

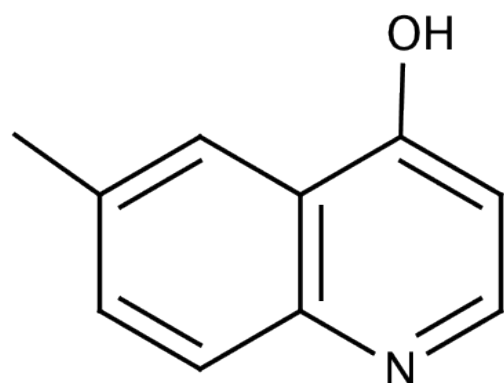
化学構造式画像 認識システム

ChemInfty (ケムインフティ)

研究プロジェクト

科学技術振興機構 研究成果最適展開支援事業（育成研究）

平成 21～23 年 文献や特許データベース中の化学構造式の認識と検索



化学構造式
画像

原子部分の記述		結合部分の記述	
M	V30 BEGIN ATOM	M	V30 BEGIN BOND
M	V30 1 C -1.76 0 0 0	M	V30 1 1 1 2
M	V30 2 C -3.0936 0.77 0 0	M	V30 2 2 1 6
M	V30 3 C -4.4274 0 0 0	M	V30 3 1 1 10
M	V30 4 C -4.4274 -1.54 0 0	M	V30 4 2 2 3
M	V30 5 C -3.0936 -2.31 0 0	M	V30 5 1 3 4
M	V30 6 C -1.76 -1.54 0 0	M	V30 6 2 4 5
M	V30 7 N -0.4264 -2.31 0 0	M	V30 7 1 5 6
M	V30 8 C 0.9074 -1.54 0 0	M	V30 8 1 6 7
M	V30 9 C 0.9074 0 0 0	M	V30 9 2 7 8
M	V30 10 C -0.4264 0.77 0 0	M	V30 10 1 8 9
M	V30 11 C -5.28 0.715 0 0	M	V30 11 2 9 10
M	V30 12 O -0.55 1.98 0 0	M	V30 12 1 11 3
M	V30 END ATOM	M	V30 13 1 12 10
		M	V30 END BOND

ChemInfty

業界標準形式
MOL, SDF等

データベース

検索サービス

描画
ツール

化学構造式画像を
検索可能な標準形式へ変換！

光学的文字認識技術(OCR)と構造解析技術により、印刷物中の化学構造式画像を認識し検索可能な業界標準形式に変換します。

これにより、従来は手作業で行っていた化学構造式画像の電子化作業が効率化します。

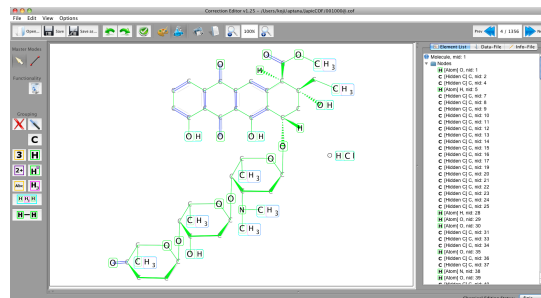
電子化されたデータは業界標準形式のため、既存のツールと連携することができます。



ChemInfty は、画像の認識エンジンと修正インタフェースから構成されています。

認識エンジンに化学構造式画像を入力すると認識結果が得られます。認識結果が誤認識を含んでいる場合には、操作性の高い修正インタフェースにより、人手で効率よく修正できます。

修正結果は、業界標準の形式(SDF, MOL 形式等)に変換されますので、化学構造式を取り扱う既存の様々なツールやサービスと連携した処理が行えます。



修正インタフェース

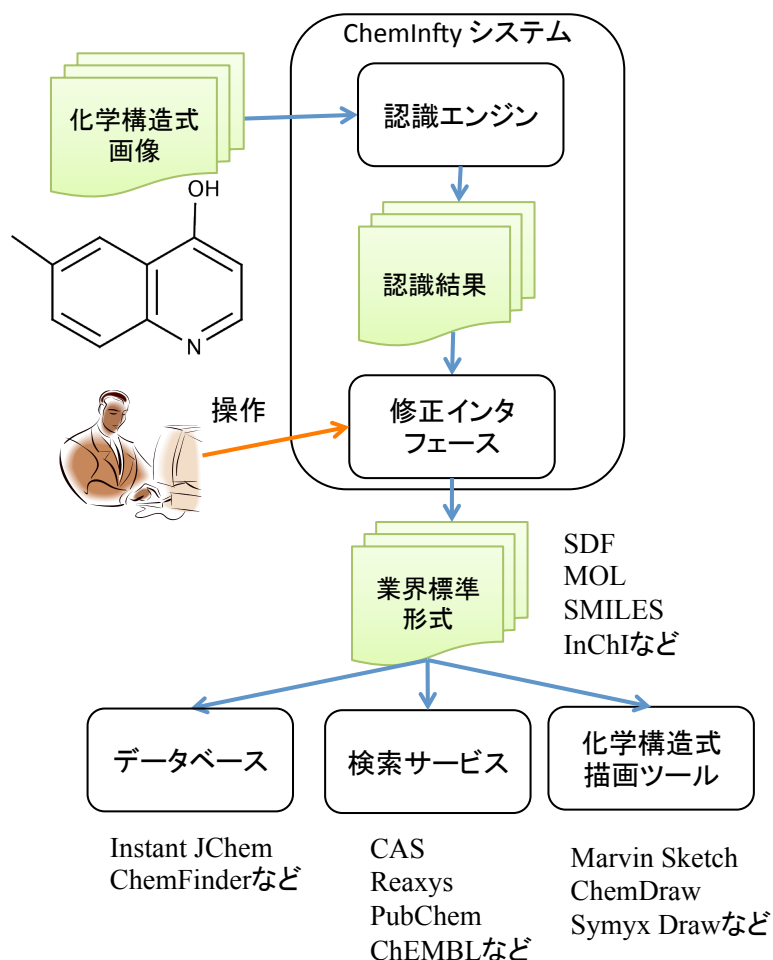
➤ 接触画像に対応した認識エンジン

認識エンジンは、スキャン時や描画時に発生する線と文字の接触や文字同士の接触画像も認識します。

➤ 操作性の高い修正インタフェース

化学構造式としての整合性や画像の情報を利用した、認識結果を修正するための操作性の高い修正インタフェースを提供します。

現在、実用化に向けたプロトタイプシステムを開発中です。ご意見、ご相談お待ちしております。



認識対応要素

- 立体結合
- 示性式表現
- Markush 構造

動作環境

- 認識プログラム
 - Windows XP 以上
- 修正インタフェース
 - Java の動作する環境
 - ◇ Windows
 - ◇ MacOS X

連絡先

〒 814-0001

福岡市早良区百道3丁目8-3 4

JST イノベーションプラザ福岡

九州大学数理学研究院 鈴木研究室
(H24 年 2 月末まで)

E-mail : contact@infty.biz

ウェブページ:

<http://www.inftyproject.org/jp/ChemInfty>

平成 23 年 12 月 1 日発行