

Enquadramento e historial do Segundo HAREM

Diana Santos

Cristina Mota e Diana Santos, editoras, *Desafios na avaliação conjunta do reconhecimento de entidades mencionadas: O Segundo HAREM*, 2008, p. 1–7.

O HAREM é uma avaliação conjunta na área do reconhecimento de entidades mencionadas (REM) em português, organizada pela Linguateca. Uma avaliação conjunta, em poucas palavras, é uma tarefa que vários sistemas concordam em tentar executar, de forma a comparar o desempenho entre eles, com base em medidas consensuais e com recursos criados por uma comissão idónea – veja-se Santos (2007b) para mais informação sobre este paradigma de avaliação.

Neste texto pretendo fazer um historial da actividade da Linguateca nesta área e deixar expresso quais as muitas pessoas que participaram, e em que grau, na já longa história do HAREM.

O interesse pela avaliação na área foi inicialmente levantado no Encontro Preparatório sobre avaliação conjunta em 2002 (EPAV 2002), e uma parte significativa do trabalho da Cristina Mota, contratada na altura pela Linguateca para trabalhar no então pólo da Linguateca no LabEL, foi o de fazer um estudo preliminar das necessidades e interesses da comunidade, cujo relato foi feito no Avalon 2003 (Mota, 2003).

Quando houve finalmente oportunidade para organizar o que veio a ser chamado o Primeiro HAREM, já a Cristina se encontrava a trabalhar na sua tese de doutoramento sem ligação à Linguateca, e coube ao pólo da Linguateca no XLDB, em colaboração com o pólo de Oslo, a organização do Primeiro HAREM (de Setembro de 2004 a Julho de 2006).

A parte de leão coube ao Nuno Cardoso e a mim, embora contássemos com o apoio e colaboração de várias pessoas em fases diferentes dessa avaliação conjunta, que foi descrita em forma de livro em Santos e Cardoso (2007a) e onde essas contribuições estão bem documentadas. O Nuno acabou também por fazer a sua tese de mestrado, orientada pelo Mário J. Silva pelo XLDB e pelo Eugénio Oliveira pela FEUP sobre a organização e validação dos resultados do HAREM (Cardoso, 2006), num exemplo raro de uma tese em avaliação em Portugal, descrevendo, não um protótipo nem uma ideia para confirmação futura, mas uma actividade passada de interesse para toda a comunidade.

Quando a Linguateca recebeu financiamento para a sua terceira fase, de 15 de Dezembro de 2006 até fim de Dezembro de 2008, uma das iniciativas que constava do seu plano de actividades era precisamente a organização de um Segundo HAREM. Infelizmente era agora a vez de o Nuno Cardoso estar em trabalhos de doutoramento – aliás a maior parte das pessoas ligadas ao Primeiro HAREM também tinham deixado de ter ligação contratual à Linguateca – e, portanto, foi preciso criar uma nova equipa para organizar o Segundo HAREM.

Trabalho que me coube mais uma vez, mas com cujo resultado não me podia dar por mais satisfeita: A primeira “aquisição” foi a Cláudia Freitas, recém-doutorada em semântica computacional (Freitas, 2007) e uma fã do HAREM, que tinha vindo trabalhar para a Linguateca desde Julho de 2007 e graças em grande parte ao HAREM (com o qual tomara conhecimento no PROPOR 2006 e seus satélites). O segundo foi o Hugo Oliveira, também a trabalhar em semântica computacional (Gonçalo Oliveira et al., 2008) e chegado à Linguateca, em boa hora, em Setembro de 2007. Ambos no pólo de Coimbra. Quando o Segundo HAREM arrancou, por essa altura, éramos portanto apenas três, mas a vinda da Paula Carvalho, também recém-doutorada (Carvalho, 2007), em Dezembro de 2007, para engrossar as fileiras do HAREM, foi providencial, porque à data já estávamos a compreender que nos tínhamos comprometido a muito (talvez demais). A Paula encontrava-se afecta ao pólo da Linguateca no XLDB, de onde também o David Cruz foi chamado para ajudar à parte informática, tendo contudo sido substituído pela Cristina Mota, que ao acabar a sua bolsa de doutoramento da FCT foi recontratada pela Linguateca para o HAREM.

Note-se aliás que a Cristina apenas começou a trabalhar para a organização do HAREM a partir da data da avaliação conjunta, visto que participou no HAREM com o seu sistema. E ficou com a avaliação do TEMPO porque foi condição para a sua contratação não concorrer ao TEMPO.

Desde essa altura, esta equipa de cinco pessoas esteve a dar o seu melhor (embora em percentagens de afectação diferentes) à organização do HAREM, só se considerando o Segundo HAREM terminado à data da disponibilização do seu pacote de recursos, a LÂMPADA, o que ocorreu a 17 de Novembro de 2008.

Existem algumas diferenças fundamentais entre o Primeiro e o Segundo HAREM que convém serem destacadas aqui, de forma a este livro poder ser comparado de forma justa com o anterior.

1. Em primeiro lugar, o Segundo HAREM estava limitado pelo tempo escasso que restava à Linguateca, e que é mais compreensível na secção que faz o historial detalhado do Segundo HAREM, mas que é sumariado na figura 1.
2. Em segundo lugar, por se tratar de uma segunda edição, tivemos em muitos casos que efectuar compromissos entre razões históricas e de continuidade e razões científicas e de progresso, compromissos esses que não são fáceis e que o presente livro tenta documentar.
3. Em terceiro lugar, apostámos numa internacionalização do HAREM que não funcionou, tendo mandado todas as mensagens (e principal documentação) sempre em paralelo, em português e em inglês. Contudo, e ao contrário do Primeiro HAREM em que até houve quatro capítulos em inglês no livro, não houve nesta edição qualquer participação não-lusófona.
4. Em quarto lugar, tivemos a oferta de uma nova tarefa para realizar no âmbito do HAREM, a normalização do TEMPO, que foi muito gratificante e enriquecedora mas que deu origem a muitíssimos problemas e dificuldades que não podíamos prever, e de que fazemos aqui o historial mais detalhado, para explicar, entre outras coisas, a existência de dois capítulos do presente volume dedicados a essa pista. Convém contudo indicar que larga troca de correspondência e discussão de diversos pormenores da proposta do TEMPO teve lugar nos bastidores, entre o grupo e a organização do HAREM, e em que não só ficou clara a nossa divergência teórica como a autoria diferenciada da proposta. Quanto à forma de avaliar o TEMPO, devido à necessidade de integração com o resto do HAREM, essa ficou desde sempre atribuída à organização do HAREM, que teria também a obrigação de calcular os resultados e publicá-los. Isso explica, como já mencionado, a existência de dois capítulos sobre a organização desta pista: o capítulo 2 e o 3.
5. Em quinto lugar, e para garantir algum progresso também ao nível das tarefas a que os sistemas eram desafiados, lançámos mais uma pista, o ReReLEM, para identificar relações entre EM (capítulo 4).
6. Em sexto lugar, e talvez esta seja a diferença mais importante, muito do trabalho de estudo e reflexão sobre o Segundo HAREM terá de ficar por fazer – ou terá de ser feito no âmbito de um terceiro se vier a ser organizado, devido ao pouco tempo a que já aludi.

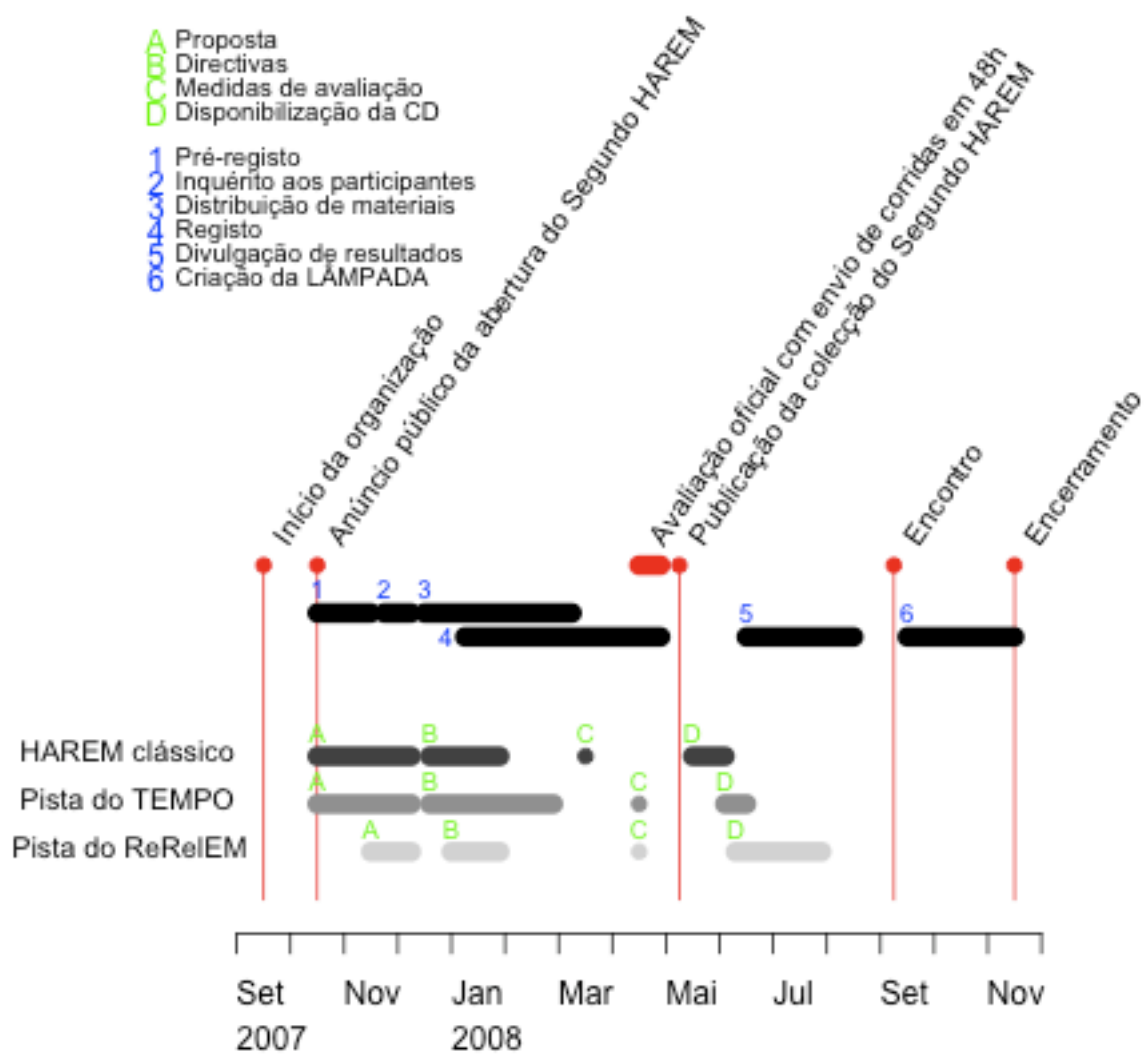


Figura 1: Diagrama temporal do Segundo HAREM

Em relação ao conteúdo do presente livro, a organização do HAREM como um todo considerou que, embora houvesse diferentes datas de início do trabalho e diferentes percentagens de afectação ao mesmo (a única que esteve a 100% no HAREM foi a Cristina), a equipa foi tão coesa e trabalhou tão bem em conjunto que não havia razão para apresentar artigos em separado, e por isso todos os artigos têm os cinco autores como autores. Refira-se, contudo, que a Cristina não participou na elaboração dos programas de avaliação do HAREM clássico visto que foi participante. Isso não nos pareceu razão suficiente para não ser co-autora do capítulo 1, que redigiu em conjunto com os outros autores.

Mais uma vez, quero realçar que o trabalho apresentado aqui é fruto indirecto de toda a equipa da Linguateca, e que em particular toda a infraestrutura, apoio organizativo e informático muito devem ao Luís Costa e ao Luís Miguel Cabral do pólo de Oslo.

Finalmente nunca é demais salientar que uma avaliação conjunta só funciona graças aos seus participantes, e que a comunidade do processamento da língua portuguesa é a principal beneficiada, mas também responsável e autora, do HAREM.

Historial detalhado

Aqui apresento as datas mais importantes do Segundo HAREM, para que fique registado o processo.

15 de Outubro de 2007 Abertura do Segundo HAREM, publicitada em várias listas.

15 de Outubro de 2007 Abordagem, por parte do Nuno Mamede, Caroline Hagège e Jorge Baptista (mais tarde chamados “grupo do TEMPO”), anunciando-nos a sua intenção de propor uma tarefa nova em relação ao tempo.

10 de Novembro de 2007 Prazo limite para registo dos participantes, tivemos manifestações de interesse de 22 grupos.

12 de Novembro de 2007 Mencionada a possibilidade da existência de uma pista de co-referência, em colaboração com a Renata Vieira (que mais tarde preferiu apenas participar).

14 de Novembro de 2007 Envio da primeira versão das directivas relativas ao TEMPO pelo respectivo grupo, por nós reenviada para a lista.

22 de Novembro de 2007 Inquérito aos participantes sobre os moldes de participação no TEMPO e no ReReLEM assim como outras questões.

3 de Dezembro de 2007 Prazo limite para discussão sobre os moldes do Segundo HAREM.

4 de Dezembro de 2007 Conclusões da discussão tornadas públicas, em particular com a proposta do TEMPO cabalmente aprovada.

10 de Dezembro de 2007 Exemplos preliminares do HAREM (primeira colecção de exemplo).

18 de Dezembro de 2007 Directivas preliminares publicadas, do HAREM e do TEMPO.

4 de Janeiro de 2008 Directivas iniciais do ReReLEM publicadas.

- 9 de Janeiro de 2008** Foi criado um formulário para que todos os participantes se registassem com uma definição precisa das tarefas que iam tentar.
- 17 de Janeiro de 2008** Como resultado, dois grupos afirmaram querer tentar o TEMPO completo, e cinco apenas o TEMPO clássico (ou seja, categoria, tipo e subtipo de EM tipo TEMPO). Em relação ao ReReLEM, houve quatro grupos inicialmente interessados. Finalmente, quanto a SUBTIPOS, apenas os dos LOCAL foram aceites, tendo sido rejeitados os de PESSOA e ORGANIZACAO propostos.
- 31 de Janeiro de 2008** Directivas finais publicadas, assim como material de teste congelado.
- 15 de Fevereiro de 2008** Adenda sobre o TEMPO divulgada.
- 15 de Fevereiro de 2008** Primeira versão do validador do Segundo HAREM disponível.
- 26 de Fevereiro de 2008** Nova versão das directivas do TEMPO divulgada.
- 4 de Março de 2008** Disponibilização do exemplário (para o HAREM clássico).
- 10 de Março de 2008** Primeira versão das medidas de avaliação, para o HAREM clássico.
- 10 de Abril de 2008** Primeira versão das medidas de avaliação para o TEMPO e para o ReRelEM.
- 14 de Abril de 2008** Abertura do prazo oficial do Segundo HAREM, tendo os participantes 48 horas desde que fossem buscar a colecção para a devolver anotada.
- 28 de Abril de 2008** Encerramento do prazo oficial do Segundo HAREM.
- 8 de Maio de 2008** Publicação da colecção do Segundo HAREM.
- 16 de Maio de 2008** Disponibilização da primeira versão da colecção dourada do Segundo HAREM para inspecção.
- 4 de Junho de 2008** Disponibilização da primeira versão da colecção dourada para o TEMPO completo para inspecção.
- 4 de Junho de 2008** Disponibilização da versão final² da colecção dourada do Segundo HAREM.
- 6 de Junho de 2008** Disponibilização da primeira versão da colecção dourada para o ReRelEM para inspecção.
- 12 de Junho de 2008** Disponibilização da versão final da colecção dourada para o TEMPO completo.
- 18 de Junho de 2008** Disponibilização da primeira versão dos programas de avaliação.

² Na altura foi divulgada como final, contudo foram sendo descobertos alguns problemas e houve uma CD oficial – para os resultados – e mais tarde ainda foram feitas umas últimas correcções até à versão derradeira disponibilizada na LÂMPADA. Note-se, todavia, que a lista dessas alterações foi enviada a todos os participantes, o mesmo acontecendo em relação às outras colecções douradas.

- 19 de Junho de 2008** Divulgação dos resultados preliminares do HAREM clássico.
- 25 de Junho de 2008** Divulgação dos resultados preliminares da avaliação do TEMPO.
- 31 de Julho de 2008** Disponibilização da versão final da colecção dourada para o ReReEM.
- 6 de Agosto de 2008** Divulgação dos resultados preliminares da pista ReReEM.
- 8 de Agosto de 2008** Divulgação dos relatórios finais individuais (exceptuando a pista ReReEM).
- 21 de Agosto de 2008** Divulgação de novos resultados da pista ReReEM.
- 7 de Setembro de 2008** Encontro do Segundo HAREM, como satélite do PROPOR 2008.
- 12 de Novembro de 2008** Divulgação das últimas alterações às colecções douradas.
- 13 de Novembro de 2008** Divulgação das últimas alterações aos programas de avaliação.
- 17 de Novembro de 2008** Pacote de recursos finais do Segundo HAREM, a LÂMPADA, disponibilizado.