



STN 新プラットフォームの クロスオーバー検索



1

本日の内容

1. クロスオーバー検索
2. 索引情報を利用するクロスオーバー検索
 - 2-1. 物質 → 文献のクロスオーバー
 - 2-2. 文献 → 物質のクロスオーバー
3. 共通の情報の抽出・検索による
クロスオーバー検索



クロスオーバー検索とは

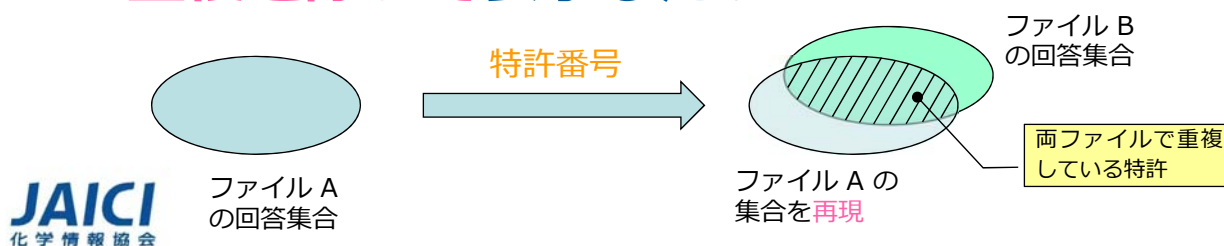
- ファイル間で共通する**特有の情報**を用いた、**複数のファイルを横断した検索**のこと
 - 物質 - 文献ファイルを組み合わせた検索
 - 複数の文献ファイルを用いた網羅的な検索
- **特有の情報の例**
 - 化学物質の番号 (CAS 登録番号など), 名称
 - 特許番号 など

クロスオーバー検索の利用例

- ある**物質**の合成法について書かれた**文献**を調べたい



- 複数の**特許**ファイルの検索結果を**重複を除いて表示**したい



クロスオーバー検索の方法

- **クロスオーバー検索の方法は 3 種類**
 - 物質，文献ファイルの**索引情報**を利用
 1. Get References/Get Substances ボタン
 2. REFX/SUBX コマンド
 - ファイル間で共通する**特有の情報**を検索
 3. Create Term List 機能

化学物質名称，
特許番号など

本日の内容

1. クロスオーバー検索
2. **索引情報を利用するクロスオーバー検索**
 - 2-1. 物質 → 文献のクロスオーバー
 - 2-2. 文献 → 物質のクロスオーバー
3. 共通の情報の抽出・検索による
クロスオーバー検索



物質から文献へのクロスオーバー

REGISTRY ファイル

1. CAS Registry Number: **145632-68-4**

CA Index Name: 2H-Azepin-2-one, hexahydro-7-(2-methylpropyl)-
 Molecular Formula: C₁₀H₁₉N O
 References in CAplus: 5

Get References
REFX コマンド

CAplus ファイル

2. A convenient synthesis of medium-sized lactams through RCM reaction of oxoxazolidinones

Author: Kamimura, Akio
 Tanaka, Keiichi
 Hayashi, Takahiro
 Omata, Yoji

Corporate Source: Yamaguchi University, Department of Applied Chemistry, Faculty of Engineering, Ube 755-8611, Japan (JP)

Source: Tetrahedron Letters (20060529), 47(22), pp. 3625-3627. In English
 CODEN: TELEAY ISSN: 0040-4039

Substance Indexing 26

102539-72-0P
 105-00-2P
145632-68-4P
 679-66-5P
 896718-19-7P
 896718-25-5P
 896718-29-9P
 896718-31-3P

特定の化学物質はCAS登録番号で索引

共通の情報：
CAS登録番号

Get References (1)

- **Get References** ボタンでクロスオーバー検索を実行できる

REGISTRY (1)

Get References

1. Guanidine, N-cyano-N'-[2-[[[(4-methyl-1H-imidazol-5-yl)]...]

シメチジンの物質レコード (REGISTRY)

ボタンをクリックするだけ!

CAplus (7,179)

Get Substances

1. Patent 2016 **Irisin-related cancer treatments**
 In one embodiment, the invention provides methods of treating of a subject in need thereof a ...
 Substance Indexing
51481-61-9, Cimetidine
 Notes: irisin-related cancer treatments
 Roles: Biological Study (BIOL),Therapeutic Use (THU),Uses

2. Patent 2016 **Edible cat toy**
 An edible toy for cats is made from a base ma...
 modified cellulose; hypromellose; an edible, ...
 Substance Indexing
51481-61-9, Cimetidine
 Notes: Edible cat toy
 Roles: Biological Study (BIOL),Biological Use, Unclassified (B

シメチジンが索引された文献レコード (CAplus)

ファイルは自動選択!

Get References (2)

- 選択した回答のみのクロスオーバー

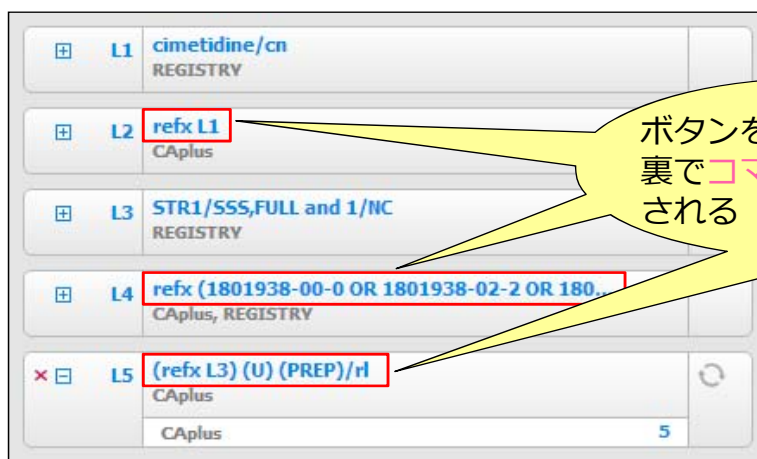
- CAS ロールによる限定

Get References ボタンでクロスオーバー可能なファイル

- 化学物質データベースで
Get References ボタンをクリックすると、自動的に文献データベースでクロスオーバー検索が実行される

化学物質データベース	文献データベース
REGISTRY, MARPAT	CAplus
DCR, DWPIM	DWPI
REAXYSFILESUB	REAXYSFILEBIB

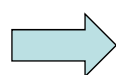
Get References によるクロスオーバー検索



→ **REFX コマンドを手入力しても検索できます！**

コマンドによるクロスオーバー検索では・・・

- ボタンで定義されたファイル以外のファイルにもクロスオーバー可能
- 複数ファイルからまたは複数ファイルへのクロスオーバー検索を実行できる
- CAS ロールやキーワードとの組み合わせなど細かい検索が可能



これらの検索を行うには、**コマンド**によるクロスオーバー

REFX コマンド (物質 → 文献)

- 物質ファイルと文献ファイルのマルチファイル環境に入り、REFX コマンドを実行する

REFX L#

L#は物質ファイルの回答集合

– CPlus で CAS ロールで限定する場合

REFX L# (U) ロールコード/RL

(U) 演算子

Get References および REFX コマンドでクロスオーバー可能なファイル

- Get References によるクロスオーバー

化学物質データベース	文献データベース
REGISTRY, MARPAT	CPlus
DCR, DWPIM	DWPI
REAXYSFILESUB	REAXYSFILEBIB

- REFX コマンドによるクロスオーバー

化学物質データベース	文献データベース
REGISTRY	CPlus, BIOSIS, EMBASE, MEDLINE
MARPAT	CPlus
DCR, DWPIM	DWPI
REAXYSFILESUB	REAXYSFILEBIB

文献から物質へのクロスオーバー

The screenshot illustrates a cross-over search process. On the left, a 'REGISTRY ファイル' (Registry File) shows a search result for '1. CAS Registry Number: 145632-68-4'. The entry includes the 'CA Index Name', 'Molecular Formula', and 'References in CAplus'. A chemical structure is displayed below. A yellow arrow labeled 'Get Substances' and 'SUBX コマンド' points from this entry to the right. On the right, a 'CAplus ファイル' (CAplus File) shows a search result for '2. A convenient synthesis of medium-sized lactams through RCM reaction of oxoxazolones'. The entry includes 'Author', 'Corporate Source', and 'Source' information. A yellow arrow labeled 'Get References' and 'REFX コマンド' points from this entry back to the left. A yellow box highlights the CAS number '145632-68-4P' in the 'Substance Indexing' section of the right file, with a callout stating '特定の化学物質はCAS登録番号で索引' (Specific chemical substances are indexed by CAS registration number).

Get Substances

- **Get Substances** ボタンでクロスオーバー検索
を実行できる
- 選択した回答のみのクロスオーバーも可能

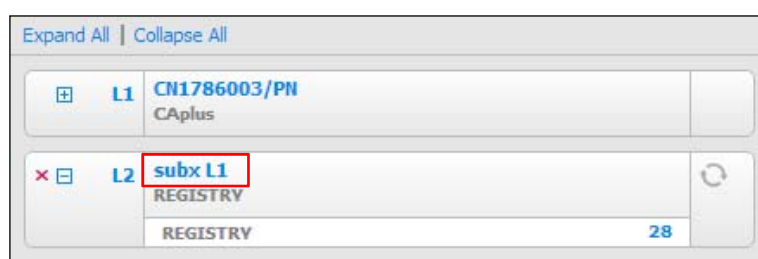
The screenshot shows two panels of the database interface. The left panel, labeled 'L1', shows a search result for '1. Patent 2001 Preparation of naphthothiazole'. A yellow box highlights the 'Get Substances' button. A callout bubble points to this button with the text '文献から物質の検索もボタンのクリックだけ!' (Searching for substances from literature is also just a button click!). The right panel, labeled 'L2', shows the search result for '1. 1H-Benzo[de]thiazolo[4,5-h]isoquinoline-1,3(2H)-dione, 2-[2-(dimethylamino)C18H17N3O3S'. A chemical structure is displayed, and a callout bubble points to it with the text '物質レコード (REGISTRY)' (Substance Record (REGISTRY)).

SUBX コマンド (文献 → 物質)

- 文献ファイルと物質ファイルのマルチファイル環境に入り、SUBX コマンドを実行する

SUBX L#

L#は文献ファイルの回答集合



Get Substances および SUBX コマンドでクロスオーバー可能なファイル

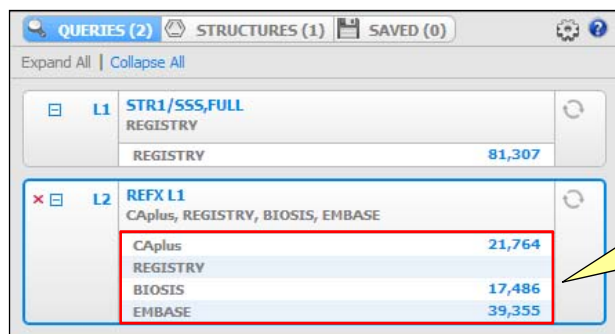
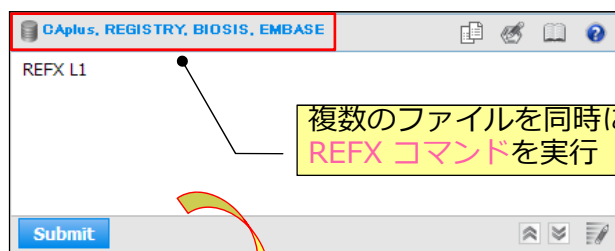
- Get Substances によるクロスオーバー

文献データベース	化学物質データベース
CAplus	REGISTRY
DWPI	DCR
REAXYSFILEBib	REAXYSFILESub

- SUBX コマンドによるクロスオーバー

文献データベース	化学物質データベース
CAplus, BIOSIS, EMBASE, MEDLINE	REGISTRY
CAplus	MARPAT
DWPI	DCR, DWPIM
REAXYSFILEBib	REAXYSFILESub

複数ファイルへのクロスオーバー検索



検索例 1

• クラリスロマイシンの合成文献の調査 (REGISTRY - CAplus ファイル)

- 検索の流れ

- 1) REGISTRY ファイルで化学物質名称検索を行う
 - Clarithromycin/CN
- 2) CAplus ファイルにクロスオーバー検索を行う
 - Get References ボタン
 - Preparation で限定

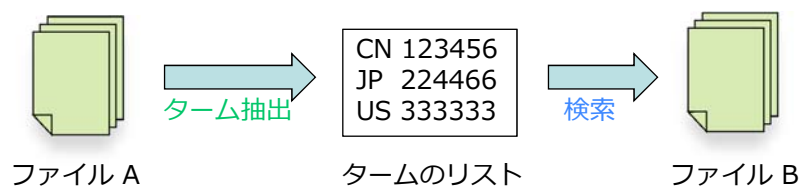
本日の内容

1. クロスオーバー検索
2. 索引情報を利用するクロスオーバー検索
 - 2-1. 物質 → 文献のクロスオーバー
 - 2-2. 文献 → 物質のクロスオーバー
3. 共通の情報の抽出・検索による
クロスオーバー検索



ファイル間で共通する情報による クロスオーバー検索

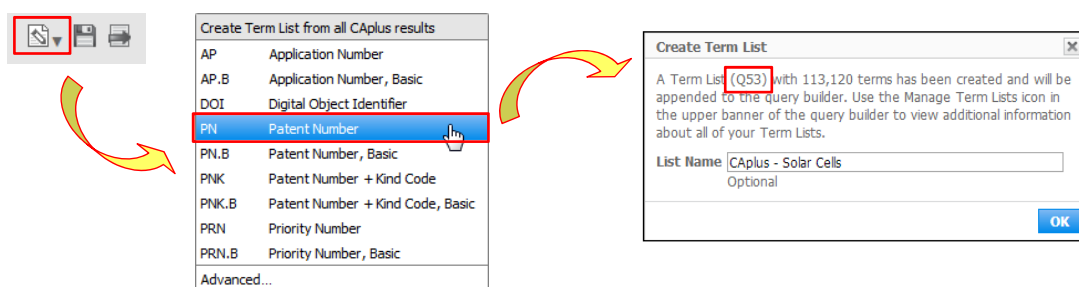
- ボタンやコマンドでのクロスオーバーができないファイルでは、**両ファイルで共通**、かつ**特有の情報**を用いて検索する
– 特許番号, レコード番号, 化学物質名称など



- タームの抽出は **Create Term List 機能**

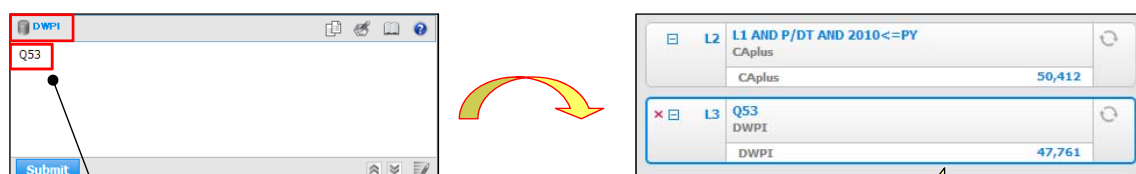
Create Term List 機能

- 指定したフィールドの情報を抽出してタームリストを作成する機能
 - タームリストには Q 番号が付与される
 - Q 番号は他のプロジェクトでも利用可能



Q 番号による検索

- ファイルを選択し, Q 番号で検索する
 - クエリーパネルで検索フィールドの変更も可能



Q 番号が作成されると,
自動でクエリーパネルに
入力される

DWPI ファイルで
CAplus ファイルの
回答を再現

Q 番号の内容の確認

- タームは Manage Term List で確認できる

Manage Term Lists

← Return

List name: CPlus - Solar Cells (Q53)
Term Count: 113,120

page 1 of 114

Source Field	Term
PN	AP2149
PN	AP2180
PN	AR48446
PN	AR48604
PN	AR50885
PN	AR51682
PN	AR53209
PN	AR54454
PN	AR54455
PN	AR54456
PN	AR54457
PN	AR54458
PN	AR54459
PN	AR54472
PN	AR54762
PN	AR54763
PN	AR55122
PN	AR56164
PN	AR57079
PN	AR57250
PN	AR57251
PN	AR57544
PN	AR58020

検索フィールドも一緒に抽出されている

Close

検索例 2

- 検索例 1 で得られた CPlus ファイルの回答に含まれる米国特許のクレームを USFULL ファイルで確認する

- 検索の流れ

- 1) CPlus ファイルで Create Term List により特許番号の抽出を行う
- 2) USFULL ファイルで、作成された Q 番号による検索を行う

まとめ

- **Get References** , **Get Substances** ボタンは
クリック操作のみで簡単に
クロスオーバー検索が可能
- **REFX, SUBX** コマンドを用いれば
様々なファイルへのクロスオーバー検索
も可能
- **Create Term List** 機能で情報を抽出し、
作成されたタームの集合 (Q 番号) を
他のファイルで検索できる

参考資料

- **STN 利用ガイド**
- http://www.jaici.or.jp/newstn/pdf/newSTN_guide.pdf
- **STN CA/REG**
- http://www.jaici.or.jp/newstn/pdf/newSTN_CAREG.pdf
- **STN 新プラットフォーム 日本語ヘルプ**
- <https://www.stn.org/help/stn/jp/index.htm>

