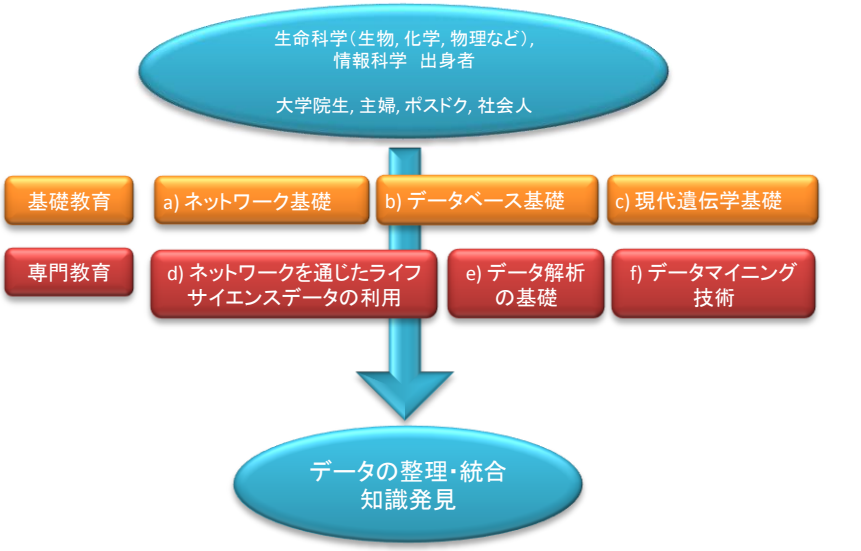
DB高度利用者の養成：
情報量が増え続けるデータベースを使いこなすために

お茶の水女子大学　生命情報学教育研究センター
金子　聡子
お茶の水女子大学　大学院人間文化創成科学研究科
瀬々　潤

データベース高度利用者養成



趣旨

実験技術の進歩に伴い、最近の生命科学では情報爆発が起こっている。この情報を活かし生命科学で効率よく知識発見に繋げるために、データ管理と利用の中間に位置し、様々なデータの整理・統合及びデータからの知識発見を行う「DB高度利用者」の役割が重要になってきている。しかし、DB高度利用者には、生命科学の知識のみならず、情報科学の技術など、多様なスキルが要求されるため、人材養成はあまり進んでいなかった。

お茶の水女子大学では、この「DB高度利用者」に必要なスキルを調査した結果、今後のDB高度利用者に必要且つ今までの人材養成では不足していた内容として、ネットワークを通じてデータを利用する手法と、知識発見技術の習得が浮かび上がった。そこで、これらの知識を体系的に扱うカリキュラムを策定し、現在はDB高度利用者養成の講義・演習を行っている。

基礎教育

体系的な講義により情報科学と生命科学の知識を補完

a) ネットワーク基礎
コンピュータネットワークおよびインターネットの歴史と仕組み／階層プロトコルとTCP/IP／ネームサーバ／TELNETとFTP／ルーティングなど

b) データベース基礎
生物学におけるデータベースの歴史／リレーショナルデータベースについて／MySQLによる演習など

c) 現代遺伝学基礎
細胞やゲノムの構成／転写調節／翻訳／タンパクの立体構造／分子生物学実験の原理／遺伝学の歴史／進化学の概論など

専門教育

d) ネットワークを通じたライフサイエンスデータの利用
ライフサイエンス統合データベースセンターが提供するWebAPIサービス(TogoWS)の利用とデータ処理で頻出する作業の効率化と自動化を学ぶ

UNIXの基礎／ゲノムビューアの使い方／相同性検索(Web, ローカル, EMBOSSの利用)／WebAPIの利用(Google Chart API, TogoWS(SOAP, REST))／相同性検索の自動化／系統樹を描くなど

e) データ解析の基礎
ライフサイエンスデータを扱うためのプログラム手法と基礎的なデータの統計処理を学ぶ

プログラミングの流れ／変数と型／制御文／メソッド／配列／文字列／ArrayList／JFreeChartを利用したグラフ描画／相同性検索など

f) データマイニング技術
高度なデータ処理技法を学ぶ事で、新しいデータベースの創発を目指す

決定木／Naive Bayes／AdaBoost／k-Nearest Neighbor／Support Vector Machine／k-means／クラスタリング／Association Ruleなど

自習教材

http://togodb.sel.is.ocha.ac.jp/

実績, 感想, 進路

-実績-
平成19年度
専門　延べ受講人数 18人 (90分×10コマ×2)

平成20年度
基礎　延べ受講人数 14人 (90分×3コマ×3)
専門　延べ受講人数 31人 (90分×15コマ×2)

-受講者の感想 (平成20年度実施分)-
・塩基配列の扱い方や系統樹の作成方法は、論文を書く上で必要なことだったので、助けになった。
・Javaでのプログラミングを段階的に教えてもらえたので、分かりやすかった。
・普段から相同性検索は行っていたので、講習内容を用いて作業が速くなった。

-受講者の進路など-
卒業生は、製薬会社やシンクタンクのSE, 化粧品メーカーなどに就職しています。
在学生は、統合TVや自習教材のコンテンツ作成に活躍しています。

講義風景

平成20年度　専門　(ネットワークを通じたライフサイエンスデータの利用, データ解析の基礎)

今年度の予定

DB高度利用者基礎
現代遺伝学基礎　7月28日
ネットワーク基礎　7月30日
データベース基礎　7月31日

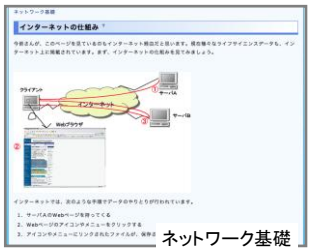
DB高度利用者専門
ネットワークを通じたライフサイエンスデータの利用
5月11日-7月13日(毎週月曜),
8月3日-7日, 9月7日-11日 *同じ内容を行います

データ解析の基礎
8月10日-12,18,19日, 9月14日-18日
*同じ内容を行います

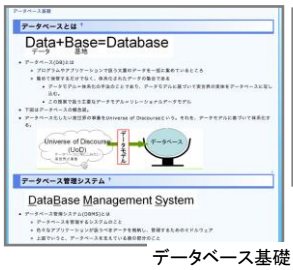
データマイニング技術
後期(毎週木曜)

受講者を募集しています。
togodb-info@sel.is.ocha.ac.jpまでご連絡下さい。

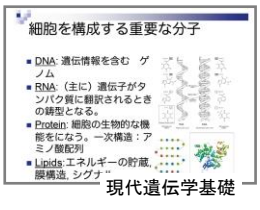
講義資料 (例)



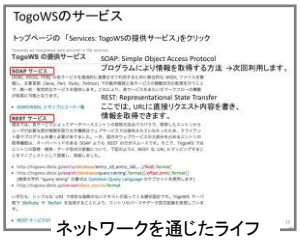
ネットワーク基礎



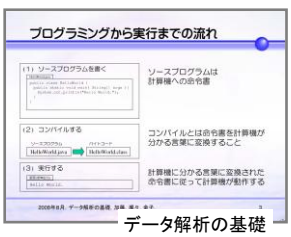
データベース基礎



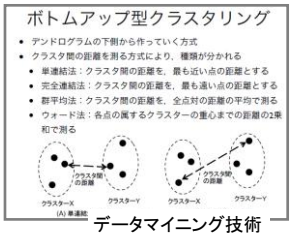
現代遺伝学基礎



ネットワークを通じたライフサイエンスデータの利用



データ解析の基礎



データマイニング技術