

# Museu da Pessoa — Arquitectura

J. João Almeida

jj@di.uminho.pt

J. Gustavo Rocha

jgr@di.uminho.pt

P. Rangel Henriques

prh@di.uminho.pt

Sónia Moreira

sonya@alfarrabio.di.uminho.pt

Alberto Simões

albie@alfarrabio.di.uminho.pt

## Resumo

É conhecido o extremo valor patrimonial e sociológico das Histórias de Vida. No entanto, apesar do encanto irresistível das histórias, trata-se de um património em rápido desaparecimento.

O Museu da Pessoa é um arquivo fundamentalmente acessível pela Internet, que apresenta um conjunto de especificidades:

- arquivo electrónico
- documentos multi-formato e contendo documentos
- complexidade do ciclo de vida de aquisição das histórias
- meta-informação, documentos, estruturas classificativas, postos de construção e consulta, tudo no mesmo suporte
- trabalho cooperativo multi-instituição, internacional

Neste documento descreve-se o projecto e apresenta-se a arquitectura do **Museu da Pessoa** discutindo-se a problemática associada à sua concretização.

## 1 Introdução

O Museu da Pessoa tem como objectivo colectar depoimentos de todo o ser humano, célebre ou anónimo, para eternizar a sua história. A partir da história do indivíduo, escrevem-se as histórias das famílias, das comunidades, das instituições e dos povos. O Museu da Pessoa interessa-se pelo "eu". Trata o homem comum não como um objecto experimental mas como um informador, bem informado sobre os acontecimentos e experiências que vivenciou. Em cada história de vida, para além do perfil pessoal, há um conjunto de representações que a pessoa faz de fenómenos sociais. Ao reportarem as suas vidas, os narradores evocam eventos e outros temas, daí que as histórias de vida podem servir como elemento essencial de pesquisa para identificação de algumas problemáticas. Através dos relatos das histórias de vida pode-se reconstruir um universo social.

As histórias de vida não constituem um inquérito verificatório, não visam estabelecer leis ou sequer comprovar hipóteses, têm por função recolher testemunhos, elucidá-los e descrever acontecimentos vividos. São testemunhos pessoais, que transportam um carácter social e histórico, que devem ser preservados e trabalhados para se tornarem num património humano incomensurável.

A partir do registo de depoimentos em áudio e/ou vídeo e de pesquisa directa em arquivos, o Museu da Pessoa: disponibiliza dados na Internet; organiza e sistematiza acervos históricos; produz bases de dados, CD-ROM, publicações, documentários em vídeo, exposições e seminários. A pesquisa histórica baseada em depoimentos permite ainda aplicações como: marketing institucional e promocional; valorização de recursos humanos; actividades pedagógicas, culturais e entretenimento.

## 1.1 História do Núcleo Português do Museu da Pessoa

A ideia de um espaço virtual que congregasse um infindável número de histórias de vida contadas na primeira pessoa surgiu no Brasil, em São Paulo, por um grupo de jornalistas e historiadores, em 1992. O Museu da Pessoa brasileiro é actualmente uma fonte de informação única, com um acervo de 500 depoimentos e mais de mil fotografias digitalizadas e várias publicações. As televisões brasileiras, em particular a Rede Globo, têm recorrido a este material para criar personagens para telenovelas. Cativados por este projecto tão criativo e humano, um grupo de docentes da Universidade do Minho criaram, em 1999, o Núcleo Português do Museu da Pessoa. O primeiro trabalho que está a ser realizado em parceria com o Museu de São Paulo, é a recolha de histórias de personagens anónimas do Porto, de modo a reconstruir a história de certos lugares e preservar um património social de tradições ricas e únicas, que tendem a ser esquecidas pelos mais novos. O promotor deste primeiro projecto é a sociedade Porto 2001. Depois foram surgindo outras iniciativas e parcerias com a Biblioteca Camilo Castelo Branco de Famalicão, com o projecto Nónio e com a Câmara Municipal de Montalegre.

O Núcleo quer-se manter e solidificar com novos projectos e iniciativas, mas tem como objectivo último a criação, juntamente com o núcleo brasileiro, e outros que poderão surgir, de uma rede de Museus da Pessoa que valorize e divulgue a língua lusófona a nível mundial.

## 1.2 Declaração de princípios e filosofia

O conceito de Museu da Pessoa está muito ligado a um conjunto de convicções e objectivos que sumariamente podemos resumir em:

- acreditamos que é urgente recolher as histórias individuais e pequenos episódios;
- julgamos que devemos manter, sempre que possível, uma história contada na primeira pessoa, com toda a sua riqueza adicional;
- queremos tornar o acervo o mais acessível possível (Internet, réplicas, livros, etc.), uma vez que se trata de um património comum;
- queremos ver respeitadas as regras da ética;
- queremos que seja usado o máximo rigor e que sejam seguidas as abordagens mais metódicas que for possível;
- sabemos que o conjunto de potenciais usos de cada história é quase infinito.

Tentaremos sempre que possível:

- ir definindo e usando formatos comuns, de modo a facilitar intercâmbio de histórias e de ferramentas. Independência de plataformas. Independência de aplicações;
- construir e partilhar ferramentas que facilitem o trabalho;
- construir estruturas classificativas comuns.

Para que a riqueza seja máxima pretendemos também que haja o maior número possível de ligações entre documentos.

### Postura adoptada

Como consequência, no núcleo têm-se a adoptada uma postura

- aberta
- faça você mesmo

- disponibilização dos documentos em formatos ricos (XML) que tenham capacidade de estabelecer relações entre os documentos
- apoio a iniciativas que registem histórias
- tentativa de rigor (formação, equipe multi-disciplinar)

### 1.3 Ciclo de tratamento de uma história de vida

O Tratamento da história tem um ciclo de vida, uma série de etapas de produção até se tornar num texto classificado, resumido, condensado e cronologicamente ordenado. A recolha da história de vida realiza-se sempre a partir da entrevista, pedindo ao entrevistado que se conte, que descreva a sua história pessoal. É preciso deixar o entrevistado contar-se ao seu próprio ritmo, podendo alongar-se e fazer as pausas necessárias. No entanto, o entrevistador nunca se deve apagar no seu papel, deve sempre orientar, mudar de tema consoante o guião pré-estabelecido para a entrevista.

A etapa seguinte é a transcrição da entrevista em formato XML, uma linguagem que nos permite fazer anotações de texto, que corresponde à reprodução fiel do discurso registado. As pausas, os risos, as repetições, os ritmos, os suspiros, os assobios e as variações morfológicas devem ser cuidadosamente anotadas. A entrevista em bruto deve preservar todas as redundâncias do discurso.

O passo que se segue é a revisão da entrevista transcrita, de modo a apreender as repetições e as ideias-chaves. Destina-se a eliminar as redundâncias do discurso e as repetições, a fazer as devidas correcções ortográficas e as variações sintácticas, a reagrupar partes do discurso que se referem ao mesmo tema mas que se encontram desordenadas na história de vida, a acrescentar meta-informação ao documento, a marcar as referências, as entidades, os locais e as pessoas. Atribuir uma certa continuidade e linearidade à entrevista, para que não se verifiquem repetições temáticas ao longo do texto e, ao mesmo tempo, enriquecê-la com informação adicional. Surge a entrevista editada, reagrupada por temas e já com a catalogação de uma ou outra história, ou frases que merecem destaque e que possam cativar o leitor a ler a entrevista na íntegra. A acompanhar a entrevista editada há a mini-biografia, um pequeno texto que retrata, em forma de síntese, o percurso da vida do narrador, referenciando os momentos mais pertinentes e fomentando a curiosidade do leitor para uma pesquisa.

Concluídas estas etapas, o depoimento está pronto para ser indexado. A indexação tem como primeiro objectivo fornecer de modo condensado uma representação simplificada da entrevista. Apreendem-se as respostas, eliminam-se as perguntas e determinam-se os indexadores que vão permitir juntar um conjunto de histórias num discurso único, ordenado cronologicamente, de modo a conservar as variações pessoais. Esta estrutura torna a história de vida mais legível e compreensível para o leitor. A fase de indexação, é das mais complicadas de todo o ciclo de vida das histórias, porque obriga a uma classificação de acontecimentos e a dificuldade reside em saber como classificar sem deformar a narrativa, sem deixar de fora elementos significantes. É uma problemática que cabe ao indexador resolver de modo ponderado, sem deturpar o sentido e o fundamento da história.

Paralelamente ao tratamento da história, há o tratamento do som, da imagem e dos documentos pessoais. O som da entrevista é normalmente guardado em minidisc, posteriormente é convertido para MP3, onde será indexado, marcado e colocado no site, para que haja a possibilidade de o leitor ler e ouvir a história em simultâneo. Esta é uma forma de preservar a oralidade, as terminologias e os vocábulos portugueses, tão característicos de cada região, que vão desaparecendo. A imagem é digitalizada e tratada, ajustam-se as cores, eliminam-se as sombras e recuperam-se zonas da fotografia gastas pelo tempo. Seguidamente procede-se à legendagem; depois é altura de localizar os lugares, referir as datas e identificar as pessoas. Tratada e legendada a fotografia pode, não só, ser inserida na entrevista editada para ilustrar a história, como também, pode ser utilizada na criação de um álbum fotográfico pessoal.

O ciclo de tratamento da história não termina no momento em que acaba o tratamento; os derivados que pode originar são infindáveis. A conversão da história de XML para HTML e

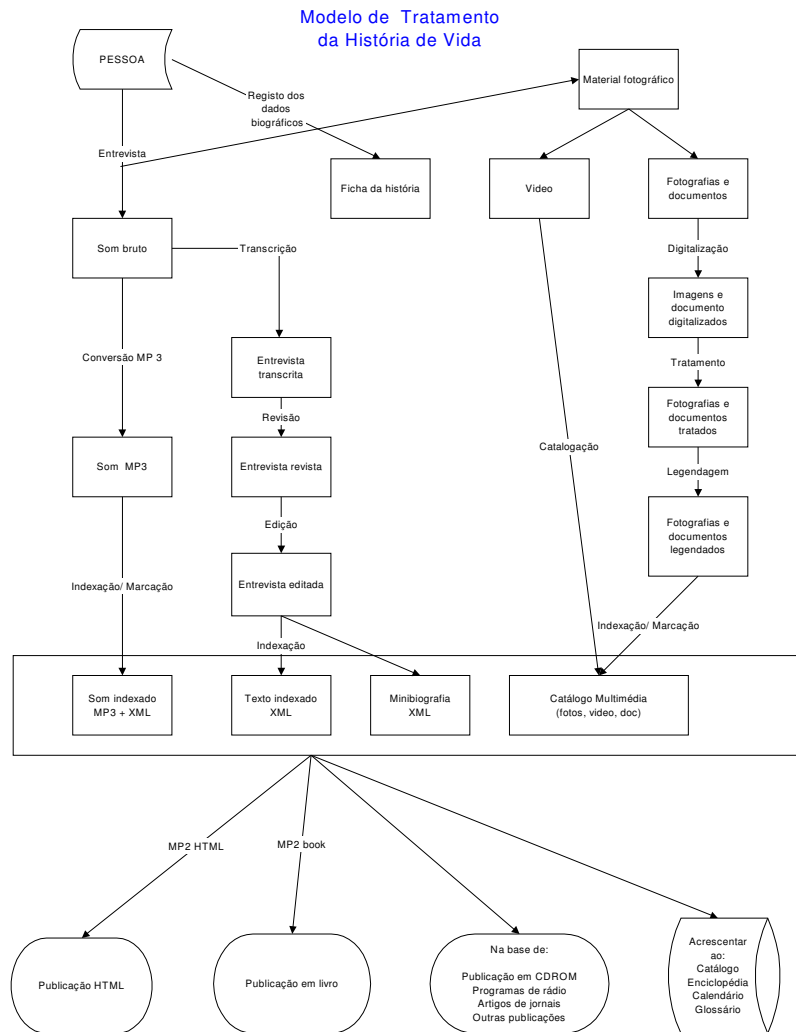


Figura 1: Ciclo de vida de uma entrevista

posteriormente a sua publicação na Internet é apenas um dos vários destinos do depoimento. Há sempre a possibilidade da criação de CD ROM, da feitura de livros, da participação em programas de rádio, da apresentação em jornais, entre várias.

O Museu da Pessoa não se restringe só a trabalhar histórias de vida recolhidas no terreno; todas as pessoas podem enviar e registar na Internet o seu testemunho, as suas histórias.

Nestes casos, em que o narrador escreve sem ser interpelado por um entrevistador, trata-se de um tipo de documento bastante diferente. O ciclo de vida destes documentos enviados é diferente embora haja um conjunto de fases comuns.

## 2 Concretização do Museu da Pessoa

### 2.1 Armazém vs arquivo

O conjunto dos depoimentos recolhidos constitui o acervo: algo semelhante a um armazém de documentos. Para que se constitua um arquivo é necessário construir um conjunto de componentes extra que permitam:

- a navegação conceptual
- o acesso ao catálogo de documentos existentes
- uma visão da evolução temporal do arquivo
- uma interligação entre o que for relacionado

Seguidamente analisaremos a organização das ferramentas e componentes que suportem o arquivo.

### 2.2 Arquitectura do MP

O Museu da Pessoa é composto por um conjunto de projectos, um sistemas classificativo comum e um conjunto de recursos de ligação (que se explicarão à frente):

MP = Projecto*	x
Sistema Classificativo	x
Enciclopédia	x
Glossário	x
Calendário	x
Boletim	

A Estrutura classificativa é uma ontologia com informação geográfica, de eventos, de profissões, temática, etc. Está descrita através de um thesaurus enriquecido com uma maior variedade de relações e de um conjunto de propriedades matemáticas.

Os Recursos gerais de ligação, são um conjunto de suportes onde se colecta informação das histórias individuais juntando-as para formarem uma colecção de conhecimentos agrupados de determinada maneira:

- Enciclopédia - onde se agrupam referências a eventos, personalidades, monumentos, etc
- Glossário - onde se juntam termos menos usuais e sua explicação
- Calendário - onde se vão colocando acontecimentos que tenham uma ligação temporal forte
- Boletim: destaques, novidades, agenda

Cada Projecto é composto por documentos e subprojectos mantendo ainda um conjunto de informação geral associada:

```

Projecto =
  identificação                x
  Descrição                    x
  Publicações                  x
  Responsável: pessoa          x
  Colaboradores: pessoa*      x
  Classificação: (termo x relacao)* x
  ( Projecto | Documento )*

```

Os documentos são talvez a parte mais importante dos projectos. De entre os possíveis documentos damos especial realce às histórias de vida que agrupam um conjunto de subhistórias, som, fotografias e informação de catalogação e classificação associada:

```

Historia =
  meta-informação              x
  (texto x mp3 | História | foto ) x
  relevância

meta-informação =
  classificação: (termo x relacao)* x
  ( papel x pessoa )*          x
  data

```

## 2.3 Ferramentas

No sentido de apoiar as actividades referidas nas secções anteriores tem-se vindo a desenvolver um conjunto de ferramentas (que se encontram em diferentes graus de maturidade e acabamento):

- ajuda à transcrição/estruturação:
  - detector de nomes próprios (dado um texto permite acrescenta etiquetas de referência em volta dos nomes próprios)
  - editor de texto com controlo via teclas de um tocador de MP3
- associadas à estrutura classificativa:
  - extracção de todos os termos usados em classificação de documentos para possível revisão da estrutura classificativa
  - módulo de gestão de ontologias e thesaurus (para permitir uma pesquisa e navegação conceptual nas histórias e documentos em geral)
- associadas à publicação:
  - conversor XML -> HTML multi-página (ver descrição e exemplos na secção seguinte)
  - conversor XML -> livro (ver descrição e exemplos na secção seguinte)
- associada à gestão de recursos comuns e de ligação:
  - extracção de datas (extraí datas de nascimento e de eventos em geral, para enriquecimento do calendário)
  - construtor de catálogos (extraí a metainformação associada às várias histórias de vida e suas sub-histórias)
  - módulo construtor/editor de novidade (de modo a permitir ter uma leitura da evolução temporal do arquivo. Exemplo de novidades são a descrição (metainformação) dos últimos documentos inseridos no arquivo )
  - construtor de glossários
  - construtor de agenda/calendário

## 3 Exemplo: uma história de vida e seus derivados

### 3.1 Anotação de uma história de vida

Vamos analisar cuidadosamente qual o tratamento que foi dado a uma história de vida e quais as suas consequências. Em particular, este é um extracto da entrevista do marinheiro José Teixeira, conhecido por Naná.

Após a definição de início de um documento do museu da pessoa (mp), e em particular o de uma história de vida (hv), introduzimos meta-informação (meta) referente à entrevista:

```
<mp>
<hv>
<meta>
  <projecto>Projecto Comércio Tradicional da Ribeira</projecto>
  <autor tipo="Transcrito por">Sónia Moreira</autor>
  <autor tipo="Entrevistadores">
    Patrícia Sousa, Lígia Costa e Sónia Moreira</autor>
  <autor tipo="Depoente">José Fernandes Teixeira (Naná)</autor>
  <nascimento onde="Muro dos Bacalhoeiros - Ribeira"
    data="1946/12/22"></nascimento>
  <profissao> Marinheiro ; Construtor de barcos</profissao>
  <data>Ribeira, Dezembro de 1999</data>
</meta>
```

Deste pequeno extracto podemos retirar informação variada. A quarta linha permite retirar o nome do projecto onde este documento se enquadra. Ou seja, ao encontrarmos um documento solto facilmente podemos identificar a sua origem.

As quatro linhas seguintes definem vários autores. No entanto, devemos ter em consideração os vários tipos de autores. Os primeiros dois conjuntos definem quem transcreveu o documento (a partir do som em bruto produziu a entrevista em bruto) e quem entrevistou o indivíduo sobre o qual trata o depoimento. Por fim, o último autor é o depoente, sobre quem trata a história de vida. Poderíamos definir outros nomes distintos para estas etiquetas, mas quisemos frisar que se trata, realmente, de um autor.

A etiqueta da nona linha permite indicar em que dia e onde nasceu o depoente. Esta informação pode ser extraída automaticamente para ser incorporada no calendário onde pode ser indicada a sua data de aniversário.

O local onde o depoente nasceu pode ser extraído para o thesaurus (permitindo posteriormente a navegação geográfica) e pode ser adicionado à enciclopédia que o deve caracterizar geograficamente.

Em relação à profissão, deve ser tratada da forma que foi referida no parágrafo anterior: incorporada no thesaurus para possibilitar uma futura pesquisa temática e incorporada na enciclopédia para a caracterizar.

Finalmente, a data referente à entrevista permite que se saiba quando e onde foi realizada.

Depois da meta-informação, pode aparecer um parágrafo como:

```
<identificacao>Sou José Fernandes Marques Teixeira.
Nasci no Muro dos Bacalhoeiros nº111 a 22 de Dezembro de 1946.
Actualmente, moro na Rua da Fonte Taurina.
</identificacao>
```

A etiquetação de um parágrafo desta forma permite que o leitor saiba que aí encontra toda a informação referente à identificação do depoente. Ao converter este documento para um outro formato devemos ter em consideração esta zona para a destacarmos de alguma forma.

Um parágrafo seguinte pode ser:

```
<infancia> A minha infância ..... a andar descalço pelas ruas, que era
proibido.

<h tipo="multas; desinfeção"
  tit="'Deitavam-nos creolina na cabeça" >
```

Se fôssemos apanhados descalços a polícia prendia-nos e éramos multados. A multa era à volta dos 12 ou 13 escudos. Para não nos multarem, nós fazíamos ligaduras para os pés, mas o meu pai às vezes arranjava socas em madeira. Também não podíamos andar com o cabelo grande porque havia a desinfecção. Éramos presos, rapavam-nos o cabelo todo e deitavam-nos creolina na cabeça para matar os piolhos.

</h>  
</infancia>

Neste extracto podemos encontrar dois pormenores importantes. Por um lado, ao anotarmos um conjunto de texto como descrevendo a infância do depoente, podemos dar-lhe algum destaque e classificá-lo de forma adequada. Permite que um leitor que esteja interessado nas histórias das crianças naquele tempo as encontre facilmente.

Em relação à etiqueta denominada h, de história, anota determinado texto como algo que se pode ler independentemente do depoimento completo, no fundo uma sub-história contida na história de vida. Esta anotação deve conter um título, identificativo da história, e um tipo de história. Este tipo é extraído automaticamente para ser introduzido no thesaurus.

<h tipo="Episódio" tit="O Dione que ficou sem leme">  
E numa dessas viagens aos Açores, no navio Dione, à vinda para Lisboa o barco largou o leme. A gente vinha a governar em cima e o leme andava sempre à roda em baixo. Tentávamos guiar, mas não conseguíamos governar o barco porque não tinha o leme debaixo. Andámos oito dias à deriva em alto-mar a apanhar vagas muito altas. Nenhum salva-vidas nos foi buscar, foi um arrastão de pesca que nos deu reboque até aos estaleiros. Tivemos o cozinheiro que ficou louco, começou a ver coisas e queria matar-me. Houve uma altura, em que pensámos matar o cão para comer, porque não tínhamos nada para comer. O cão, infelizmente, também caiu pela borda fora e foi para o mar. Há coisas que naquele tempo passavam-se com maior dificuldade. Agora, os barcos andam sozinhos, não precisam de ninguém a guiá-los.  
</h>

<h tipo="naufrágio" tit="O naufrágio da Badalhoc">  
<destaque>Eram barcos de passageiros mesmo, havia muitos que traziam para baixo e levavam para cima, por exemplo, mais tarde fizeram umas lanchas, era a <ref tipo="barco">Foz de Sousa</ref>, mais conhecida pelo nome de "Badalhoc". A "Badalhoc" foi ao fundo em 5 de Abril de 1950 mesmo em frente à Quinta da Paradela. Bateu nas pedras, morreram 28 pessoas, foi o desastre maior que houve no rio Douro até á data. Eu era miúdo tinha para aí 12 ou 13 anos. Ela ia para cima carregada de pessoal, bateu numa pedra no fundo encostada à margem. Bateu, arrombou e começou a meter água e o pessoal caiu todo para o lado, o barco virou e as pessoas morreram.</destaque>  
</h>

</hv>  
</mp>

Este extracto demonstra novamente a etiquetação de histórias. Em relação à etiqueta *destaque*, permite que se especifique algum texto a ser incluído de forma especial nas publicações. Por exemplo, pode ser colocado com letras um pouco maiores numa das margens duma publicação. Em relação à etiquetagem de *Foz de Sousa* como referência ao termo *barco* permite que seja introduzido no catálogo para que a navegação do thesaurus os apresente.



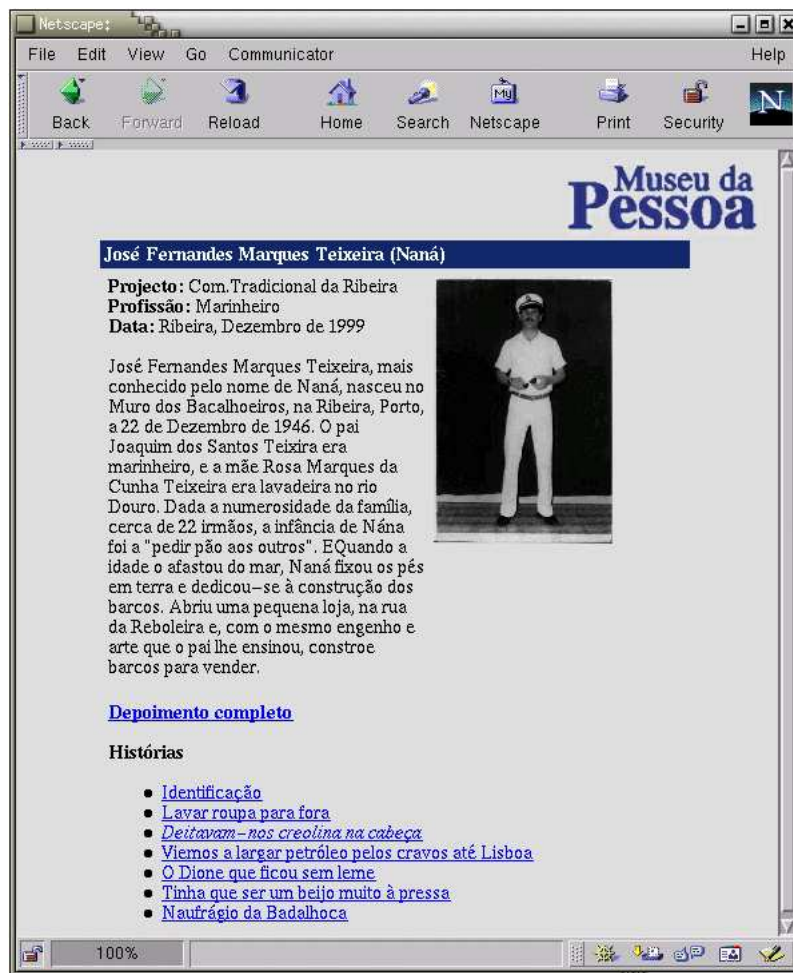


Figura 2: Exemplo de uma página html índice

### 3.2 Publicação para a Internet

Para a apresentação do depoimento na Internet é necessária a transformação da linguagem de anotação em que está etiquetado (XML) para a utilizada na Internet (HTML).

Consideremos a entrevista editada que é constituída, basicamente, por meta informação referente ao depoente e um texto anotado. Este texto pode estar dividido em histórias. Uma forma possível seria publicar a entrevista toda de uma vez. No entanto, qualquer leitor se cansará passadas algumas linhas, ou porque lhe parece que o texto é demasiado grande, ou porque não é o tipo de história com que estava a contar.

Assim, uma forma mais correcta de publicar a entrevista é dividir a mesma em várias partes. Para apresentar a história deve existir uma página que possa ser como que o bilhete de identidade do depoente. Esta deve conter toda a informação que se sabe sobre ele, uma fotografia (se a houver) e uma listagem dos títulos das histórias existentes.. Na informação sobre o depoente deve-se aumentar uma pequena biografia para que o leitor possa saber sobre que trata a entrevista.

Esta página pode conter referências a outro tipo de catalogação como a que vimos no documento: *infância*. De forma idêntica, podem existir etiquetas idênticas: *casamento*, *ofício*, ... Cada um destes textos também devem ser incluídos em páginas: uma para cada um.

Cada uma das histórias contadas deve estar numa página separada para que se possa aceder directamente à história desejada. Essa página deve conter o título da história, um tipo classificativo

e o autor da mesma.

Por fim, continua a ser importante uma folha com a entrevista completa para que seja possível ler as partes da entrevista que não foram incluídas nas sub-histórias.

As referências devem permitir, utilizando os *links* do HTML, aceder aos documentos do catálogo desse mesmo tipo.

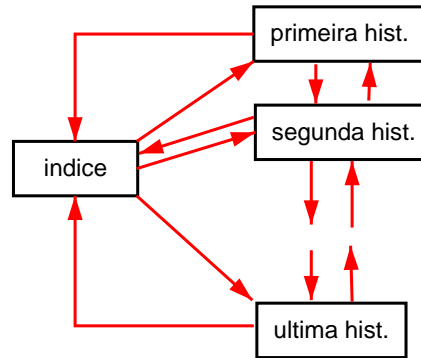


Figura 3: Esquema de interligação das sub-histórias

### 3.3 Publicação em Papel

Frequentemente as histórias podem ter tamanho correspondente a várias páginas de texto impresso. Por vezes constituem autênticos livros. Dado que não é cómodo ler textos grandes em écran de computador, decidiu-se que era importante a existência de ferramentas de publicação em papel para uso interno e externo.

As ferramentas de publicação em papel actualmente existentes no Museu da Pessoa permitem construir "livros" em vários formatos (a4, a5, a6, agenda) com histórias individuais, ou com grupos de histórias, totalmente automaticamente.

Internamente está a ser produzido formato Latex que é sucessivamente transformado para PostScript e finalmente também para PDF (portable document format).

Optou-se também por permitir a impressão directa de cadernos já paginados de modo a que quando impressos numa impressora dúplex, seja apenas necessário agrafar para que o "livro" esteja pronto.

As seguintes mostra-se o resultado de imprimir o exemplo atrás referido em formato de agenda.

## José Fernandes Teixeira (Naná)

Museu da Pessoa  
Projecto Comércio Tradicional da Ribeira  
15 de Fevereiro de 2001

### Conteúdo

1 José Fernandes Teixeira (Naná)	
Marinheiro ; Construtor de barcos	2
1.1 Identificação	3
1.2 Deitavam-nos creolina na cabeça	3
1.3 O Dione que ficou sem leme	3
1.4 Naufrágio da Badalhoca	3

1

## 1 José Fernandes Teixeira (Naná)

Marinheiro ; Construtor de barcos

Projecto Comércio Tradicional da Ribeira

Transcrito por - Sónia Moreira

Entrevistadores - Patrícia Sousa, e Sónia Moreira

Deponente - José Fernandes Teixeira (Naná)

Nascido em Muro dos Bacalhoiros - Ribeira 1946/12/22

...Eram barcos de passageiros mesmo, havia muitos que traziam para baixo e levavam para cima, por exemplo, mais tarde fizeram umas lanchas, era a **Foz de Sousa**, mais conhecida pela nome de "Badalhoca". A "Badalhoca" foi ao fundo em 5 de Abril de 1950 mesmo em frente à Quinta da Paradeira. Bateu nas pedras, morreram 28 pessoas, foi o desastre maior que houve no rio Douro até à data. Eu era miúdo tinha para aí 12 ou 13 anos. Ela ia para cima carregada de pessoal, bateu numa pedra no fundo encostada à margem. Bateu, arrombou e começou a meter água e o pessoal caiu todo para o lado, o barco virou e as pessoas morreram....

2

### 1.1 Identificação

Sou José Fernandes Marques Teixeira. Nasci no Muro dos Bacalhoiros nº111 a 22 de Dezembro de 1946. Actualmente, moro na Rua da Fonte Taurina.

A minha infância ..... a andar descalço pelas ruas, que era proibido.

### 1.2 Deitavam-nos creolina na cabeça

*multas, desinfeção*

Se fôssemos apanhados descalços a polícia prendia-nos e éramos multados. A multa era à volta dos 12 ou 13 escudos. Para não nos multarem, nós fazíamos ligaduras para os pés, mas o meu pai às vezes arranjava socas em madeira. Também não podíamos andar com o cabelo grande porque havia a desinfeção. Éramos presos, rapavám-nos o cabelo todo e deitavam-nos creolina na cabeça para matar os piolhos.

### 1.3 O Dione que ficou sem leme

*Epidémio*

E numa dessas viagens aos Açores, no navio Dione, à vinda para Lisboa o barco largou o leme. A gente vinha a governar em cima e o leme andava sempre à roda em baixo. Tentávamos guiar, mas não conseguíamos governar o barco porque não tinha o leme debaixo. Andámos oito dias à deriva em alto-mar a apanhar vagas muito altas. Nenhum salva-vidas nos foi buscar, foi um arrastão de pesca que nos deu reboque até aos estaleiros. Tivemos o cozinheiro que ficou louco, começou a ver coisas e queria matar-me. Houve uma altura, em que pensámos matar o cão para comer, porque não tínhamos nada para comer. O cão, infelizmente, também caiu pela borda fora e foi para o mar. Há coisas que naquele tempo passavam-se com maior dificuldade. Agora, os barcos andam sozinhos, não precisam de ninguém a guiá-los.

### 1.4 Naufrágio da Badalhoca

*naufrágio*

Eram barcos de passageiros mesmo, havia muitos que traziam para baixo e levavam para cima, por exemplo, mais tarde fizeram umas lanchas, era a **Foz de Sousa**, mais conhecida pela nome de "Badalhoca". A "Badalhoca" foi ao fundo em 5 de Abril de 1950 mesmo em frente à Quinta da Paradeira. Bateu nas pedras, morreram 28 pessoas, foi o desastre maior que houve no rio Douro até à data. Eu era miúdo tinha para aí

3

Museu da  
**Pessoa**

Dezembro 2001

D	S	T	Q	Q	S	S
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31					

Acontecimentos:

dia 17: [Abertura da Taberna](#)  
A 17 de Dezembro de 1961 o Sr. Manuel de Oliveira Chaminé abriu a taberna "Neca Chaminé"

dia 1: [Aniversário](#)  
A 1 de Dezembro de 1931 nasceu o Sr. António Machado

dia 22: [Aniversário do Naná](#)  
A 22 de Dezembro de 1946 nasceu o Senhor José Fernando Marques Teixeira, mais conhecido por Naná.

Dado que o texto inicial estava marcado conceptualmente (em XML), foi possível escolher

quais os elementos que se pretende que apareçam no livro final, foi possível gerar índices, escolher títulos, copiar destaques para páginas iniciais, realçar referências, etc.

### 3.4 Extracção de informação

Do mesmo modo que se extraiu e transformou os textos etiquetados para HTML e para PDF, também se construíram extractores de meta-informação para os catálogos, de entradas para o calendário, de termos para o glossário, etc.

No que diz respeito à criação de catálogo, é retirado tanto a meta-informação geral da história de vida como as meta-informações referentes a cada sub-história. Ou seja: as sub-histórias comportam-se como documentos autónomos permitindo assim uma granularidade mais fina.

Tanto as Histórias de Vida como as sub-histórias contêm informação classificativa – classificadores (retirada por exemplo do atributo `tipo` das histórias, e dos elementos `profissão` da meta-informação) que para além de constarem do catálogo, são directamente usados na navegação conceptual, juntamente com o thesaurus.

No exemplo anterior aparece explícita a referência à data de nascimento. Torna-se portanto possível a extracção de pequenos registos de destaque que se acrescenta ao calendário:

```
<evento id="21" data="22-12-*/>
<titulo>Aniversário de José Fernandes Teixeira</titulo>
<url>/mp/art2html?file=Nana.xml</url>
<desc>A 22 de Dezembro de 1946 nasceu
    José Fernando Teixeira (Naná).</desc>
</evento>
```

Este destaque fica ainda ligado através do campo `<url>` à história de vida ou à sub-história correspondente.

De um modo semelhante, quando existe alguma ocorrência de termos pouco usuais, estes podem ser marcados com a sua explicação permitindo que seja possível extrair um glossário de termos invulgares, consultável separadamente e com ligações às histórias.

## 4 Conclusões e trabalho futuro

Na vida atarefada dos nossos dias, não nos resta tempo para nos questionarmos sobre as nossas origens. Quem somos e de onde vimos? São perguntas às quais muitos de nós não sabem responder. Parar um pouco para escutar os nossos avós e os nossos pais, dar atenção àqueles que fizeram a nossa história, que têm memórias ricas e únicas para partilhar, não custa muito e enriquece-nos por dentro.

Com este projecto pretendemos:

- preservar todos os episódios e pequenas histórias que se vão apagando das memórias ou morrendo com aqueles que as protagonizaram,
- motivar o diálogo entre gerações e valorizar a história de cada pessoa, independentemente do seu estatuto social ou da sua raça,
- aumentar o nosso acervo de histórias de forma a permitir analisar de várias perspectivas os antepassados da nossa sociedade.

A recolha de histórias permite não só caracterizar as populações de várias zonas, mas também instituições que desejem reconstruir a sua história através das vidas dos seus trabalhadores.

## 4.1 Evolução e estado actual

Como se mostrou ao longo deste artigo, para transformar o armazém que se vai constituindo ao recolher todos esses depoimentos, sons e imagens, num verdadeiro arquivo digital aonde as histórias pudessem ser consultadas de maneira fácil, por si sós ou relacionadas com outras histórias, pessoas, locais, eventos, épocas, etc., tivemos necessidade de desenvolver uma série de ferramentas informáticas que trabalhassem cada contributo individual de modo a criar todas as facilidades de manipulação e navegação pretendidas.

A exequibilidade de tal tarefa impôs que o material de base recolhido nas entrevistas fosse armazenado em formato digital de forma estruturada e sistemática, enriquecido com as devidas anotações, e fosse classificado à luz de determinada ontologia. Para isso tivemos de perceber ao pormenor o ciclo de vida de uma história . , desde a entrevista à sua publicação, o que nos levou a identificar com rigor e exaustão todos os tipos de documentos que suportam o funcionamento do Museu da Pessoa. Pudemos então definir aquilo a que chamámos a *arquitectura do Museu da Pessoa*, conforme se mostrou atrás

A facilidade com que anotámos a texto das entrevistas, fizemos a sua classificação, produzimos livros em formatos variados e pusémos o museu acessível via Web com as possibilidades de navegação esperadas, permitem-nos concluir que a abordagem seguida com suporte em XML foi uma escolha acertada. De igual modo, a opção pela linguagem de programação Perl para implementação das ferramentas, mostrou ser ajustada, pois essa linguagem é expedita para o manuseamento de documentos multi-formato e mesmo para o reaproveitamento de ferramentas existentes constituindo um ambiente cooperativo.

O processo de captação, tratamento e disponibilização de história snunca acaba. Iniciado em fins de 1999, já estão disponíveis, dos 70 depoimentos recolhidos, cerca de 30 histórias de vida, cada uma delas com qualquer coisa como 10 sub-histórias e perto de 3 fotografias. Com os actuais recursos, prevê-se, para os próximos tempos, um ritmo semelhante de crescimento nas recolhas, mas uma manifesta aceleração no processo produtivo interno. Contudo, nada obsta a que o resultado das várias parcerias, que tem vindo ou possam vir a surgir, se traduza num grande aumento.

Quanto à evolução técnica, muito há ainda a fazer, desde acabar/refinar algumas das ferramentas desenvolvidas até à criação de novas facilidades. Antes de mais, reconhecemos a necessidade urgente de repensar o site (quer o seu aspecto, quer a estrutura/formas de navegação) e de activar o thesaurus já construído para outros projectos (procedendo, então, ao seu enriquecimento).

Depois, é fundamental criar o suporte que permita o envio directo de contributos, tais como histórias (texto, fotografias, etc.) individuais, ou mesmo projectos de grupo (Museu da Pessoa nas Escolas, etc.).

Outra linha de trabalho futuro está relacionada coma construção de ferramentas genealógicas.

Com o aumento da quantidade de depoimentos podemos vir a conseguir construir árvores genealógicas e cronologias para que possamos relacionar as vidas dos vários intervenientes.

## Referências

- [1] Projecto museu da pessoa. <http://alfarrabio.um.geira.pt/mp>.
- [2] *eXtended Markup Language (XML) version 1.0 recommendation*. World Wide Web Consortium, 10 February 1998. <http://www.w3.org/TR/1998/REC-xml-19980210.html/>.
- [3] Almeida, J.J. Dicionário multi-fonte. In *Actas do XII Congresso da Associação Portuguesa de Linguística*, Braga 1996, 1997.
- [4] Almeida, José João. Projecto Natura - natural language processing, 1995. <http://natura.di.uminho.pt/>.
- [5] J.C. Ramalho & J.J. Almeida & P.R. Henriques. Algebraic specification of documents. *Theoretical Computer Science*, (199):231–247, 1998.

- [6] Leite, Carolina. Contos e histórias de vida nas ciências sociais. *Cadernos do Noroeste, Série Comunicação*, 12 (1-2):216–227, 1999.
- [7] Poirier, J. & et al. *Histórias de vida, teoria e prática*. Celta Editora, Oeiras, 1999.
- [8] Pollard, Richard. *Hypertext presentation of thesauri used in online searching*. Graduate School of Library & Information Science — University of Tennessee.
- [9] Ramalho, José Carlos Leite. *Anotação Estrutural de Documentos e sua Semântica*. Departamento de Informática, Escola de Engenharia, Universidade do Minho, 2000.
- [10] Veerasamy, Aravindan & Hudson, Scott & Navathe, Shamkant. *Querying, Navigating & Visualizing an Online Library Catalog*. College of Computing — Georgia Institute of Technology.