

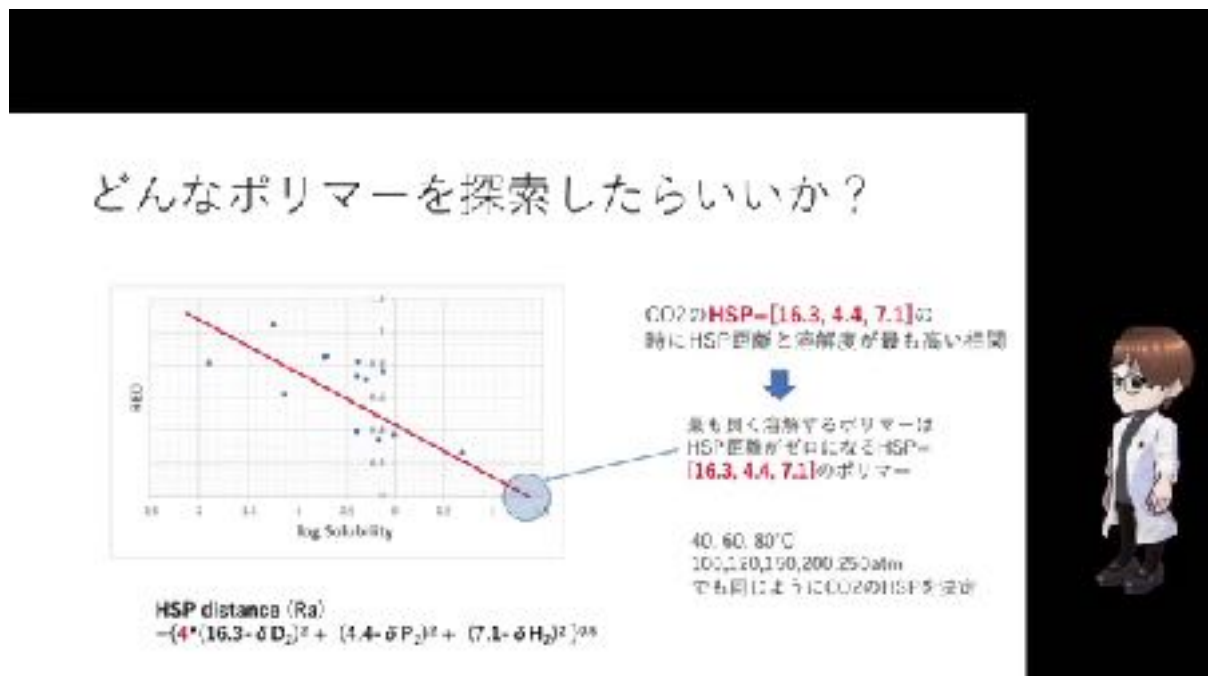
pirika.comの研究記録

Pirika News

V-Tube & ブログ

たまには環境問題に取り組もう。

紙おむつが水を吸うように炭酸ガスを吸うポリマーを開発してみましょう。



水はH₂O、炭酸ガスはCO₂。CO₂分子の方が大きいけれど、水素結合しないのでなかなか凝縮できない。天然ガス中に含まれているCO₂や排ガス中のCO₂が紙オムツに吸着されたら面白いな。

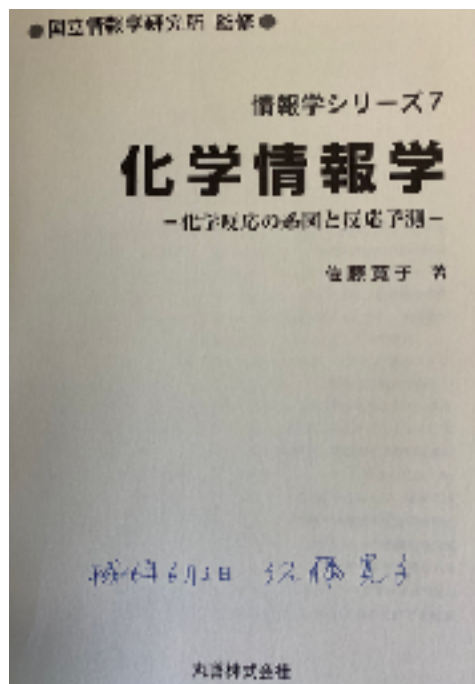
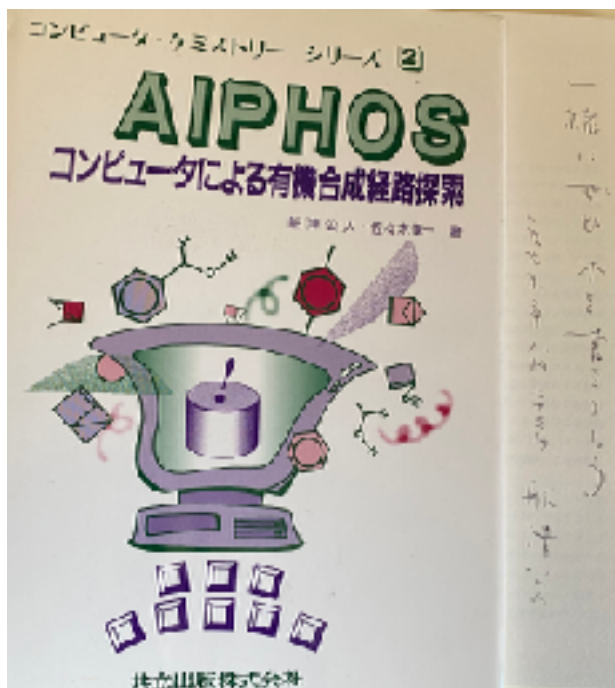
ほとんど妄想でやってみると、可能性のあるポリマーが設計できた。

しかし、単位空間あたりCOOが一番大きくなるという設計で、HSPはあまり役に立たなかった。誰か酔狂な人がいたら実際に重合してほしい。

書籍のEditor その後

"Drug Development Supported by Informatics"というSpringerの書籍のEditorの仕事をしている。自分でも2章ほど書いたことは先月触れた。

他のEditorは、奈良先端科学技術大学院大学の船津公人先生と、University of Zurich の佐藤寛子先生だ。両先生とは古くからの付き合いだ。



「一緒にぜひ本を書きましょう」と船津先生からサインを頂いて、四半世紀。やっと夢が実現した。

先生たちは計算機化学といえば量子化学だった時代に、ケモインフォマティクスで頑張っていた。

私はちょうど筑波での国家プロジェクト(1995-1999)で代替フロンを研究していた頃だ。私はCALTECHのGoddard教授のところに留学(1990-1991)して計算機化学を学んではいたが、元は、合成系の高分子屋。国研には量子化学の大御所がいて計算しているので、私なんかはお呼びでない。そこで、どうせMO計算には時間がかかるので、jobは放り込んで、その間、ケモインフォマティクスに力を入れていた。お茶の水女子大学の長嶋雲兵先生に助けてもらいながら、自分でニューラルネットワークのプログラムを作成し、物性推算、逆設計に取り組んだ。その時に両先生と出会った。

そんな、(色々な紆余曲折はあったが)ご縁で、インフォマティクスの書籍の編集に取り組んでいる。(今はそんなわけで他のことはほとんどできていない)

Editor本来の仕事は、他の先生の手原稿をSpringerの流儀に合っているかチェックする。そして、参考文献や図表の参照が正しいかをチェックする。

それぞれが専門の先生が書いているものなので、中身に関するチェックはほとんどしない（というかできない）。

私は先月にも書いたが、フォーマットは不得意だ。参照論文のチェックと、ロジックチェックを主に担当した。

というか、他の先生の手原稿チェックをしながら（修正されたところを見ながら）自分の原稿をコソコソ直していた。

編集の作業は今回が初めてだったが、本当に勉強になった。

多分、今後の書籍編集でものすごく重要になること。

生成系AIに書かせたものが入り込んでいないかチェック。

ChatGPTに『XXXXに関するレポートはお前が書いたものか?』と聞くと、YYY%で私の書いたものです。って答えたとかいう話も聞いたことがある。

学生の手書いたレポートをチェックするAIもあるという。

大学によってはレポートを生成系AIに書かせたら一発退場だという。

ChatGPTに書かせた日本語のものを、DEEP-Lで英語にして、それをGoogle Translateでフランス語にして、ChatGPTに英語に変換させたら、見破るかな?

まー、でも結局は使う人のレベル次第だろうな。

中身がわかっていなければ、生成系AIが知ったかぶりで書いた文章を本気に受け止めてしまう。変換ルートを複雑にしても最終結果はあり得ない知ったかぶりの内容になる。

プロンプト当てクイズ!

どのようなプロンプトを与えたら、ChatGPTは次のような答えを返すでしょうか?

ケモインフォマティクスは、化学データの収集、分析、解釈に情報技術と計算手法を応用する学問分野です。ドラッグデザインにおけるケモインフォマティクスの役割は、AI（人工知能）とML（機械学習）を通じて、新しい薬剤候補の同定、薬剤の効果や安全性の予測、および薬剤作用のメカニズムの解明にあります。AIとMLは以下のような具体的な方法でケモインフォマティクスを強化しています。

1. ****薬剤候補の同定****: MLモデルは、大規模な化合物ライブラリから潜在的な薬剤候補を迅速にスクリーニングし、有望な化合物を特定します。
2. ****活性予測と副作用の評価****: MLアルゴリズムは、化合物の構造データからその生物学的活性や潜在的な副作用を予測し、より安全な薬剤設計を促進します。

3. ****薬剤作用メカニズムの解明****: AI技術は、複雑な生物学的ネットワーク内での化合物の作用機序を解析し、薬剤の標的となる生物分子や経路を特定します。

4. ****最適化とパーソナライズ****: AIを用いたモデリングは、薬剤の化学構造を最適化し、個々の患者に最適な薬剤を設計するためのパーソナライズドメディシンへの道を開きます。

ケモインフォマティクスにおけるAIとMLの応用は、ドラッグデザインプロセスを加速し、薬剤開発のコストを削減し、新しい治療薬の発見を促進する大きな可能性を秘めています。これにより、より効果的で安全な薬剤が患者に迅速に届けられるようになります。

「産業洗浄」執筆

日本産業洗浄協議会というところから会誌への執筆を依頼された。まー、2月中に書き上げよう。以前に書いた「化学と教育」もそうだけど、紙面は白黒だ。こういうpdfとかwebページのような見せ方はできない。そこで原稿に2次元QRコードを埋め込むことを許可してもらった。



<https://www.pirika.com/Program/SV/PSto1S.html>

例えば、こんなふうに。

原稿を読みながらiPhoneでぐりぐりする。新しい雑誌の読み方になると思う。

2年目決算終了

合同会社、pirika.comを設立してから2度目の決算が終了した。たっぷり税金も取られた。でも、内部留保も順調に増えている。

交際費、会議費、使途不明金、使途秘匿金：

会社を経営しているとこの違いには敏感になる。

使途不明金：お金の支出は領収書があり明確だけど、何に使われたのか但し書きがない。経費には計上できないので支払う法人税が増える。

使途秘匿金：使ったお金を故意に隠していると判断されるお金のこと。政治家や総会屋などに渡すお金。経費にできないだけでなく、悪質なので重加算税を取られる。

交際費：事業を行う上で必要な接待のこと。私のサイズの会社なら交際費は800万円までか、飲食費の50%を損金算入、簡単に言えば経費にできる。いつ・どこで・誰と・何のために飲食したかを明確にする必要がある。一人当たり5000円以下であれば会議費として計上できる。

会議費：社内で行われた会議や取引先との打ち合わせで発生した費用のこと。会議費は全額経費として計上できるが、通常の領収書に加え、参加した者の氏名、名称と関係。参加者の人数。

先月だかに、固定資産税の額っておかしくないか書いた。交際費の限度も物価の上昇に合わせて上げるべきだ。そういうのを決めるのが政治だろうが、政治家は裏金作りに忙しくてそれどころではないらしい。政治家のパーティーへの参加はどうしているんだろう？使途秘匿金にしているなら、今の1000倍ぐらい重加算税を取ればどうだろう？誰も政治パーティーに参加したいと思わせなくすることが大事だ。どっちみち政治家は「覚えていない」ので投資効果はない。

政治家の報告書。使途不明で済ましているが、これは使途秘匿金だろう。これも今の1000倍ぐらい重加算税を取れば良い。そうすれば、政治家がパーティーをすればするほど、税収が増える。政治と癒着してボロ儲けしている会社からの税収が増える。

政治家の支出のうち、交際費、会議費は相手を公開しないので認められない。相手はいるはずなので、使途秘匿金だ。政治家はすぐに、個人的支出はなかったとかいうが、政治家の胃袋は公と個人に別れているのだろうか？ 会社員は出張しても、食事代は経費精算できない。家では食べないわけではないから。

使途として、書籍代を3000万と記載した政治家がいた。当然領収書があるはずだし、検収

が必要だ。政治家が1万円の本を書く。それを3000冊購入した。それは賄賂だろ。それを地元で配った？買収だろう？

それを秘書のせいにするなら、罰則を懲役1000年ぐらいにすればいい。もしくは公民権100年停止とか。汚れ役を被れば後で引き上げてもらえると思っているから、ヤクザと同じで身代わりになる。娑婆に出て来れなくなれば誰も身代わりにならない。

アカデミアへの対応

ある大学から問い合わせを受けて、こんな返事を書いた。

最初に、誤解の無いようにお伝えしておきます。

私は、pirika.comという合同会社のCDOです。

主な業務はHSPiPの販売とコンサルです。

従って、アカデミアから今回のような相談があったところで、費用をお願いすることはありません。先生のお手伝いをすることで、良い論文が出れば、HSPiPの販売が増えれば十分と考えています。論文や特許に関しても先生のご自由にやっていただいて結構です。

関西の方の大学のように、HSPiPを使いながら、一言もそれに触れずに論文を書いたり、Y-MBのクラッキングを博士論文としているケースもありますが、それはアカデミアの自殺だと認識しています。基本普通の引用があれば、こちらは問題ありません。逆に、アイデアなどを実現できる可能性（今の私には手足はありませんので）に対してワクワクします。

他の大学でのやり方の一例では、研究室の雑誌会のようなもの、学部の外部講師の講演会みたいなもので話すのはよくあるやり方です。

どんな話が聞きたいかを教えていただければ、合やすことは可能です。

それとは全く別のフェーズとしては、大阪の大学の先生から、pirikaの研究会で学生を受け入れてほしいとお願いされたこともあります。

HSPだけでなく、データサイエンスも含めての育成です。

私なんかは、なんのネームバリューもない者ですが、どうせなら日本と思いながら、くる話は、コンサルも含め外資系ばかりです。

時期も含め、ゆるゆる詰めていかれればと思います。

まともな大学には協力するということだ。

最近聴いた音楽

知床旅情の中の「ピリカが笑う」ってアイヌ語のピリカで、pirika.comの語源だったんだ。びっくり。

そんなこと言っていたら、ピリカピリカと言うアイヌの民謡があることを教わった。Apple Musicで検索するとその曲が出てきた。すごいなー。これ社歌にしよう。

最近作ったプログラム

Excelを使ったHSPiP用のデータ作成用のOfficeScriptのプログラムは以前作ってV-tubeでも説明している。

DXによるHSP用データの生産性向上

<https://youtu.be/rYRH3zz11ZU>

今回は、データを10Kに拡張して、hsdxだけでなくsofxのデータも作成できるようにした。sofxというのは、Solvent optimizer File xml の略だ。

以前のフォーマットはhsdとsofだった。これは単なるタブ区切りのテキストファイルなので、Excelに持って行って処理するのが楽だった。Abbott先生はユーザーにExcelを使って欲しくない。HSPiPだけで全部やってほしい。そこでxmlに変えてしまった。

そこでコンバータは自分で作らなくてはならない。

まー簡単なのでトライしてみるのも良いだろう。

どれでも良いから、HSPiP Dataフォルダーにある、XXX.sofxをテキストエディターで開いてみよう。

```
<?xml version="1.0" encoding="utf-8"?>
<HSP_Data>
  <Chemical>
    <Solvent>Acetone</Solvent>
    <δ D>15.5</δ D>
    <δ P>10.4</δ P>
    <δ H>7</δ H>
    <δ HD_x002F_A>0.1/7.0</δ HD_x002F_A>
    <MVol>73.8</MVol> //分子体積
    <Other_x0020_Names></Other_x0020_Names>
    <RER>560</RER> //Relative Evaporation Ratio
```

```

<Weight>100</Weight> //分子量
<Vol_x0025_>-</Vol_x0025_>
<S>FALSE</S>
<Distance>6.98</Distance>
<AA>7.117</AA> //Antoine A
<AB>1210</AB> //Antoine B
<AC>229.7</AC> //Antoine C
<BPt>56</BPt> //沸点
<FPt>-17</FPt> // Flash point
<VP_x0040_25>232</VP_x0040_25> //25°Cの蒸気圧
<MPt>-95</MPt> //融点
<MWt>58</MWt> //分子量
<LogS>2</LogS> //水への溶解度 g/100g水
<CAS>67-64-1</CAS>
<SMILES>CC(C)=O</SMILES>
<Settings>-</Settings>
</Chemical>
<Chemical>

</Chemical>

</HSP_Data>

```

このフォーマットに合わせてsofxファイル作れば、自分用のファイルを作ることができる。と口で言うほど簡単ではない。

以前なら、色々なsofファイルから必要な溶媒をExcel上で集め、タブ区切りでセーブすればよかった。今はコンバータを作らない限り、HSPiP上で一つずつ手入力しなくてはならない。その時に実験値、オフィシャル値があればそれを使い、それがない場合にはY-MBで計算しなくてはならない。

Abbott先生の作った様々なsofxファイルにどこまで実験値の値が入っているかは、わからない。それは、我々のソフトを逆解析して自分達のパラメータだと称するものがあるので、わからないようにしている。でも、そうするとsofxやhsdxをHSPiPだけで作るのはとても大変な作業になる。そこで、自分の実験値を入れたテーブルとY-MB23で計算した推算値を入れたマスターテーブルを作っておき、CASをキーにして必要なsofxやhsdxを作成するソフトを作った。差し当たってはJavaScriptで作ってブラウザ上で動かす。CASは日付になるものは先頭に”_”をつける。

hsdxを作る場合にはCASとScoreのペアを用意してテキストエリアにペーストする。make

M	N	O
	CAS	Score
	110-09-1	1
	872-10-4	1
	75-09-2	1
	646-06-0	1
	70-90-1	1
	109-69-6	1
	67-66-1	1
	67-64-1	1
	68-21-2	1
	67-66-6	1
	123-10-3	1
	111-18-9	1
	110-10-9	1
	108-19-4	1
	108-18-2	1
	106-07-4	1
	71-21-3	1
	820-06-1	1
	73-01-3	1
	75-09-8	3
	80-78-3	3
	110-02-7	3
	108-20-3	3
	110-04-2	3
	64-17-0	3
	67-50-1	3
	7732-10-5	3

The screenshot shows the HSP file builder interface. It has two radio buttons: 'HSD Data' (unselected) and 'sof data' (selected). Below them, it says 'The first line is title "CAS"'. A text area contains the instruction 'Please paste CAS and Score'. A table lists CAS numbers and scores, with the first row highlighted. Below the table are 'Make file' and 'Close' buttons. The bottom part of the screenshot shows the generated XML code, including a header section with various parameters and an 'sof format' section with XML tags for the selected data row.

fileボタンを押せば、タブ区切りのフォーマットとxmlフォーマットのテキストを作成する。

sofxを作る時にはCASだけで良い。新しい溶媒や、溶媒のデータはエクセル上で管理する。変更したらプログラムに流し込んで変更する。プログラム作成はバグ潰しをChatGPTに助けてもらって、2時間仕事。マスターテーブルを作るのに3時間。まー1日仕事にしては、有用なプログラムが作成できた。

GreenSolvent24.sofxを作った。