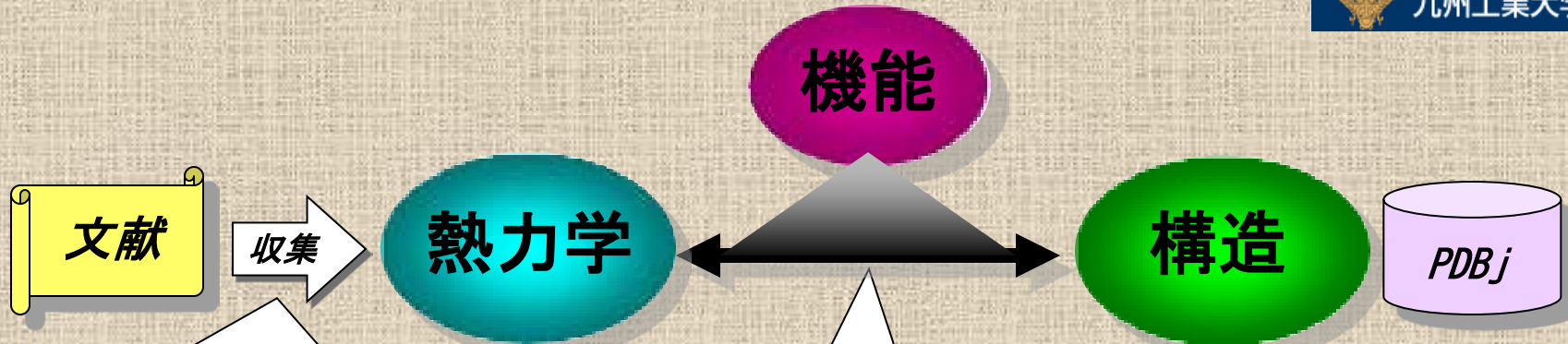


# 生体分子の熱力学データと構造データの統合



## テキストマイニング技術

- データを含む論文の自動収集
- テキストからのデータの自動抽出

## 統合化技術

- 熱力学情報と構造情報のクロスレファレンス
- オントロジーの整備
- データ交換フォーマットの整備
- 検索の統合 (APIの共通仕様)

# 今年度の経過

- 蛋白質と変異体の熱力学データと構造データを対応させるクロスレファレンスを約600件作成
- 蛋白質と核酸の相互作用の熱力学データを対応させるクロスレファレンスを約500件作成
- 蛋白質・蛋白質相互作用データベースのスキーマの設計を行い、データベースシステムとWebページを作成中
- 蛋白質・核酸相互作用熱力学データについてXMLフォーマットのプロトタイプを作成
- テキストマイニングにより文献検索と文献からのデータ抽出を自動化する方法を統合DBセンターと連携して開発中