## João Correia MIEIC 2015

### Pendente

* Ver
* [fcorreia@glint.pt](mailto:fcorreia@glint.pt)

### Arquivo

* 2016-05-19
  + Está a indexar o EResults (sistema de documentos) e a produzir templates EHR para o EHRServer; vai passar a fazer isto documento a documento
  + Em processo separado, indexa o EHRServer com o Solr
  + Falta ver o mecanismo de tornar incremental cada um dos processos
  + Como extra, há que acrescentar uma pesquisa avançada com gamas de datas (fica como sugestão as categorias de doenças e a área geográfica – a proximidade poderia ser critério de relevância)
  + Pedir à Carla um dicionário de sinónimos de doenças em português
* 2016-04-22
  + Querem usar o ERH como repositório comum para indexação quer do EResults quer do EPR; esses serão o resultado pretendido
  + Estão a pôr a hipótese de migrar parte da informação para o OpenEHR
  + Devolver o arquétipo com alguns dados como o nome do doente, a aplicação que produziu o documento, o diagnóstico, tipologia documental
  + Converter os dados do EResults para o EHR
  + Indexar os arquétipos com os documentos
  + Montar o alimentador incremental
  + Criar a interface com a pesquisa geral e eventual pesquisa avançada para restringir por datas ou por tipologias documentais
  + Mostrar o resultado num passo ou em dois para expandir o conteúdo antes de saltar para o documento
  + Conversão de termos da pesquisa com dicionários específicos
* [10:50:07] Joao Correia: ja estive a falar com eles e já estão esclarecidas as dúvidas de que falamos ontem
* [10:52:06] Joao Correia: esse modelo que acabei de enviar, diz respeito ao eresults, que é uma base de dados que praticamente só guarda documentos e metainformação associada a cada um dos documentos
* [10:52:56] Joao Correia: estes documentos também são para indexar mas como discutimos ontem não têm grande interesse usar o ehr neles
* [10:56:09] Joao Correia: a utilização do ehr é util noutra base de dados, DIGITALSIGNATURE, que guarda informações sobre pescrições e diagnósticos
* [10:59:13] Joao Correia: esta informação está distribuida por diversas tabelas, mas eles geram ficheiros xml desta forma:
* [10:59:15] Joao Correia: select
* xmlelement("ArrayOfClinicalNote", xmlagg(
* xmlelement(
* "ClinicalNote",
* xmlforest(papm.especialidade\_nota as "Specialty"),
* xmlforest(papm.dt\_cri as "Date"),
* xmlforest(papm.tip\_nota as "Type"),
* xmlforest(papm.observacao as "Observation")
* )
* ))
* from gh\_pc.pc\_area\_priv\_medico papm
* where papm.id\_area\_priv = :ID\_AREA\_PRIV
* [11:00:57] Joao Correia: resumindo, o xml resultante penso ja ter interesse de utilizar juntamente com o ehr
* [11:06:01] Joao Correia: ou seja a informação que estava nas tabelas passa a estar nesse ficheiro xml que também está guardado como BLOB numa tabela da base de dados
* [11:44:45] Gabriel David: OK. Embora esta BD seja muito virada para documentos, há um conjunto de dados que devem dar para preencher vários arquétipos. Se reparar, está ali muito clara a noção de episódio.
* [11:48:20] Gabriel David: Para sermos realistas, acho que se deve concentrar na BD e2\_results e no XML da DigitalSignature e não inventar mais. Tente arrumar os atributos de uma e de outra em arquétipos para poder justificar com a orientação pelo OpenEHR, embora sem a preocupação de ser exaustivo ou de estar a converter as BD para o CaboLabs. Mas sobre isto podemos falar depois.
* [11:52:11] Gabriel David: Interessa como resultado da sua indexação devolver listas com dois tipos de itens: 1) links para documentos do E2\_results, embrulhados nos arquétipos mínimos que identificar; 2) links para documentos XML da DigitalSignatures, também embrulhados nos arquétipos que identificar.
* FEUP, 2016-04-06
  + Tarefa: mapear os dados da Oracle/GLINT para o MySQL/OpenEHR
    - Delimitar este trabalho por frequência ou relevância dos dados
    - Arquitetura do sistema com sincronização
  + Definir quais os templates a indexar, por frequência ou relevância de acordo com o acima
    - Podem corresponder a documentos GLINT ou a templates EPSOS
  + Indexar uma palavra num arquétipo deve ir buscar o template respetivo ou o PDF respetivo e ordenar por preferência
  + Verificar se existe o conceito de episódio no EHR
* Francisco só falou em BLOB com documentos para indexar
* Mensagem no Skype 2016-03-29
  + Boas notícias, amanhã vou ter acesso aos dados. Relativamente às dúvidas com que fiquei na reunião anterior já estão todas esclarecidas à exceção do “dv\_coded\_text\_index” que eles também não têm grande certeza na sua importância para a indexação. Aparentemente diz respeito a códigos tipo o SNOMED, mas o Francisco diz que os documentos não terão grande informação com esses códigos (se bem que durante a pesquisa até pode ser importante, isto é a minha opinião).
  + Quanto à informação a indexar o Francisco acha que devo começar com texto e datas. A meta-informação também deve ser possível de ser pesquisada por isso também vai ter que ser indexada de alguma forma e não fazer o pedido à db na altura da pesquisa como tínhamos falado na reunião. Isto como é óbvio aumentará novamente a quantidade de pedidos à base de dados por parte do solr. No entanto vou tentar manter o menor número de pedidos possível.
  + Relativamente ao openEHR, uma coisa que percebi mal no início, é que não sou obrigado a utilizar o cabolabs (aquele que lhe mostrei na ultima reunião). Torna-se um pequeno problema devido ao tempo mas a boa noticia e que já escrevi sobre outros ehrs para o estado da arte por isso já sei quais os que vou testar a ver se são melhores ou não. Quanto à estrutura de cada não deve haver problema uma vez que são todos semelhantes. Penso demorar no máximo hoje e amanhã de manhã pois não quero mesmo perder muito mais tempo com testes.
  + Uma vez que ontem ainda não tinha falado com o Francisco fui experimentando criar uma api em .Net e não parece nada complicado. Aliás ontem cheguei mesmo a conseguir criar uma api simples a pesquisar um produto por id.
* Resposta
  + São de facto boas notícias. Assim pode avançar com mais segurança. A questão da indexação dos metadados tem que ser vista com cuidado. Acho bem que se indexem aqueles metadados que são úteis na pesquisa, tais como o nome e, eventualmente, a data de nascimento. Mas há vários outros aspetos que podem ser interessantes numa perspetiva de pesquisa e que tenho dúvidas que devam ser indexados. Mas há um segundo nível que é o da apresentação dos resultados. Quando temos documentos, é uma questão de guardar os links e mostrar os documentos. No caso das BD, o mais próximo é o template ou formulário em que a informação aparece ou que se constrói explicitamente para a resposta. Isso implica ir consultar a BD na hora. E até se podem prever alguns filtros simples no formulário do estilo gamas de datas de nascimento ou de ocorrência dos episódios, áreas geográficas, profissões dos doentes, etc. para ajudar a filtrar os resultados da pesquisa em texto integral. Se é necessário ir à BD, então podem-se implementar estes critérios.
* 2016-03-03
  + Usar especificação EHR
    - Implementação da CaboLabs
  + Passar de Oracle para mysql
  + Andou a fazer testes com OpenEHR em mysql
  + Vai indexar o conteúdo do OpenEHR
    - Possivelmente não vai ter que indexar ficheiros
* 2016-02-26
  + Nota de Projeto: 12
  + Nota de estado da arte: 11
  + Documento de PDIS
    - Aumentar um bocadinho a descrição da Glint e do ambiente computacional no cliente típico da Glint, onde e como é a informação produzida e armazenada
    - Como visualizar o resultado da pesquisa?
* FEUP, 2016-02-08
  + Apresentação a 1 de fevereiro
  + É necessário categorizar
  + Como relacionar
* GLINT, 2015-11-18
  + Indexação de documentos clínicos
    - Em ficheiros fora e dentro da BD
    - Em campos da BD clínica da GLINT
  + Inovação está na integração e na perspetiva clínica
  + Como apresentar o resultado, no caso da BD? Talvez a tabela, linha e coluna; depois relacionar; ver linked data; chaves externas a sair e a entrar
  + Ferramentas
    - Analisar e comparar as ferramentas
    - Lucene
    - SolR
  + Público alvo
    - Médicos
  + Indexação na perspetiva clínica
  + Resultados em API RDF? Para reutilizar
  + Como interrogar?
    - Alertas em contexto de prescrição
    - Já há em contexto de diagnóstico (serviços que dão informação relacionada com um dado diagnóstico
    - Para já partir de uma caixa de pesquisa com uma query texto
  + Não há muita equipa a trabalhar nisto, é mais o supervisor, com o Pedro ou o Correia
  + Usar os sistemas da GLINT