

大学図書館員に向けた 情報技術の学びの 最初の一歩を考える

2022年5月21日

大学図書館研究会 第30回オープンカレッジ

田辺浩介（物質・材料研究機構）

TANABE.Kosuke@nims.go.jp

自己紹介

- 物質・材料研究機構 (NIMS) 統合型材料開発・情報基盤部門
材料データプラットフォームセンター 主任エンジニア
- 図書館システム・研究者総覧・機関リポジトリなどの
学術情報サービスの開発・運用に携わる

Materials Data Repository

The screenshot shows the MDR website interface. At the top left is the MDR logo and the text 'MATERIALS DATA REPOSITORY'. On the top right are navigation links: Home, About, Help, Contact, and Login. Below the navigation is a search bar with the text 'Search MDR' and 'Enter search terms', and a 'GO' button. The main content area is divided into a left sidebar and a main results area. The sidebar, titled 'Limit your search', contains several filter categories: Type of work (with a dropdown arrow), Collection, Keyword, Language, Publisher, Resource type, Visibility, Computational methods, Data origin, Properties addressed, and Synthesis and. The 'Type of work' filter is expanded to show 'Dataset' with a red 'x' and the number '1,925'. The main results area shows a filter bar with 'Filtering by: TYPE OF WORK > DATASET' and a 'Start Over' button. Below the filter bar is a pagination bar with '« Previous | 1 - 10 of 1,925 | Next »', a 'SORT BY RELEVANCE' button, a '10 PER PAGE' button, and icons for list, grid, and print views. The first result is titled 'HAXPES spectra of Au'. It includes a small thumbnail image of HAXPES spectra, a 'Description/Abstract' section stating 'This dataset consists of HAXPES spectra at Au measured at SPring-8 BL46XU, and is a part of HAXPES database (MDR HAXPES DB, https://doi.o...', a 'Keyword' section with 'Au, BL46XU, HAXPES, SPring-8, and collection - MDR HAXPES DB', a 'Material/Specimen' section with 'Au', a 'Resource Type' section with 'Dataset', a 'Data origin' section with 'experiments', a 'Contact person' section with 'Industrial Application and Partnership Division, JASRI', and a 'Date Uploaded' section with '05/04/2022'. Below this result is another result titled 'HAXPES spectra of Au Sputtered', which also includes a thumbnail and a 'Description/Abstract' section.

材料研究の論文や研究データを
公開するデータリポジトリ。
研究データはほぼ全件に
DataCite DOIを付与

<https://mdr.nims.go.jp>

みなさんに聞いてみたいこと

なぜ情報技術を学びたいと思うのか？

- つまり、何ができるようになりたいのか
- たとえば「デジタルアーカイブの構築」といっても、いくつかに分解できる
 - 調達の仕様書が書けるようになりたい
 - メタデータを効率よく書けるようになりたい
 - 登録するファイルの作成(ファイル形式の変換など)をしたい
 - 公開用ソフトウェアの導入や、導入後の修正作業をしたい
- おそらくこれは人それぞれ、始め方も人それぞれ

自分自身の始め方を
思い出してみる

きっかけ

- 大学は文学部
 - コンピュータにはなんとなく興味があったため、とりあえず「情報」と名前のつく専攻を選んだ
 - しかし、特にプログラミングの授業はなかった
- ある日、授業で「OPACの評価」というレポート課題が出た

「日本中の大学図書館の本が探せるのって すごいなあ」

NACSIS Webcat

総合目録データベースWWW検索サービス

Webcatは、学術研究利用のために供するものであり、営利のための利用はできません。
なお、Webcatで検索した資料について、図書館に利用を申し込む際には、各図書館で利用条件が異なる場合がありますので、あらかじめ電話等で御確認ください。
[\[利用の手引き\]](#) || [\[English version here\]](#)

全資料 図書 雑誌

タイトル :
著者名 :
出版者 :
出版年 :
標準番号 :
フリーワード:

[\[目録情報課ホームページ\]](#) || [\[学術情報センターホームページ\]](#)

NACSIS webcat: 簡略表示

[\[利用の手引き\]](#) || [\[検索画面に戻る\]](#)

該当件数は 17 件です

1. [WORLD CUP USA94への道](#) -- スポニチ・エンタープライズ, 1994
2. [すべてのサッカー場で逢おう : ワールドカップのスターたちと日本代表の光と影 / 加部究, 伯井寛著](#) -- パラス, 1994
3. [ぼくたちの盃\(ワールドカップ\) : サポーターが見た!フランスへの熱き軌跡 / 吉沢康一監修](#) -- 光文社, 1998
4. [グレート・ゴール : ワールドカップ・サッカー 1930~1986 \[ビデオ \(カセット\)\]](#) -- 文藝春秋, 19-- -- (Sports Graphic Number Video Version)
5. [サッカー世界のプレー : ワールドカップ西ドイツ大会 / 牛木素吉郎著](#) -- 講談社, 1975. -- (講談社スポーツシリーズ)
6. [サッカー世界のプレー : ワールドカップのスターたち / 牛木素吉郎構成 ; 岸本健写真](#) -- 講談社, 1970
7. [ザ・ワールドカップ : 強者たちの壮絶なる戦い / 永井透著](#) -- フットワーク出版, 1998
8. [スキー・ワールドカップのすべて / 志賀仁郎著](#) -- ベースボール・マガジン社, 1976
9. [マイウェイ・ミスキー : ワールドカップへの挑戦 / 海和俊宏著](#) -- 筑摩書房, 1987
10. [燃えつきるまで : ワールドカップ・サッカー・ストーリー / ビート・デイヴィス著 ; 大塚一志訳](#) -- 図書出版社, 1994

Webcat(総合目録データベースWWW検索サービス)

本運用開始. <https://www.nii.ac.jp/CNEWS/N44/04.html>

きっかけ

- **NACISIS Webcatを見て感動した**
 - 「日本中の大学図書館の本が探せるのってすごいなあ」
 - どうやって動いているんだろう
 - こういうサービスを自分で作ってみたいなあ

その後

- 就職先の大学の図書室に、図書管理システムがなかった
- 「今がチャンスだ」と開発を思い立った
- **蔵書検索のサンプルが載っているプログラミングの参考書を何回も読んだ**

実際に作っているもの

図書館システム
Next-L Enju
(フリーソフト
ウェアとして公開)



貸出と予約 ▾

開館日と催し物 ▾

資料の検索

検索語:

(条件を詳しく指定して検索)

163 件の資料があります。

新着資料

ピックアップ

件名:

タグ:

No
Image

POPSマニュアル・シリーズ

広島大学附属図書館 広島大学附属図書館

人気のあるタグ

(タグの一覧)

Next-L Enju Leaf 1.4.0, オープンソース統合図書館システム | [このシステムについて](#) | [不具合を報告する](#) | [マニュアル](#)

Developed by [Kosuke Tanabe](#) and [Project Next-L](#).

<https://enju.next-l.jp>

今に至ります

The screenshot displays the NIMS Digital Library website. At the top, it features the NIMS logo and the text "Library of Materials Science" and "材料科学専門図書館". Navigation menus include "図書館案内", "図書館への依頼", "所内イントラネット", and "ログイン". A search bar is prominently displayed with the text "資料の検索" and a description of the search capabilities. Below the search bar, there are tabs for "新着図書", "おすすめ", "NIMS ソフトウェア", "コレクション", and "便利なサイト". A list of new books is shown, including titles like "マイクロ・ナノ熱工学の進展" and "バイオミメティクス・エコミメティクス". On the right side, there is a vertical stack of database logos including NIMS Papers Database, NIMS Patents Database, SAMURAI, MDR, MatNavi, STAM, WEB OF SCIENCE, Journal Citation Reports, Essential Science Indicators, Scopus, SciVal, arXiv, SciFinder, Alloy Phase Diagrams, and ICSD.

機構の図書管理システム
Enjuをベースにしている

具体的な始め方

(再掲)なぜ情報技術を学びたいと思うのか？

- つまり、何ができるようになりたいのか
- たとえば「デジタルアーカイブの構築」といっても、いくつかに分解できる
 - 調達の仕様書が書けるようになりたい
 - メタデータを効率よく書けるようになりたい
 - 登録するファイルの作成(ファイル形式の変換など)をしたい
 - 公開用ソフトウェアの導入や、導入後の修正作業をしたい
- おそらくこれは人それぞれ、始め方も人それぞれ

その中で、あえてふたつ
キーワードを選ぶとしたら

1. テキストファイル

2. コマンドライン

1. テキストファイル

テキストファイルとは

- 乱暴に言えば、メモ帳で読み書きできるファイル
- いろいろな拡張子があるが、中身はすべてテキスト
 - .txt
 - .csv
 - .tsv
 - .html
 - .md
 - その他いろいろ

なぜテキストファイルが重要か

- どのようなソフトウェアでも読み書きできる
 - 図書館の重要な役目である「長期保存」にも役に立つ
- コマンドラインのツールを使って、さまざまな処理ができる

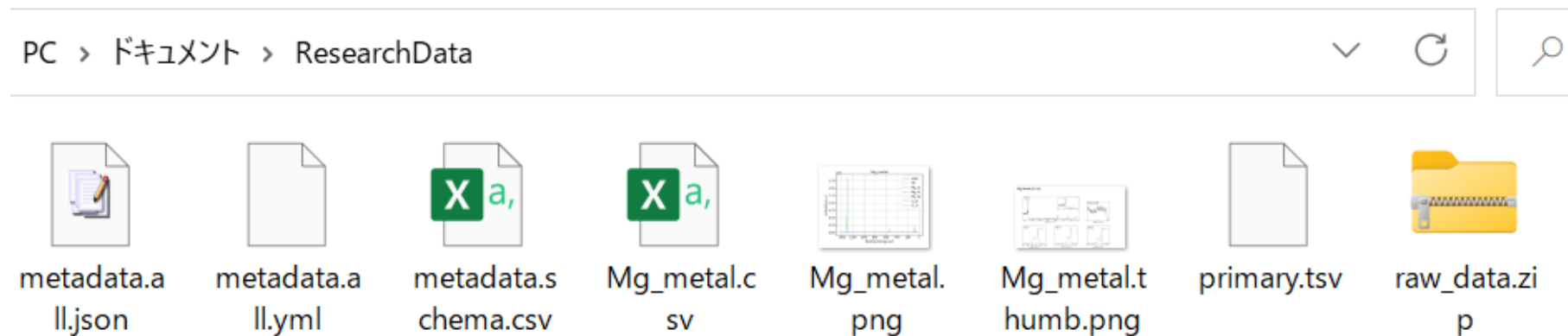
2. コマンドライン

なぜコマンドラインが重要か

- コンピュータを(マウスではなく)文字(テキスト)の入力で操作する方法
- テキストファイルと組み合わせることで、いろいろな操作をかんたんに自動化できる

例題

- あなたはデジタルアーカイブの管理者です
- デジタルアーカイブのサーバを新しくすることになり、
現在登録しているファイルを新しいサーバにコピーすることになりました
- ファイルの移行作業の際、ファイルが壊れたり改ざんされたり、紛失したりしていないことを確かめたいと思っています



入手元: HAXPES spectra of Mg. <https://mdr.nims.go.jp/concern/datasets/0v838351f?locale=en>

ファイルの「ハッシュ値」を取得

```
kosuke@ayumi: /mnt/c/Docur x + v
kosuke@ayumi:/mnt/c/Documents and Settings/kosuke/Documents$ find ./ResearchData/ -type f -print0 | xargs -0 sha512sum >
checksum.txt
kosuke@ayumi:/mnt/c/Documents and Settings/kosuke/Documents$ cat checksum.txt
43dc20c54d8ba56f9adedafe63c81c01b038c4b3f0136275e94f8ec193493c6ab9e6fee6270ff1bd5d30ab8421330230305988cae5c90291da9e752b
e0789738 ./ResearchData/metadata.all.json
c3b46e9a0266f5fa8f6d0dd78fe2c4e253badd00afde8c3ac8975b6c795b18d7e4691db028ae38c5fc403505358d426a70c59edd5e1dc2013b45edba
5c1d49ed ./ResearchData/metadata.all.yml
fc904be3750393e2671c9c209a58dddfeb3fdff9825bcf7841900ec55664c9d399a1279e3c8835da9b34d3eacaf2d7d7dfbe5649d4f7852c4bc05873
7079a1ef ./ResearchData/metadata.schema.csv
4dd53f4076a054c39a0d433e269dc26af59138500b192583377bcc48e3a4d034faf9259ed84699f837f18c3f41d8216ad129401a45818e93675b6fc9
e095a755 ./ResearchData/Mg_metal.csv
a2a954eb82996ee89725e1c0e3dd335b1b26acd8c0525633580c5d43e2bd890886e0080793e9bdb38b45e3d7bba7e4115a117399ec4ad670ac2107b2
c2767fa0 ./ResearchData/Mg_metal.png
a0d3f5a0b530d40c10f0705404ba750f970b261fe51903ad5d21a25d0e1f3a207d7f4f9f8b4a547367fe86eb3fd5edd9acd6f8999d616bc3dfbd2fed
3c0291c5 ./ResearchData/Mg_metal.thumb.png
acddd453a6f57d86fd596fd355b13e27d852f11693aebcc8843d287dad3aa0fe7bbf
3819dd89 ./ResearchData/primary.tsv
3e494bb414cb18bc1dd0e2d05da45b2352eec2e2f7ddaa372963fe4480bb8b8e195c
6a5dd48c ./ResearchData/raw_data.zip
kosuke@ayumi:/mnt/c/Documents and Settings/kosuke/Documents$ |
```

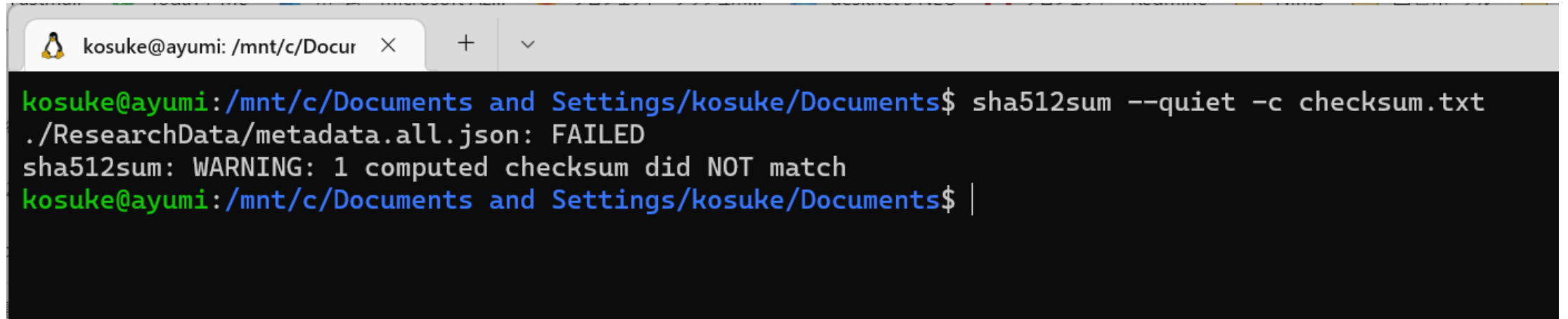
- 「ハッシュ値」は、ファイルの中身が同じなら、必ず同じ文字列になる
- 1文字でも違っていたら、全く違う文字列になる

ファイルを壊してみる

ファイル 編集 表示

```
{
  "@data_info@identifier@pid":"spring8.db7d7261-9ee3-464f
  "@data_info@identifier@register_name":"/SPring-8/BL46XU
  "ファイルの値を不正に書き換えてみる",
  "@data_info@identifier@proposal_number":"2019S0000",
  "@data_info@date@create_time":"2020-01-26 10:34:10",
  "@data_info@date@update_time":"2021-07-30 08:01:21",
  "@data_info@facility":"SPring-8",
  "@data_info@class_name":"BL46XU",
```


ファイルが壊れていることの検知



```
kosuke@ayumi: /mnt/c/Docur × + v
kosuke@ayumi:/mnt/c/Documents and Settings/kosuke/Documents$ sha512sum --quiet -c checksum.txt
./ResearchData/metadata.all.json: FAILED
sha512sum: WARNING: 1 computed checksum did NOT match
kosuke@ayumi:/mnt/c/Documents and Settings/kosuke/Documents$ |
```

使用したコマンドの説明

内容を分割してみる

```
find ./ResearchData -type f -print0 ¥  
| ¥  
xargs -0 sha512sum ¥  
> checksum.txt
```

```
find ./ResearchData -type f -print0
```

- "./ResearchData"ディレクトリ (フォルダ) を検索する (find ./researchdata)
- 検索対象は、種類がファイルであるもの (-type f)
- 検索結果のファイル名を出力する (-print0)

- |
- この文字よりも前に実行したコマンドの出力結果を、この文字の後のコマンドに引き渡す

```
xargs -0 sha512sum
```

- 受け取った出力結果 (ファイル名) の一覧を、後ろに記述したコマンドに渡す (xargs -0)
- 指定されたファイル名のファイルに対して、ハッシュ値を計算する (sha512sum)

```
> checksum.txt
```

- ファイルのハッシュ値の出力結果を、checksum.txt に書き出す

テキストファイルと
コマンドライン

よく言われる「DX」の前提とは

- 「コンピュータに作業をやらせる」こと
- コンピュータに作業をやってもらうには、その作業の「手順書」を書かなければならない
 - 一般に「(プログラムの)ソースコード」と呼ばれるもの
- ほとんどのコンピュータでは、この「手順書」はテキストファイルで書くようになっている

例題

- 研究者の論文一覧を作成したい
- タイトルや著者名、出版年、ジャーナル名を含める
- 手動で作ると時間がかかる上に入力の間違いも多いので、**論文のDOIを使って、書誌事項を自動で取得したい**

手順書の例

1. DOIのリストのファイルを開く
2. 最初の行に書いてあるDOIを読み込む
3. 読み込んだDOIをCrossrefで検索する
4. Crossrefの検索結果から、タイトルや著者名を取り出して表示する
5. ファイルの次の行に移動する
6. ファイルの終わりにたどりつくまで、2から5を繰り返す

<> doi.rb

```
1  #!/usr/bin/env ruby
2  require 'open-uri'
3  require 'json'
4  require 'csv'
5
6  # DOIから論文著者・タイトル・出版年・ジャーナル名・助成機関名を
7  # 取得するプログラム
8
9  File.open('doi1.txt').each do |line|
10   # ファイル"doi1.txt"から1行ずつDOIを読み込み、Crossref APIに問い合わせる
11   url = "https://api.crossref.org/works/#{line.strip}?mailto=nabeta@gmail.com"
12   data = JSON.parse(URI.open(url).read)
13
14   # 検索結果から各項目を取り出す
15   puts [
16     # 出版年
17     data.dig('message', 'published-print', 'date-parts')&.first&.first,
18     # タイトル
19     data['message']['title'].first,
20     # ジャーナル名
21     data['message']['container-title'].first,
22     # 著者名
23     data['message']['author'].map{|author| "#{author['given']} #{author['family']}" }.join(', '),
24     # 助成機関名
25     data['message']['funder']&.first&.dig('name')
26     # これらの項目をcsvとして出力する
27   ].to_csv
28
29   # Crossref APIにひんばんにアクセスしないよう、次のアクセスまで1秒待つ
30   sleep 1
31 end
```

前ページの「手順書」を
「Ruby」というプログラミング
言語で書いたもの

出力例

	A	B	C	D
1	2022	Development of Co-lean (Sm,Y)(Fe,Co,Ti) ₁₂ compo	Applied Physics Express	Ministry of Education, Culture, Sports, Science and Technology
2	2022	Friction in clay-bearing faults increases with the ionic radius of int	Communications Earth & Environment	MEXT Japan Society for the Promotion of Science
3	2022	Influence of interatomic potential and simulation procedures on th	Journal of Non-Crystalline Solids	
4	2022	Coordination State Control of Citric Acid Molecules on Europium(I	ACS Applied Nano Materials	Japan Society for the Promotion of Science
5		Facile Fabrication of N-Type Flexible CoSb ₃ -xTex Skutterudite/PI	Polymers	Japan Science and Technology Agency
6		Spin Polarization of Mn Could Enhance Grain Boundary Sliding in	Materials	National Natural Science Foundation of China
7	2022	Orbital Gating Driven by Giant Stark Effect in Tunneling Phototr	Advanced Materials	National Research Foundation of Korea
8	2022	Distribution coefficients for rare-earth doping in Y ₃ Ga ₅ O ₁₂ garnet	Journal of Crystal Growth	
9	2022	Fabrication, microstructure, and mechanical properties of all-oxide	Journal of Materials Science	Japan Society for the Promotion of Science
10	2022	Rational synthesis of ruthenium-based metallo-supramolecular pc	Applied Catalysis B: Environmental	National Institute for Materials Science
11				

応用: 研究者総覧SAMURAI

The screenshot displays the SAMURAI NIMS Researchers Directory Service interface. At the top, the logo and name 'SAMURAI NIMS Researchers Directory Service' are visible, along with 'ログイン' and 'English' buttons. A navigation bar contains links for 'SAMURAIについて', '研究紹介アルバム', '使い方', and 'NIMS著作物@NIMS Library'. The breadcrumb trail reads 'HOME > プロフィール > 田邊 浩介'. The main content area features a profile for 田邊 浩介 (TANABE, Kosuke), including a photo, contact information (Email: TANABE.Kosuke@nims.go.jp, TEL: 029-860-4786, Address: 305-0044 茨城県つくば市並木1-1 [アクセス]), and links to GPG and GitHub. A blue callout box on the right states: '研究者の業績の情報を ORCIDやCrossrefなど 機構内外のデータベースから自動的に取得・更新'. The '研究内容' section lists keywords like 'FRBR, Digital library, Next-L Enju' and mentions involvement in SAMURAI and related systems. A preview of the SAMURAI website is shown at the bottom.

<https://samurai.nims.go.jp>



Pelin Tozman, Hossein Sepehri-Amin, Xin Tang, Tadakatsu Ohkubo, Kazuhiro Hono.
Development of Co-lean (Sm,Y)(Fe,Co,Ti)₁₂ compounds with large saturation magnetization. Applied Physics Express. **15** [4] (2022) 045505 [10.35848/1882-0786/ac5b34](https://doi.org/10.35848/1882-0786/ac5b34)

Hiroshi Sakuma, David A. Lockner, John Solum, Nicholas C. Davatzes. Friction in clay-bearing faults increases with the ionic radius of interlayer cations. Communications Earth & Environment. **3** [1] (2022) [10.1038/s43247-022-00444-3](https://doi.org/10.1038/s43247-022-00444-3)

Jayani Kalahe, Yohei Onodera, Yasuyuki Takimoto, Hiroyuki Hijiya, Madoka Miyatani, Shinji Kohara, Shingo Urata, Jincheng Du. Influence of interatomic simulation procedures on the structures and properties of sodium aluminosilicate glasses from molecular dynamics simulations. Journal of Non-Crystalline Solids. **588** (2022) 121639 [10.1016/j.jnoncrysol.2022.121639](https://doi.org/10.1016/j.jnoncrysol.2022.121639)

Takuya Kataoka, Takumi Hashimoto, Sadaki Samitsu, Zizhen Liu, Motohiro Tagaya. Coordination State Control of Citric Acid Molecules on Europium(III) Ion-Doped Hydroxyapatite Nanoparticles for Highly Efficient Photoluminescence toward Biomedical Applications. ACS Applied Nano Materials. **5** [2] (2022) 2305-2315 [10.1021/acsanm.1c03999](https://doi.org/10.1021/acsanm.1c03999)

機構の研究者全員の
新しい論文の情報が
毎日自動的に更新されます

ここまでのまとめ

- テキストファイルの中身は文字
- コンピュータは文字で操作できる(コマンドライン)
- つまり、テキストファイルでコンピュータを操作できる

- テキストファイルとコマンドラインは組み合わせで使える
 - 「ファイルの一覧」というテキストファイルと「ファイルのチェック」というコマンドラインの組み合わせ
- なにかの「一覧」のテキストファイルがあれば、コマンドラインで作業を自動化できる可能性が高くなる

最初の一歩

テキストファイルに慣れる

- 毎日の仕事で作るファイルのうち、
テキストファイルで作れそうなものを探してみよう
- 特に「一覧表」がテキストファイルで作れないか見てみよう
 - グラフを含まない・必要としないExcelファイルが狙い目
- ほかにも、毎日の作業メモや議事録を、テキストファイルで書いてみよう

練習問題

- みなさんの手元に、どんな「一覧表」がありますか
 - その表に書いてあるデータで、どのような業務をしていますか
 - その業務の中で、「似たような作業を繰り返している」と感じていることはありますか
-
- ある程度まとまったら、上記の「回答」を提出しましょう
 - 待っている間にコマンドラインの操作環境の準備をしておきましょう

コマンドラインの操作環境を準備する

- Unix系OSのコマンドラインの実行環境"bash"を用意してみよう
- Windowsの人は次のいずれかがおすすめ
 - Git for Windowsをインストールする
<https://gitforwindows.org/>
 - Microsoft StoreからLinuxをインストールする
- Macの人は最初からインストール済み
 - 「ターミナル」を開く
 - (最近のMacはzshがデフォルトですが、bashも最初から入っています)

提出先(?)

- 大図研メーリングリスト
- Code4Lib JAPANのSlackチャンネル
 - <https://wiki.code4lib.jp/wiki/Slack>
- システムに詳しい同僚
- システム担当の部署

- 「こんなことがやりたい」「わからない」ということを、誰かに伝えてみるのが、ほんとうの最初の一步です