

IBI5013 – Bancos de Dados para Bioinformática

Projeto 2 – Escrita de um Artigo sobre Bancos de Dados Científicos / Bancos de Dados Biológicos

Prazo para a definição do grupo e do tema: 07/06/2017

Prazo para a entrega do projeto: 30/06/2017

1 Descrição do Trabalho

Para o segundo trabalho de IBI5013, cada grupo deve selecionar um tema para estudo relacionado aos listados na Seção 2 deste documento.

Cada grupo deverá entregar no Paca até o dia 07/06/2017 a composição do grupo e o tema escolhido. Cada grupo deve ter 2 ou 3 componentes.

Até o dia 30/06/2017, cada grupo deve entregar no Paca um texto com os resultados do estudo realizado. O texto deverá ter o formato de um artigo técnico-científico e deverá conter as seguintes informações:

1. Contextualização e motivação do tema;
2. Principais conceitos relacionados ao tema;
3. Descrição do estado da arte de trabalhos teóricos e também de ferramentas de *software* ou sistemas *web* relacionados ao tema;
4. Um pouco mais de detalhes sobre ao menos uma ferramenta ou sistema *web* escolhido para ilustrar conceitos importantes abordados no trabalho. Se pertinente, use um estudo de caso para reportar o uso da ferramenta ou sistema *web* escolhido na criação/manipulação de um banco de dados de exemplo.

Note que a lista acima se refere ao conteúdo do texto, mas não à sua estrutura. Cada grupo terá a liberdade de escolher a estrutura que julgar mais conveniente para o tema que abordará. Entretanto, os grupos devem se atentar para o fato de que deverão produzir um texto científico e que, portanto, deverão ter um cuidado especial com relação à apresentação dos resultados obtidos a partir dos estudos realizados. Os conceitos devem ser apresentados formalmente. A discussão sobre o estado da arte deve considerar referências bibliográficas de origens confiáveis. O estudo de caso deve apresentar todos os detalhes necessários para que um leitor do artigo seja capaz de reproduzi-lo com sucesso (caso assim deseje).

2 Temas Propostos para o Estudo

Os temas propostos aos grupos são os seguintes:

- Modelos de dados e formatos de arquivos padronizados para dados e metadados de bioinformática
Exemplos:

- Formatos de arquivos: FASTA, Genbank/EMBL/DDBJ, etc.
- Modelos de dados: *guidelines* reunidos pelo projeto MIBBI (https://biosharing.org/standards/?selected_facets=isMIBBI:true)
- Bancos de dados biológicos públicos (falar sobre uma categoria específica)
Exemplos de categorias de bancos de dados biológicos:
 - Sequências de nucleotídios; Sequências de RNA; Sequências de proteínas; Genômica (não vertebrados); Vias metabólicas e de sinalização; Genoma humano e de outros vertebrados; Doenças e genes humanos; Dados de microarranjos e outras expressões gênicas, etc.
- O uso de sistemas gerenciadores de bancos de dados científicos em bioinformática
Exemplos de SGBDs científicos:
 - SciDB (<http://www.paradigm4.com/>), rasdaman (<http://www.rasdaman.org/>), MonetDB (<https://www.monetdb.org/>)
- O uso de sistemas NoSQL em bioinformática
Exemplos de sistemas NoSQL:
 - Neo4J (<https://neo4j.com/>), MongoDB (<https://www.mongodb.com/>), Cassandra (<http://cassandra.apache.org/>)
- Sistemas de gerenciamento de *workflows* (= processos de análise de dados) para Bioinformática
Exemplos:
 - Taverna (<http://www.taverna.org.uk/>), Galaxy (<https://usegalaxy.org/>)

Quando o tema envolver ferramentas de software, tentem priorizar a escolha de ferramentas de software livre, porque assim vocês poderão instalá-las e testá-las sem restrições.

Aviso: A professora poderá “interferir” na atribuição dos temas aos grupos caso muitos grupos escolham um mesmo tema. Caso isso seja necessário, usaremos o final da aula do dia 08/06/2017 para chegar a um consenso sobre a distribuição dos assuntos.

Caso um grupo tenha alguma outra proposta de tema pertinente e que gostaria de desenvolver neste trabalho, então o grupo deve conversar com a professora a respeito disso antes do dia 07/06/2017.

3 Critérios de Avaliação

O texto entregue como resultado do Projeto 2 será avaliado quanto ao seu conteúdo segundo os requisitos especificados na Seção 1. Quanto à forma, serão avaliadas:

- A conformidade do texto com a estrutura e a formalidade esperada de um artigo técnico-científico;
- A qualidade das referências bibliográficas usadas e da apresentação das mesmas no texto;
- A qualidade das figuras e tabelas usadas para complementar as informações no texto.

O uso de \LaTeX para a confecção do artigo a ser entregue é **fortemente** recomendado.