

STN[®]

CAplusSM Search Strategies in an Evolving Database

John Zabilski

Senior Application Specialist, CAS

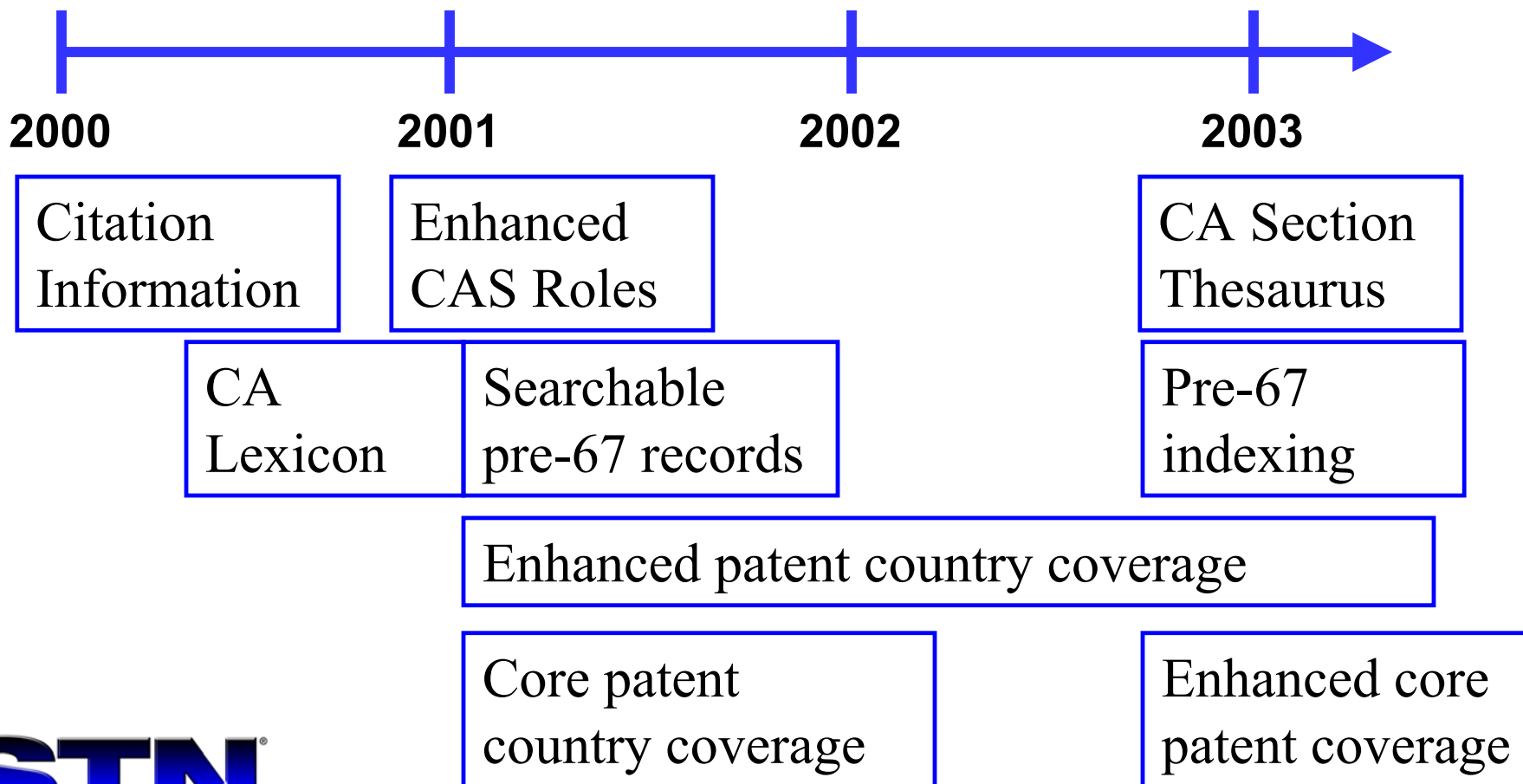
CAS is a division of the American Chemical Society

Session Agenda

- **Overview**
- Comprehensive substance searching
- Maximizing subject search results



Overview of CAplus enhancements



Session Agenda

- Overview
- **Comprehensive substance searching**
- Maximizing subject search results



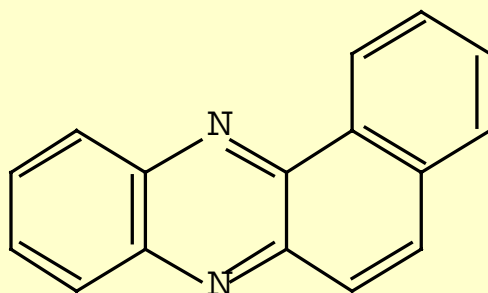
Substances may be represented many ways in CAPLUS

AN 1960:42299 CAPLUS
DN 54:42299
OREF 54:8296b-c
TI Electron spin resonance and bivalency of some dithiocarbamates of the coinage metals (Cu, Ag, Au)
SO Nature (1959), 184(Suppl. No. 4), 183-4
CC 3 (Electronic Phenomena and Spectra)
AB • • • benzene solns. of R_2NCS_2Au and $(R_2NCS_2)_2$ are mixed or if $(Et_2NCS_2)_3Au$ or $[(iso-Pr)_2NCS_2]_3Au$ are dissolved in benzene. It is concluded that bivalent Au is present in these solns.
IT Magnetic resonance absorption
(of dithiocarbamates of coinage metals, valency and)
IT ~~Carbamic acid, diisopropyldithio-, copper, Au and Ag salts~~
~~Carbanilic acid, N-methyldithio-, copper complex~~
IT 13681-87-3, Carbamic acid, diethyldithio-, copper salt
(magnetic resonance absorption of, valency and)

Substance indexing is different across the CAplus database

In this time period	Substances are indexed using	And refined using
1967–present	CAS RNs	CAS Roles
1907–1967	CAS RNs Nomenclature	Text terms

Search Question: Locate literature describing the preparation of benzo[a]phenazines such as



Search Strategy

To obtain comprehensive substance search results

Step 1 Conduct structure search.

Step 2 Identify name segments.

Step 3 Locate CAplus references.

Step 4 Locate unique CAOLDSM references.



Conduct a structure search:

=> FILE REGISTRY

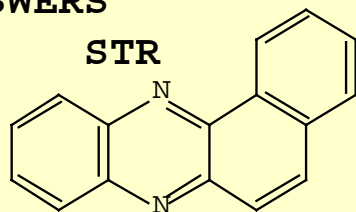
Uploading C:\Program Files\stnexp\Queries\benzophenazine.str

L1 STRUCTURE UPLOADED

=> D

L1 HAS NO ANSWERS

L1



Rings are not isolated.

=> S L1 SSS SAM

L2 50 SEA SSS SAM L1

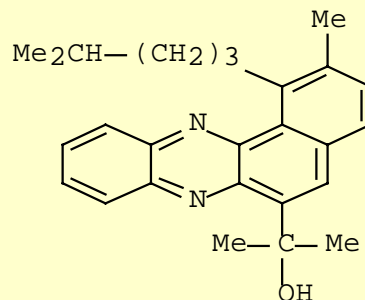
=> S L1 SSS FULL

L3 1864 SEA SSS FUL L1

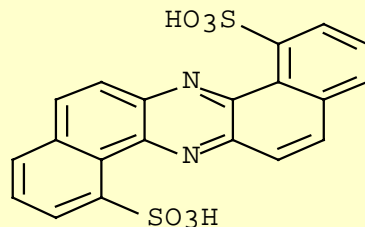
Identify name segments:

=> D SCAN L3

L3 1864 ANSWERS REGISTRY COPYRIGHT 2003 ACS on STN
IN Benzo[a]phenazine-6-methanol, $\alpha,\alpha,2$ -trimethyl-1-(4-
methylpentyl)- (9CI)
MF C26 H30 N2 O



L3 1864 ANSWERS REGISTRY COPYRIGHT 2003 ACS on STN
IN Dibenzo[a,h]phenazine-1,8-disulfonic acid (9CI)
MF C20 H12 N2 O6 S2
CI COM



Different strategies are needed for different portions of the file

- CAS RNs and CAS Roles for indexed records 1967–present
- CAS RNs and text terms for indexed records pre-1967
- Name segments and text terms for very new records and pre-1967 records

Locate CAPLUS references:

CAS RN indexed records, 1967–present

```
=> FILE CAPLUS
```

```
=> S L3/PREP
```

```
L4          291 L3/PREP  
              (L3 (L) PREP/RL)
```

CAS RN indexed records, pre-1967:

```
=> S L3 (S) (PREP? OR ISOL?) RANGE =,1966  
L5          270 L3 (S) (PREP? OR ISOL?)
```

Use RANGE to restrict answers by entry date into the file from the earliest record through 1966.

Locate CPlus references (cont'd):

Nomenclature-indexed records, pre-1967:

```
=> S ?BENZO (S) PHENAZIN? (S) (PREP? OR ISOL?) RANGE =,1966
L6          261 ?BENZO (S) PHENAZIN?(S)(PREP? OR ISOL?)
```

Combine results:

```
=> S L5 OR L6
L7          312 L5 OR L6
```

```
=> S L7 OR L4
L8          603 L7 OR L4
```

Most nomenclature records also contain CAS RNs. 42 unique answers were added.

```
=> S L7 NOT L5
```

Results include records indexed by CAS RN and CAS Role

=> D BIB HITSTR 12, 521

L8 ANSWER 12 OF 603 CAPLUS COPYRIGHT 2003 ACS on STN

AN 2002:323581 CAPLUS [Full-text](#)

DN 137:185468

TI Quantitative cascade condensations between o-phenylenediamines
and 1,2-dicarbonyl compounds without production of wastes •••

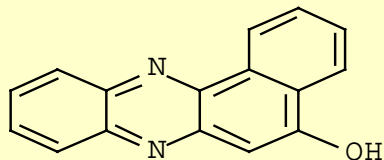
IT 2089-82-9P, Benzo[a]phenazin-5-ol

RL: SPN (Synthetic preparation); **PREP (Preparation)**

(quant. cascade condensations between o-phenylenediamines and
1,2-dicarbonyl compds. without production of wastes)

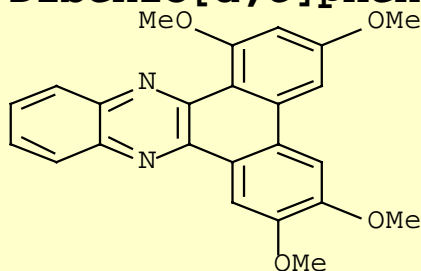
RN 2089-82-9 CAPLUS

CN Benzo[a]phenazin-5-ol (7CI, 8CI, 9CI) (CA INDEX NAME)



Results include pre-1967 records indexed by CAS RN, with text descriptions

L8 ANSWER 521 OF 603 CAPLUS COPYRIGHT 2003 ACS on STN
AN 1955:4947 CAPLUS [Full-text](#)
DN 49:4947
OREF 49:1076b-f
TI Alkaloids of Menispermaceous plants. CVIII. Alkaloids of
Stephania japonica. XVIII. On protostephanine. 8 • • •
IT 119013-93-3, **Dibenzo[a,c]phenazine,**
1,3,6,7-tetramethoxy-
(preparation of)
RN 119013-93-3 CAPLUS
CN Dibenzo[a,c]phenazine, 1,3,6,7-tetramethoxy- (6CI)(CA INDEX NAME)



Results include pre-1967 records indexed by nomenclature, without HIT indexing

=> D ALL 312

L8 ANSWER 312 OF 603 CAPLUS COPYRIGHT 2003 ACS on STN

AN 1966:61130 CAPLUS [Full-text](#)

DN 64.61130

OREF 64:11491a-b

TI Peroxidase action. XIII. The oxidation of 2-naphthylamine . . .
AB cf. CA 53, 6135°. The oxidation of 2-naphthylamine by the peroxidase + H₂O₂ system was examined. The reaction is complex, but two oxidation products were **isolated**, namely, **dibenzo(a,h)-phenazine** and a compound that appears to be 5-(2-naphthylamino)-**dibenzo[a,h]-phenazine**.

IT Peroxidases
(2 naphthylamine oxidn. by)

IT **Dibenzo[a,h]phenazin**
(formation from 2-naphthylamine)

IT 91-59-8, 2-Naphthylamine
(oxidn. of, by peroxidase)

The indexing contains only the German spelling, but alternative spellings are in the abstract.

Locate unique CAOLD references:

=> FILE REGISTRY

=> S L3 AND CAOLD/LC NOT CAPLUS/LC
L9 47 L3 AND CAOLD/LC NOT CAPLUS/LC

=> FILE HCAOLD

=> S L9
L10 31 L9

=> TRANSFER L8 1- OREF /AN
L11 TRANSFER L8 1- OREF : 2093 TERMS
L12 488 L14/AN

=> S L10 NOT L12
L13 10 L10 NOT L12

Substances unique to CAOLD can be identified in the REGISTRYSM file.

Use TRANSFER to find the references in CPlus that are also in CAOLD. OREF refers to a page location, so >1 record may be obtained for each number.

Display CAOLD references:

=> D 6 ALL HITSTR

L13 ANSWER 6 OF 10 HCAOLD COPYRIGHT 2003 ACS on STN

AN CA58:13947h CAOLD

TI azoaryls and phenazines from primary arylamines anions by autoxidn.

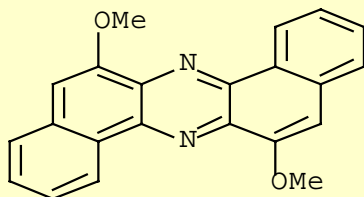
AU Horner, Leopold; Dehnert, J.

IT 226-47-1 501-58-6 501-60-0 584-90-7 613-55-8 1602-00-2
3272-91-1 3731-24-6 5692-66-0 6257-64-3 6319-23-9
7334-33-0 29418-44-8 40473-79-8 91391-90-1 92287-47-3
100436-08-6

IT 3731-24-6 100436-08-6

RN 3731-24-6 HCAOLD

CN Dibenzo[*a,h*]phenazine, 6,13-dimethoxy- (7CI, 8CI) (CA INDEX NAME)



...

Session Agenda

- Overview
- Comprehensive substance searching
- **Maximizing subject search results**



A strategy combining controlled terminology and free text will maximize subject search results

Search Question: Locate references on compositions used to remove stains from textiles.

Search Strategy

To obtain comprehensive subject search results

- Step 1** Develop a controlled term strategy using the CA Lexicon.
- Step 2** Develop a section code strategy using the CA section code thesaurus.
- Step 3** Develop a free-text strategy.
- Step 4** Combine results

Controlled term strategy – Identify terminology for *stain removers*

```
=> FILE CAPLUS
```

```
=> E STAIN REMOVERS/CT 5
```

E#	FREQUENCY	AT	TERM
E1	0	1	STAIN/CT
E2	2		STAIN NUMBER/CT
E3	0	2 -->	STAIN REMOVERS/CT
E4	0	2	STAIN REMOVERS DETE
E5	0	1	STAIN-RESISTANT/CT

The CA Lexicon defines stain removers as a linked term (qualifier) for the more general term detergents. It can be used to increase search precision.

```
=> E E3+ALL
```

```
E1 0 --> Stain removers/CT
E2 USE Detergents (L) stain removers/CT
***** END
```

Additional controlled terms used pre-1967 have been added to the CA Lexicon .

```
=> E DETERGENTS+ALL/CT
```

```
E1 11054 BT2 Materials/CT
E2 100386 BT1 Surfactants/CT
E3 41185 --> Detergents/CT
HNTD Valid heading during volume 66
(1967)to present.
E4 11759 OLD Cleaning compositions/CT
E5 UF AI 'fiya/CT
```

```
• • •
```

Develop a controlled term strategy:

=> S E4

L1 11759 "CLEANING COMPOSITIONS"/CT

=> S (L1 OR DETERGENTS/CT) (L) (STAIN# OR SPOT#) (L) REMOV?

L2 593 (L1 OR DETERGENTS/CT)(L)(STAIN# OR SPOT#)(L)REMOV?

Identify terminology for *textiles*:

=> E TEXTILES+ALL/CT

E1 69748 --> Textiles/CT

HNTE Valid heading during volume 66 (1967) to present.

NOTE Fabrics made from natural or synthetic fibers, or their blends, are indexed here.

E2 37 OLD Rags/CT

E3 UF Cloth/CT

E4 UF Fabrics/CT

E5 304 NT1 Canvas/CT

• • •

=> S L2 AND (TEXTILE# OR FABRIC#)

L3 330 L2 AND (TEXTILE# OR FABRIC#)

Free-text terms are used to obtain all records including pre-1967.

Evaluate results with D SCAN:

L3 330 ANSWERS CAPLUS COPYRIGHT 2003 ACS on STN
IC D06L001-04
CC 46-6 (Surface Active Agents and Detergents)
Section cross-reference(s): 40
TI "Semitsvet-2" composition for removing oil and
stains from the surface of **textile** products • •
IT Amides, uses and miscellaneous
(C10-13 fatty, N-hydroxyethyl, stain-removing compns. contg.,
for **textiles**)
IT **Detergents**
(**cleaning compns.** for **removal** of oil,
grease, rust and color **stains** from **textiles**)
IT 84070-05-3
RL: USES (Uses)
(stain-removing compn., for **textiles**)• • •
HOW MANY MORE ANSWERS DO YOU WISH TO SCAN? (1):1

D SCAN includes reference to CA section codes – an additional access point to relevant literature.

With D SCAN, additional relevant search terms may be observed.

D SCAN (cont'd):

Relevant literature may appear in a number of CA sections.

L3 330 ANSWERS CAPLUS COPYRIGHT 2003 ACS on STN

CC 40-8 (Textiles and Fibers)

TI Hydrocarbons as solvents for stain removal during

ST **textile** stain removal hydrocarbon solvent; spot remover hydrocarbon
textile finishing

IT Hydrocarbon oils

RL: NUU (Other use, unclassified); PRP (Properties); USES (Uses)

(Krebs 580 and Krebs 580A; hydrocarbon solvents as replacement for chlorinated hydrocarbons in stain removal from **textiles**)

IT **Textiles**

(hydrocarbon solvents as replacement for chlorinated hydrocarbons in stain removal from **textiles**) • • •

IT **Detergents**

(**spot removers**, hydrocarbon solvents as replacement for chlorinated hydrocarbons in **stain removal** from **textiles**)

IT **Detergents**

(**stain removers**, hydrocarbon solvents as replacement for chlorinated hydrocarbons in **stain removal** from **textiles**)

HOW MANY MORE ANSWERS DO YOU WISH TO SCAN? (1):0

CAplus records are categorized into sections

- One section code is assigned for main content
- Cross section codes (/SX) are assigned for secondary concepts starting in 1972
- Content of sections have changed over time to reflect changes in chemistry
- The CA section code thesaurus identifies changes in numbers, titles, and content

Section code strategy – Identify codes for *textiles*

=> E TEXTILES+ALL/CC

```
E1      184614    -->  TEXTILES/CC
E2      61602     USE  39 TEXTILES, 1967-1981/CC
E3      67822     USE  40 TEXTILES AND FIBERS, 1987 TO PRESENT/CC
E4      17081     USE  40 TEXTILES, 1982-1986/CC
E5      9319      USE  47 TEXTILES, 1963-1966/CC
E6      2214      USE  48 TEXTILES, 1962 ONLY/CC
***** END *****
```

EXPAND on the term identified in D SCAN to see the complete history.

=> E E3+ALL

```
E1      2160492   BT1  MACROMOLECULAR/CC
E2      67822    -->  40 TEXTILES AND FIBERS, 1987 TO PRESENT/CC
          NOTE THIS SECTION INCLUDES ANALYSIS, PREPARATION,
          MANUFACTURE, TESTING, PROCESSING, AND COMPOSITION OF
          NATURAL AND SYNTHETIC FIBERS AND THE CHEMICALS USED
          IN THEIR MANUFACTURE AND PROCESSING. FIBERS FOR USE
          AS REINFORCING AGENTS ARE INCLUDED IN THE
          APPROPRIATE SECTION ENCOMPASSING THE MATERIAL BEING
          REINFORCED. DYE SYNTHESIS AND COLOR-STRUCTURE
          RELATIONSHIP ARE INCLUDED IN SECTION 41. DRY
          CLEANING AND LAUNDERING ARE INCLUDED IN SECTION 46.
```

•••

Section codes (cont'd):

HNTE SUBJECT WAS COVERED IN SOMEWHAT BROADER SECTIONS PRIOR TO 1962. THESE OLD BROADER SECTIONS CAN BE DISPLAYED USING THE +OLD RELATIONSHIP CODE.

E3	58596	OLD	25 DYES AND TEXTILE CHEMISTRY, 1911-1946/CC
E4	2413	OLD	25 DYES AND TEXTILE FABRICS, 1907-1909/CC
E5	21325	OLD	25 DYES AND TEXTILES CHEMISTRY, 1947-1959/CC
E6	5251	OLD	25 DYES AND TEXTILES, 1960-1961/CC
E7	792	OLD	26 DYES AND TEXTILE FABRICS, 1910 ONLY/CC
E8	61602	OLD	39 TEXTILES, 1967-1981/CC
E9	17081	OLD	40 TEXTILES, 1982-1986/CC
E10	9319	OLD	47 TEXTILES, 1963-1966/CC
E11	2214	OLD	48 TEXTILES, 1962 ONLY/CC
E12	0	NT1	40-0 TEXTILES AND FIBERS, 1987 TO PRESENT, GENERAL/CC
E13	0	NT1	40-1 TEXTILES AND FIBERS, 1987 TO PRESENT, GENERAL/CC
E14	0	NT1	40-10 TEXTILES AND FIBERS, 1987 TO PRESENT, NONWOVEN TEXTILES AND OTHER USES/CC
E15	0	NT1	40-2 TEXTILES AND FIBERS, 1987 TO PRESENT, PREPARATION OF FIBERS/CC
• • •			
E22	0	NT1	40-9 TEXTILES AND FIBERS, 1987 TO PRESENT, TEXTILE FINISHING PROCESSES AND PROCESSING MATERIALS/CC

***** END *****

It is common that section code numbers have changed over time.

Develop a section code strategy:

=> S E2+OLD

L4 246419 "40 TEXTILES AND FIBERS, 1987 TO PRESENT"+OLD/CC
 (10 TERMS)

=> S E2/SX OR E8/SX OR E9/SX

L5 34685 "40 TEXTILES AND FIBERS, 1987 TO PRESENT"/SX OR "39
 TEXTILES, 1967-1981"/SX OR "40 TEXTILES, 1982-
 1986"/SX

=> S L4 OR L5

L6 281104 L4 OR L5

Identify codes for *detergents*:

=> E DETERGENTS+ALL/CC

E1 76498 --> DETERGENTS/CC

E2 54050 USE 46 SURFACE ACTIVE AGENTS AND DETERGENTS, 1967 TO
PRESENT/CC

***** END *****

=> E E2+OLD

E1 54050 --> 46 SURFACE ACTIVE AGENTS
PRESENT/CC

E2 660 OLD 27 FATS, FATTY OILS, AND SOAP, 1907-1909/CC

E3 2147 OLD 27 FATS, FATTY OILS, AND SOAP, 1911-1916/CC

E4 2252 OLD 27 FATS, FATTY OILS, AND SOAPS, 1917-1922/CC

E5 3165 OLD 27 FATS, FATTY OILS, WAXES, AND DETERGENTS,
1936-1939/CC

E6 16568 OLD 27 FATS, FATTY OILS, WAXES, AND DETERGENTS,
1945-1961/CC

E7 7864 OLD 27 FATS, FATTY OILS, WAXES, AND SOAPS, 1923-1935/CC

E8 2569 OLD 27 FATS, FATTY OILS, WAXES, AND SOAPS, 1940-1944/CC

E9 240 OLD 28 FATS, FATTY OILS, AND SOAP, 1910 ONLY/CC

E10 583 OLD 42 SURFACE-ACTIVE AGENTS AND DETERGENTS, 1962 ONLY/CC

E11 2132 OLD 53 SURFACE-ACTIVE AGENTS AND DETERGENTS, 1963-1966/CC

***** END *****

The OLD relationship code shows the different numbers that have been used for detergents over time.

Develop a section code strategy:

```
=> S E1+OLD OR E1/SX
```

```
L7 121294 "46 SURFACE ACTIVE AGENTS AND DETERGENTS, 1967 TO  
PRESENT"+OLD/CC OR "46 SURFACE ACTIVE AGENTS AND DETERGENTS,  
1967 TO PRESENT"/SX
```

Enhance the strategy with free text:

```
=> S (TEXTILE# OR L6) AND (L7 OR DETERGENT# OR CLEANING COMPOSITION#)  
AND (STAIN# OR SPOT#)(S)REMOV?
```

```
L8 579 (TEXTILE# OR L6) AND (L7 OR DETERGENT# OR CLEANING  
COMPOSITION#) AND (STAIN# OR SPOT#)(S)REMOV?
```

Develop a free-text strategy:

```
=> SET PLURALS ON; SET ABB ON  
SET COMMAND COMPLETED
```

```
=> S (DETERGENT OR CLEANING COMPOSITION)(S)(STAIN OR SPOT)(S)REMOV?  
AND (TEXTILE OR FABRIC)
```

```
L9          513 (DETERGENT OR CLEANING COMPOSITION)(S)(STAIN OR  
SPOT)(S)REMOV? AND (TEXTILE OR FABRIC)
```

Combine search strategies:

```
=> S L3 OR L8 OR L9
```

```
L10         780 L3 OR L8 OR L9
```

- L3 = controlled term strategy
- L8 = section code strategy
- L9 = free-text strategy

Results include records obtained with controlled terms

=> D PY HIT 55 2 753

L10 ANSWER 55 OF 780 CAPLUS COPYRIGHT 2003 ACS on STN
PY 2001

AB An alk., SDS-stable protease optimally active at pH 11
from a Bacillus sp. . . .

IT **Textiles**

(cotton, stains on; characterization and wash performance (at **stain removal**) of SDS-stable alk. protease from Bacillus sp. RGR-14)

IT **Detergents**

(laundry; characterization and wash performance (at **stain removal**) of SDS-stable alk. protease from Bacillus sp. RGR-14)

Results include records with controlled terms and current section codes

L10 ANSWER 2 OF 780 CAPLUS COPYRIGHT 2003 ACS on STN
PY 2003
CC **40-8 (Textiles and Fibers)**
AB Spun fiber is washed in a soln. contg. **stain remover**,
whitener, brightener, conditioner and/or odor **remover** (no
examples).
ST fiber cleaning soln **stain remover** whitener conditioner
odor remover; brightener **stain remover**
conditioner odor remover fiber cleaning soln
IT **Detergents**
(**stain removers**; spun fiber cleaning soln. contg.)

Results include records obtained with old section codes and controlled terms

```
L10 ANSWER 753 OF 780 CAPLUS COPYRIGHT 2003 ACS on STN
PY 1936
TI New methods for the removal of spots
CC 25 (Dyes and Textile Chemistry)
IT Cotton
    Linen
        (blood-stain removal from, chemicals for)
IT Cleaning compositions
    (for stain removal from fabrics, etc.)
IT Ink removers
    (for textiles)
IT Stains
    (removal of, from textiles)
IT Rust
    (removal of, from textiles, agent for)
IT Textiles
    (stain removal from)
```

The CAplus file has been enriched with significant new content

- Substance search strategies should take into account indexing available
 - CAS RNs and CAS Roles, 1967–present
 - CAS RNs and text descriptions, pre-1967 records
 - Nomenclature and text descriptions, pre-1967 records
- Online tools facilitate searching CAplus indexing
 - CA Lexicon
 - CA section code thesaurus



The logo for STN (Science and Technology Network) is displayed in large, bold, blue 3D letters. The letters have a slight shadow and a reflection below them. A registered trademark symbol (®) is located at the top right of the 'N'.

CAplus Search Strategies in an Evolving Database

Visit CAS at Stand 90