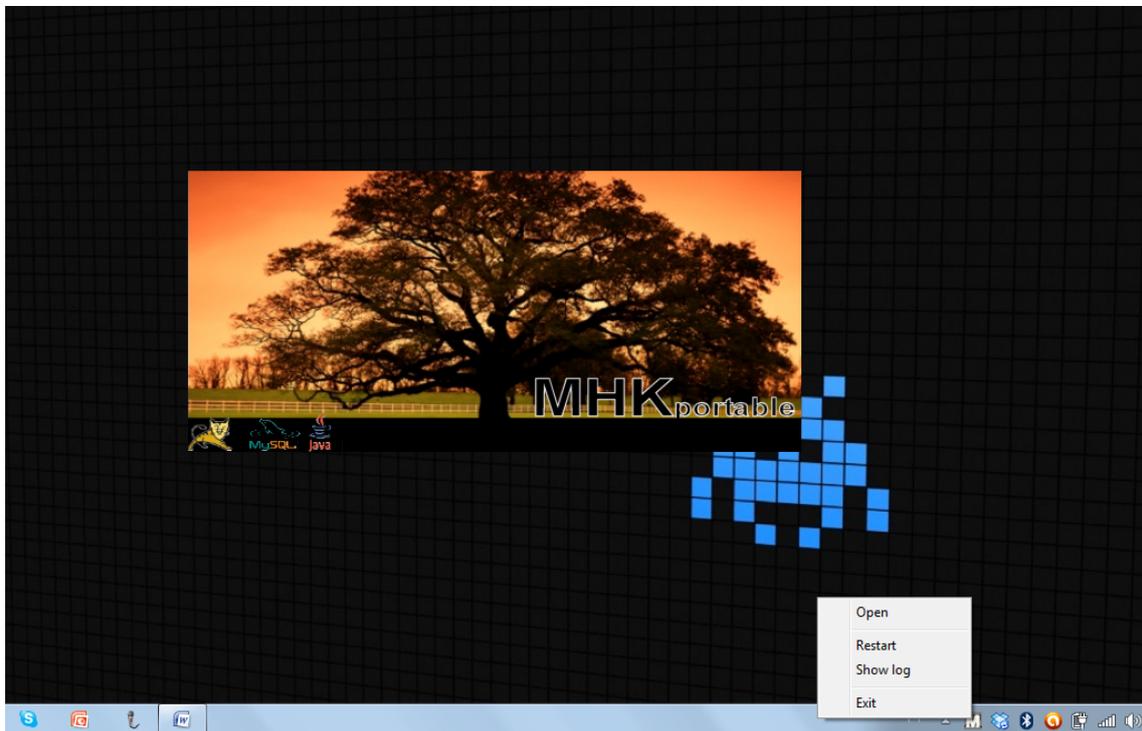


Figura 25. MHK/Time Link Portable a iniciar e respectivo ícone na barra de tarefas do Windows.

## Considerações finais

O objectivo principal do presente trabalho foi a transformação de uma ferramenta de uso académico, relativamente desconhecida, quase secreta, em algo acessível a qualquer pessoa, em qualquer parte do mundo onde existam um computador e uma ligação à Internet.

O MHK, que surgiu como uma aplicação de apoio à investigação em Micro-história e à Reconstituição de Comunidades Históricas, preparou-se para dar o salto, que permitiu finalmente uma outra visibilidade e essa ambicionada acessibilidade global.

Em 2008, resultado de uma colaboração entre a Faculdade de Letras da Universidade de Coimbra e a Câmara Municipal de Óbidos, constituía-se uma das bases de dados mais completas, tirando partido do MHK como plataforma tecnológica. Concretizada a base de dados era fundamental disponibilizá-la ao público, torná-la visível. Terá sido, de certo modo, o ponto de partida e o impulso necessário para concretizar muitas das transformações que aqui foram apresentadas.

No ano seguinte, em 2009, o Time Link ficou por fim acessível na Internet, ainda em processo de testes, devidamente acompanhado pela primeira base de dados de acesso público, que reunia os registos paroquiais de Óbidos.

A disponibilização pública do Time Link, em conjunto com a base de dados de Óbidos, teve repercussões interessantes, permitindo obter dados importantes sobre a receptividade e a projecção junto do público. À projecção mediática, obtida pela reportagem da RTP do Bom Dia Portugal e mesmo em alguma imprensa regional, seguiu-se um ritmo constante de visitas diárias, acompanhado por pedidos de acesso registado à base. Mesmo com toda a informação acessível publicamente, sem necessidade de registo, muitos preenchiam os formulários de adesão, solicitando a criação de contas de utilizador registado.

Mais recentemente, foi tornada pública uma segunda base de dados, desta vez com registos paroquiais e notariais de Eiras-Coimbra, resultante do trabalho de investigação de Ana Isabel Ribeiro (Ribeiro, 2005). As estatísticas de acesso ao endereço *timelink.fl.uc.pt*, recolhidas através das ferramentas de análise do Google Analytics, entre Março de 2009 e Outubro de 2010, revelaram mais de 15.000 visitas, com origem em 109 países diferentes, e que correspondem a 219 mil visualizações de páginas (*pageviews*). Mesmo descontando o tráfego directo<sup>58</sup>, que representa 45% dos acessos – admitindo-se que uma parte deste tráfego, embora não a sua totalidade, poderá resultar de movimentos de utilização interna – fica-se com um volume de visitas exteriores superior a 8 mil visitas, visitas estas que têm como origem os motores de pesquisa (35%) ou são direccionadas a partir de outros *websites* (20%).

No entanto, a exposição pública da base de dados e a sua abertura a um público mais alargado não significou, de modo algum, o seu término enquanto aplicação de investigação. O Time Link é utilizado actualmente como a plataforma tecnológica de vários projectos de investigação em universidades, incluindo o importante projecto de investigação Dyncoopnet, ou Dynamic Complexity of Cooperation-Based Self-Organizing Networks in the First Global Age), conduzido pela Faculdade de Letras do Porto e integrado no programa de investigação internacional TECT (The Evolution of Cooperation and Trading). Este último projecto utiliza uma fonte distinta das até aqui referidas, no caso

---

<sup>58</sup> Acesso directo refere-se a acessos realizados na posse do respectivo endereço Web, sem recurso a motores de pesquisa ou redireccionamentos de outros *websites*.

letras de câmbio e mesmo correspondência provenientes do arquivo de um importante homem de negócios do século XVI.

Também encorajador foi receber testemunho da primeira utilização autónoma e independente do Time Link, utilização essa que ocorreu sem qualquer tipo de orientação ou formação prévia da equipa de Coimbra. A preparação de um guia de transcrição, da responsabilidade de Ana Isabel Ribeiro, e a simplificação dos mecanismos de instalação do Time Link, aqui referidos como Time Link Portable, terão sido fundamentais para mais este pequeno sucesso.

Os resultados alcançados constituem assim um importante estímulo para a prossecução do trabalho aqui iniciado. O Time Link não é um produto terminado nem definitivamente encerrado. É um sistema em aberto, que continuará receptivo a melhorias e correcções. O código fonte da aplicação, acessível em [mbk.googlecode.com](http://mbk.googlecode.com), reflecte também essa abertura e permite que qualquer um possa dar a sua contribuição para o sistema ou simplesmente adaptá-lo de acordo com as suas necessidades.

O último capítulo do presente trabalho, para além de expor as transformações efectivamente concretizadas, deixa a porta aberta a novas melhorias que poderão ser introduzidas em versões futuras. A geração de redes de pessoas é um desses exemplos onde, apesar e todos os esforços realizados, continua a existir um importante espaço de progressão. O mesmo acontece na integração com Sistemas de Informação Geográfica ou na exportação, e sobretudo importação, de informação genealógica em formato GEDDOM. Os exemplos abundam e permitem perceber que o futuro Time Link será ainda melhor.

## Bibliografia

- Allen, James B., Embry, Jessie L. e Kahlile B. Mehr. 1994. "Chapter 8: Automating the Records", *BYU Studies*, vol. 34 , n.º 2.
- Allen, Paul. 2009. "My Companies." visitado em 9 de Setembro de 2009. <http://www.paulallen.net/my-companies/>.
- Amaral, Augusto Ferreira do. 2000. *Fontes da genealogia em Portugal*, Porto: Centro de Estudos de Genealogia, Heráldica e História da Família, Universidade Moderna.
- Andereck, Paul Arthur e Pence, Richard A. 1985. *Computer genealogy: a guide to research through high technology*, Salt Lake City: Ancestry.
- Araújo, Maria Marta Lobo de. 2007. "Os arquivos das misericórdias do Alto Minho: um itinerário de investigação", *Cadernos Vianenses*, Tomo 40.
- Axmark, David e Widenius, Michael. 1999. "MySQL Introduction", *Linux Journal*, 11 de Janeiro de 1999, visitado em 9 de Setembro de 2009. <http://www.linuxjournal.com/article/3609>.
- Battista, John M. 2009. "Use Cases", visitado em 17 de Outubro de 2009 <http://www.johnmichaelbattista.me/software-engineering/use-cases>.
- Brittain, Jason and Darwin, Ian F. 2008. *Tomcat: The Definitive Guide*, 2nd Edition. Sebastopol: O'Reilly Media.
- Cabral, João de Pina, Lima, Antónia Pedroso de. 2005. "Como fazer uma História de Família: Um Exercício de Contextualização Social", *Etnográfica, Revista do Centro de Estudos de Antropologia Social*, vol. IX, n.º 2.
- Campos, Maria do Rosário Castiço de. 2003. *Redes de sociabilidade e de poder : Lousã no século XVIII*, Coimbra.
- Carvalho, Joaquim Ramos de. 1989. "Expert Systems and Community Reconstruction Studies", In *History and Computing*, edited by Peter Denley, Stefan Fogelvik, 97-116, Manchester: Manchester University Press.
- Carvalho, Joaquim Ramos de. 2003. "MHK: um sistema de apoio a investigações em Micro-história", visitado 3 de Outubro de 2010, <http://timelink.fl.uc.pt/downloads/timelink.pdf>.
- Carvalho, Joaquim Ramos de. 2001. "A Comunidade como Horizonte de Inteligibilidade Máxima", Comunicação apresentada às Jornadas de História do Concelho de Gouveia.
- Carvalho, Joaquim Ramos de. 1997, *Comportamentos morais e estruturas sociais numa paróquia de Antigo Regime (Soure, 1680-1720): reconstituições, interpretações e metodologias*, Faculdade de Letras, Coimbra.

Clunies, Sandra MacLean. 2002. "Innovators: Bob Velke", *Genealogical Computing*, Vol. 22 n.º 1, <http://www.ancestry.com/learn/library/article.aspx?article=7608>, visitado a 9 de Setembro de 2009.

Cockburn, Alistair. 2001. *Writing Effective Use Cases*, Boston: Addison-Wesley Professional.

Coelho, Maria Helena da Cruz. 1997. "Os tabeliães em Portugal. Perfil profissional e sócio-económico (sécs. XIV-XV)", In *Estudios sobre el Notariado Europeo (Siglos XIV-XV)*, edited by Pilar Ostos, M<sup>a</sup> Luisa Pardo, Sevilla: Publicaciones de la Universidad de Sevilla.

Craigie, Maria João de Nogueira Ferrão Vieira. 2006. *Dicionário de Bibliografia para Genealogistas*, Lisboa: Dislivro Histórica.

Daupias D'Alcochete, Nuno. 1984. "Registo Paroquial", In *Dicionário de História de Portugal*, vol. V, Porto: Livraria Figueirinhas.

Dunn, Halbert L. 1946. "Record Linkage", *American Journal of Public Health*, vol. 36, n.º 12, 1412–1416.

Eastman, Dick. 2002. "Early Genealogy Programs, A Look Back at the Humble Beginnings of Genealogy Software", *Eastman's Online Genealogy Newsletter*, Vol. 7 N.º 37, 16 de Setembro de 2002, visitado em 24 de Abril de 2009. <http://www.eogn.com/archives/news0237.htm>.

Eastman, Dick. 2002. "GEDCOM Explained", *Eastman's Online Genealogy Newsletter*, Vol. 7 N.º 19, 13 de Maio de 2002, visitado em 24 de Abril de 2009, <http://www.eogn.com/archives/news0219.htm>.

Eastman, Dick. 2004. "GenBridge Update", *Eastman's Online Genealogy Newsletter*, 30 de Agosto de 2004, visitado em 24 de Abril de 2009, [http://blog.eogn.com/eastmans\\_online\\_genealogy/2004/08/genbridge\\_updat.html](http://blog.eogn.com/eastmans_online_genealogy/2004/08/genbridge_updat.html).

Eastman, Dick. 2006. "Wholly Genes Software's Genealogy Cruise is a Success", *Eastman's Online Genealogy Newsletter*, 19 de Novembro de 2006, visitado em 24 de Abril de 2009, [http://blog.eogn.com/eastmans\\_online\\_genealogy/2006/11/wholly\\_genes\\_so.html](http://blog.eogn.com/eastmans_online_genealogy/2006/11/wholly_genes_so.html).

Eldon, Eric. 2008. "2008 growth puts Facebook in better position to make Money", *VentureBeat* visitado em 24 de Abril de 2009, <http://venturebeat.com/2008/12/18/2008-growth-puts-facebook-in-better-position-to-make-money/>.

FamilyLink.com. 2008. "One Billion Names Now Online at WorldVitalRecords.com", visitado em 24 de Abril de 2009, <http://blog.worldvitalrecords.com/2008/03/11/one-billion-names-now-online-at-worldvitalrecordscom>

Forrest, Brandy 2006. "Controversy about our 'Web 2.0' service mark" visitado em 24 de Abril de 2009, <http://radar.oreilly.com/archives/2006/05/controversy-about-our-web-20-s.html>

Gaio, Felgueiras 1938-1941. *Nobiliário de famílias de Portugal* / Felgueiras Gaio, edit. por Agostinho de Azevedo Meirelles e Domingos de Araújo Affonso, Braga: Pax

- Greanier, Todd. 2004. *Java foundations*. San Francisco: Wiley.
- Haines, Steven e Poots, Stephen. 2002. *Java 2 primer plus*. Indianapolis: Sams Publishing.
- Gutherie, Julian. 2009. "Ancestry.com starts worldwide archiving Project", *Richmond Times-Dispatch*, visitado em 3 de Outubro de 2009  
[http://www.timesdispatch.com/rtd/lifestyles/health\\_med\\_fit/article/I-ARCH0403\\_20090416-185249/258306/](http://www.timesdispatch.com/rtd/lifestyles/health_med_fit/article/I-ARCH0403_20090416-185249/258306/).
- Harper, Simon e Yesilada, Yeliz. 2008. *Web accessibility: a foundation for research*. London: Springer
- Helm, Mathew L. e Helm, April Leigh. 2008. *Genealogy Online For Dummies, For Dummies*. Indianapolis, IN : Wiley Pub.
- Henry, Shawn Lawton. 2006. "Essential Components of Web Accessibility". W3C/WAI - World Wide Web Consortium / Web Accessibility Initiative. visitado em 6 de Setembro de 2010  
<http://www.w3.org/WAI/intro/components.php>.
- Henry, Shawn Lawton. 2006. "Understanding Web Accessibility", *Web Accessibility: Web Standards and Regulatory Compliance, 2006*, visitado em 6 de Setembro de 2010  
<http://www.uiaccess.com/understanding.html>.
- Henry, Shawn Lawton. 2007. *Just Ask: integrating accessibility throughout design*, EUA : Lulu.com.
- Holck, Jesper, Mahnke, Volker e Zicari, Roberto. 2010. Winning through incremental innovation: The case of MySQL AB, visitado em 6 de Setembro de 2010  
[http://www.iris31.se/best\\_papers/bestPaperJesper.pdf](http://www.iris31.se/best_papers/bestPaperJesper.pdf)
- Jansen, Wendy, Steenbakkens, Wilchard e Jägers, Hans. 2007. *New Business Models for the Knowledge Economy*, Aldershot : Ashgate.
- Kalakota, Ravi e Robinson, Marcia. 2000. *e-Business 2.0: Roadmap for Success*, Addison-Wesley.
- Karp, Scott. 2008. "Google AdWords: A Brief History Of Online Advertising Innovation", visitado em 2 de Dezembro de 2009, <http://publishing2.com/2008/05/27/google-adwords-a-brief-history-of-online-advertising-innovation/>
- Kelly, Sean U., Sung, Christopher e Farnham, Shelly. 2002. Designing for improved social responsibility, user participation and content in on-line communities, Conference on Human Factors in Computing Systems, ACM, New York.
- Kuo, Benjamin. 2008. "Interview with David Sacks, Geni and Yammer", socialTECH, visitado em 2 de Dezembro de 2009,  
[http://www.socaltech.com/interview\\_with\\_david\\_sacks\\_\\_geni\\_and\\_yammer/s-0017613.htm](http://www.socaltech.com/interview_with_david_sacks__geni_and_yammer/s-0017613.htm)
- Lackey, Richard S 1985. *Cite Your Sources: a manual for documenting Family Histories and Genealogical Records*, Jackson, Mississippi, University Press of Mississippi.

- Leffingwell, Dean e Widrig, Don. 1999. *Managing software requirements: a use case approach*, Addison-Wesley Professional.
- Liberge, Leila,. 2008. “Des portails pour la généalogie et l'histoire locale. L' exemple des Pays-Bas”, *Culture & Recherche*, n.º 118 -119, Automne-Hiver, 2008-2009, 15-16.
- Lytras, Miltiadis D., Damiani, Ernesto, Pablos, Patricia Ordonez. 2009. *Web 2.0, The Business Model*, New York: Springer
- Martins, Lígia de Azevedo. 1994. *Guia preliminar dos fundos de arquivo da Biblioteca Nacional*, Lisboa: Instituto da Biblioteca Nacional e do Livro.
- Mayfield, Antony. 2008. What is social media? visitado em 5 de Abril de 2009. [http://www.icrossing.co.uk/fileadmin/uploads/eBooks/What\\_is\\_Social\\_Media\\_iCrossing\\_ebook.pdf](http://www.icrossing.co.uk/fileadmin/uploads/eBooks/What_is_Social_Media_iCrossing_ebook.pdf).
- Mills, Elizabeth Shown. 1997. *Evidence! Citation & Analysis for the Family Historian*, Baltimore: Genealogical Publishing.
- Nurse, Howard L. 1996. “The ROOTS Story: Software to Make History”, visitado em 5 de Abril de 2009, <http://sonic.net/~commsoft/rstory.html>.
- O'Reilly, Tim. 2005. “What Is Web 2.0. Design Patterns and Business Models for the Next Generation of Software”, visitado em 11 de Fevereiro de 2009 <http://oreilly.com/lpt/a/6228>.
- Pizarro, José Augusto P. Sotto Mayor. 1982. “Duas ciências auxiliares da História, a Genealogia e a Heráldica”, *Humanidades*, N.º 2.
- Público. 2007. “Igreja Mórmon disponibiliza registos genealógicos até ao século XVI”. visitado a 18 de Junho de 2009, <http://ultimahora.publico.clix.pt/noticia.aspx?id=1314843&canal=62>.
- Reigel, Terry. 2005. “Why TMG?. Terry's TMG Tips”, visitado em 11 de Fevereiro de 2009, <http://tmg.reigelridge.com/whyTMG.htm>.
- Ribeiro, Ana Isabel Sampaio Ribeiro. 2005. *A comunidade de Eiras nos finais do século XVIII. Estruturas, redes e dinâmicas sociais*. Prefácio de Margarida Sobral Neto e Joaquim Ramos de Carvalho, Coimbra, Faculdade de Letras.
- Rivlin, Gary. 2006. “Wallflower at the Web Party”, *New York Times*, October 15.
- ROSEN, Ellen. 2007. “Latest Genealogy Tools Create a Need to Know”, visitado em 11 de Fevereiro de 2009, [www.nytimes.com/2007/08/18/business/18genealogy.html](http://www.nytimes.com/2007/08/18/business/18genealogy.html).
- Rosenfeld, Louis e Morville, Peter. 2006. *Information Architecture for the World Wide Web*, 3 Ed., Sebastopol: O'Reilly Media.
- Seaver , Randy. 2009. “A Tale of Two Genealogy Companies”, visitado em 11 de Fevereiro de 2009, <http://www.geneamusings.com/2009/01/tale-of-two-genealogy-companies.html>.

Thatcher, Jim, Burks , Michael R., Christian Heilmann. 2006. *Web accessibility: web standards and regulatory compliance*, New York: Springer.

Vickery, Graham, Wunsch-Vincent, Sacha. 2007. *Participative web and user-created content*, Paris: Organisation for Economic Co-operation and Development.

Wubs, H. e Huysmans, F. 2006. “Snuffelen en Graven”, Den Haag, Sociaal en Cultureel Planbureau, visitado em 13 de Outubro de 2010,  
<http://www.scp.nl/dsresource?objectid=20623&type=org>

## Anexo I. Estatísticas de acesso a *timelink.fl.uc.pt*

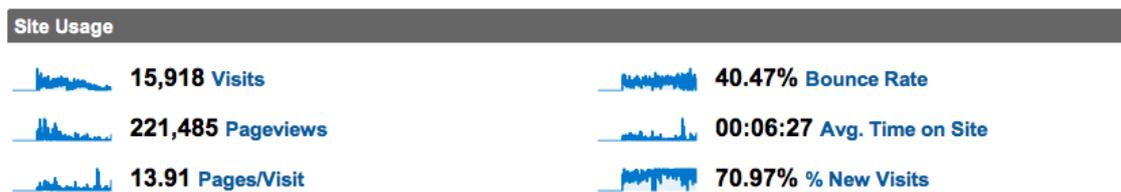


Figura 26. Total de visitas e visualizações de páginas, entre Março de 2009 e Outubro de 2010 (Google Analytics)

Posição	País de origem	Total de visitas
1.	Portugal	9,975
2.	Brasil	2,782
3.	Estados Unidos da América	718
4.	Suíça	250
5.	Bélgica	210
6.	França	187
7.	Espanha	181
8.	Canadá	178
9.	Reino Unido	168
10.	México	159
11.	Itália	105

Tabela 9. Número de visitas a *timelink.fl.uc.pt* por país entre Março de 2009 e Outubro de 2010 (Google Analytics)



Figura 27. Distribuição espacial das visitas (Google Analytics).

		Visits ↓	Pages/Visit	Avg. Time on Site	% New Visits	Bounce Rate
1.	timelink	296	31.55	00:15:22	36.15%	27.03%
2.	time link	120	38.45	00:14:16	36.67%	34.17%
3.	ofícios nocturnos	82	1.15	00:00:12	98.78%	93.90%
4.	timelink obidos	59	11.24	00:05:47	71.19%	5.08%
5.	"brígida monteira"	52	62.85	00:34:08	21.15%	5.77%
6.	timelink.fl.uc.pt	50	67.40	00:31:12	14.00%	12.00%
7.	http://timelink.fl.uc.pt/mhk/login/	40	18.95	00:18:21	2.50%	12.50%
8.	mhk	39	53.38	00:18:49	33.33%	23.08%
9.	brígida monteira	33	61.24	00:35:57	30.30%	0.00%
10.	ordens de epistola	30	106.10	00:56:32	0.00%	10.00%

Tabela 10. Expressões pesquisadas em motores de pesquisa (Google Analytics).