



STN[®]

Numeric Property Searching
in PCTFULL

Jeremias Gromotka
FIZ Karlsruhe

STN STN is available through FIZ Karlsruhe, Germany
and Chemical Abstracts Service, U.S.A.

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

Agenda

- PCTFULL Reload
- Numeric Property Search
- Numeric Property Search Examples

STN STN is available through FIZ Karlsruhe, Germany
and Chemical Abstracts Service, U.S.A.

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

The new PCTFULL – keeping pace with modern requirements

- On February 26th, the improved PCTFULL was launched, featuring
 - improved OCR-data
 - better **identification of claims** and descriptions
 - **improved bibliographical data** including priority numbers based on EPO-DOCDB
 - **updating** of backfile records (additions/corrections)
 - machine **translations** from several European languages into English
 - additional displayable languages (ZH, KO, RU)
 - searchable address data in IN, PA and AG

The PA field – offering more details

PA

POSITEC POWER TOOL (SUZHOU) CO., LTD, **Number 18 Dongwang Road, Industrial Park, Suzhou,**

Jiangsu 215123, CN, [NAT: CN, RES: CN], for all designated states except US;

HUO, Lixiang, Number 18 Dongwang Road, Industrial Park, Suzhou, Jiangsu 215123, CN, [NAT: CN, RES: CN], for US only

Address details searchable in field /PA.T

PAOR

苏州宝时得电动工具有限公司, 中国江苏省苏州市工业园区东旺路18号,

Jiangsu 215123, CN, [NAT: CN, RES: CN];

霍立祥, 中国江苏省苏州市工业园区东旺

路18号, Jiangsu 215123, CN, [NAT: CN, RES: CN]

The new PCTFULL – keeping pace with modern requirements

- On February 26th, the improved PCTFULL was launched, featuring

numeric property search

Numeric Property Search in PCTFULL

- Numeric Property Data is extracted from the full text and made available for searching by FIZ Karlsruhe for improved retrieval.
 - more than 30 physical properties in almost 400 units
 - automatic unit conversion
- Searchable property data is available in all English-language text fields (TIEN, ABEN, CLMEN and DETDEN)
 - proximity searching with numeric properties and keywords
 - flexible data input (open or closed ranges, tolerances, exact values)

Semantic Enrichment: Numbers and Units Behind the Scenes

- The task: make numeric property values and ranges searchable
 - identify numbers and units in the full-text
 - correctly recognize intervals

Semantic Enrichment: Numbers and Units Identifying Measurements

The resulting CeO_2 particle size measured by x-ray diffraction were in the range of 10 to 30 nm. Fig. 1 shows typical nano particles in a sample milled for 6 hours. In a second experiment a 1 litre attrition mill was used for milling the mixture. ... In addition it is widely accepted that the existence of a so-called 'limiting particle size' limits the practical minimum particle size that can be attained by grinding to values greater than 100nm, irrespective of the type of ball mill employed.

Semantic Enrichment: Numbers and Units Identifying Measurements

- The task: make numeric property values and ranges searchable
 - identify numbers and units in the full-text
 - correctly recognize intervals
 - extract and normalize to base units

Semantic Enrichment: Numbers and Units Normalisation

values are converted to SI base units

The resulting CeO₂ particle size measured by x-ray diffraction were in the range of 10 to 30 nm. Fig. 1 shows typical nano particles in a sample milled for 6 hours. In a second experiment a 1 litre attrition mill was used for milling the mixture. ... In addition it is widely accepted that the existence of a so-called 'limiting particle size' limits the practical minimum particle size that can be attained by grinding to values greater than 100nm, irrespective of the type of ball mill employed.

Semantic Enrichment: Numbers and Units Normalisation

values are converted to SI base units

10 to 30 nm

1.0×10^{-8} m, 3.0×10^{-8} m

6 hours

2.16×10^4 sec.

1 litre

1.0×10^{-3} m³

greater than 100nm

1.0×10^{-7} m, ∞

Note that for technical reasons, 100nm are included in this range, even though the text defines the size to be 'greater than 100nm'.

Semantic Enrichment: Numbers and Units Identifying Measurements

- The task: make numeric property values and ranges searchable
 - identify numbers and units in the full-text
 - correctly recognize intervals
 - extract and normalize to base units
 - make this information available for searching

Numeric Property Search in PCTFULL

- The generated numeric properties are searchable in the (English language) full-text
 - combination with text-based search terms
 - standard proximity operators apply
 - searching of exact values and ranges

Numeric Property Search in PCTFULL Units

almost 400 different units are searchable
in PCTFULL

Bequerel bit byte candela degree Hertz
Joule Kelvin kg/m³ kilogram kg/s lumen
Lux m² m/s meter m²/s m³ mol mol/l
N/m Ohm Pa·s Pascal percent pH rpm
second Siemens Tesla Volt Watt

Numeric Property Search Fields

Property	Field Code	Base Unit
AREA (SURFACE AREA)	/SAR	m ² (square meter)
BIT RATE	/BIR	Bit
BYTE RATE	/BYR	Byte
MOLAR CONCENTRATION (MOLARITY)	/CMOL	mol/L (mol per liter)
CONDUCTANCE (ELECTRICAL CONDUCTANCE)	/CON	S (Siemens)
DEGREE	/DEG	Degree
DENSITY (MASS DENSITY)	/DEN	kg/m ³ (kilogram per cubic meter)
VISCOSITY, DYNAMIC	/DV	PA s (pascal second)
ELECTRICAL IMPEDANCE/RESISTANCE	/RES	ohm
ENERGY	/ENE	J (Joule)
FORCE	/FOR	N (Newton)
FREQUENCY	/FRE	Hz (Hertz)

STN STN is available through FIZ Karlsruhe, Germany and Chemical Abstracts Service, U.S.A.

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

Numeric Property Search Fields

Property	Field Code	Base Unit
VISCOSITY, KINEMATIC	/KV	m ² /s (square meter per second)
LUMINOUS EMITTANCE/ILLUMINANCE	/LUME	Lux
LUMINOUS FLUX (LUMINOUS POWER)	/LUMF	Lumen
LUMINOUS INTENSITY	/LUMI	Candela
MASS	/M	Kg (kilogram)
MAGNETIC FIELD STRENGTH (MAGNETIC FLUX DENSITY)	/MFS	T (Tesla)
MASS FLOW (MASS TRANSFER)	/MFL	Kg/s (kilogram per second)
MOLECULAR WEIGHT (MOLAR MASS)	/MW	g/mol (gram per mol)
PERCENT (PROPORTIONALITY)	/PER	percent
PH VALUE	/PHV	pH
POWER	/POW	W (Watt)
PRESSURE	/PRES	Pa (Pascal)

STN STN is available through FIZ Karlsruhe, Germany and Chemical Abstracts Service, U.S.A.

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

Numeric Property Search Fields

Property	Field Code	Base Unit
RADIOACTIVITY	/RAD	Bq (Becquerel)
SPRING CONSTANT	/SCO	N/m (Newton per meter)
SIZE	/SIZ	m (meter)
SURFACE TENSION	/ST	J/m ² (Joule per square Meter)
TEMPERATURE	/TEMP	K (Kelvin)
TIME	/TIM	S (second)
VELOCITY	/VEL	m/s (meter per decond)
VELOCITY, ANGULAR	/VELA	rpm (rotations per minute)
VOLUME	/VOL	m ³ (cubic meter)
VOLTAGE	/VOLT	V (Volt)

For more details about searchable units see:
=> [HELP NPS](#) in PCTFULL

Automatic unit conversion

=> s 30-40 celsius/temp

L1 136689 30-40 CELSIUS/TEMP

=> s 30-40 c/temp
is also possible!

=> d kwic

L1 ANSWER 1 OF 136689 PCTFULL COPYRIGHT 2011 LNU on STN

DETDEN

temperature. The mixture was heated to 35 C for 3 h. . . .

L1 ANSWER 3 OF 136689 PCTFULL COPYRIGHT 2011 LNU on STN

DETDEN

formed hydrogels after incubating at 37 C for 24 hr as. . . .

L1 ANSWER 12 OF 136689 COPYRIGHT 2011 LNU on STN

DETDEN

to a temperature from about 50 degrees Fahrenheit to about 550
degrees Fahrenheit. . . .

L1 ANSWER 32 OF 136689 COPYRIGHT 2011 LNU on STN

DETDEN

temperature is within the range 263 Kelvin to 333 Kelvin.

Search example: Nanoscale

```
=> s size (3a) siz<=100nm
      756526 SIZE
      112228 SIZ<=100NM
L1    17287 SIZE (3A) SIZ<=100NM
```

```
=> d kwic 1-5
```

```
L1    ANSWER 1 OF 17287 PCTFULL COPYRIGHT 2011 LNU on STN
```

```
DETDEN . . .
particles with an average particle size of from approximately 3 to 100
nanometers, and a. . .
```

```
DETDEN . . .
30 to 300 or even 100 nanomet
```

```
DETDEN . . .
SiC powder with an average size of 90 nanometers was added and the. . .
```

```
DETDEN . . .
TiO2 powder having a particle size of about 50 nanometers is added to
finely. . .
```

For technical reasons, hitterm highlighting will typically only highlight the first number in a range. Units are not highlighted.

STN STN is available through FIZ Karlsruhe, Germany and Chemical Abstracts Service, U.S.A.

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

Search example: Nanoscale

```
=> s size (3a) siz<=100nm
      756526 SIZE
      112228 SIZ<=100NM
L1    17287 SIZE (3A) SIZ<=100NM
```

```
=> s size (3a) siz<=1.0e-7
      756526 SIZE
      112228 SIZ<=1.0E-7 M
L1    17287 SIZE (3A) SIZ<=1.0E-7 M
```

```
=> s size (3a) siz<=0.0000001
      756526 SIZE
      112228 SIZ<=0.0000001 M
L1    17287 SIZE (3A) SIZ<=0.0000001 M
```

Scientific notation is also supported.

The base unit for size is meter.

STN STN is available through FIZ Karlsruhe, Germany and Chemical Abstracts Service, U.S.A.

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

Search example: Green AlGaInP LEDs

```
=> s (light emitting diode or led) (s) AlGaInP (s) 500-570nm/siz
552948 LIGHT
```

Proximities work just as with regular text.

```
L2          16 (LIGHT EMITTING DIODE OR LED) (S) ALGAINP (S) 500-
570NM/SIZ
```

```
=> d kwic
```

```
L2          ANSWER 3 OF 16 PCTFULL COPYRIGHT 2011 LNU on STN
```

DETDEN

...The **LED** can be based on different materials, such as, without limitation, GaN, AlGaIn, InGaIn, AlInGaIn, AlInGaIn/AIN, AlInGaIn (emitting from 285 nm to 550nm), GaP, GaP:N, GaAsP, GaAsP:N, **AlGaInP** (emitting from 550nm to 660nm) SiC, GaAs, AlGaAs, BaN, InBaN, (emitting in near infrared and infrared). ...

```
L2          ANSWER 12 OF 16 PCTFULL COPYRIGHT 2011 LNU on STN
```

DETDEN

... A prior art class of **LEDs** based on GaAsP and **AlGaInP** semiconductor technology are capable of emitting at the PMMA attenuation windows of 510 and 570 nm. ...

STN STN is available through FIZ Karlsruhe, Germany
and Chemical Abstracts Service, U.S.A.

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

Search example: Elastomeric Block Co-Polymers

```
=> s ((?block?(T)?polymer? or ?block?(W)?polymer?) (S)
?elastomer?/clm and (?polymer? or ?block?)/clm (S) MW>75000)
```

```
L3          13 ((?BLOCK?(T)?POLYMER? OR ?BLOCK?(W)?POLYMER?)/CLM...
```

Searching in the claims of a patent full-text requires adjusted search strategies.

```
=> d kwic
```

```
L3          ANSWER 1 OF 13 PCTFULL COPYRIGHT
```

Note: In PCTFULL, the (S)-operator will retrieve hits within the same paragraph or claim.

CLMEN. . .

of Claim 1 wherein the **elastomeric block copolymer** is a styrene-butadiene-styrene **tn-block copolymer** having less than 1% by weight of styrene-butadiene **dt-block** based on the weight of the **elastomeric block copolymer**. 22 to 26% by weight. . . the weight of the **elastomeric block copolymer**; less than 10% by weight. . . the one or more polybutadiene **blocks**, and a molecular weight of about 100,000 to 150,000 g/mol

STN STN is available through FIZ Karlsruhe, Germany
and Chemical Abstracts Service, U.S.A.

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

Search example: Elastomeric Block Co-Polymers

L3 ANSWER 4 OF 13 PCTFULL COPYRIGHT 2011 LNU on STN

CLMEN. . .
50 weight percent of a **block copolymer** comprising: a. 5 - 98. . . and
β; and b. an **elastomeric** block; wherein said copolymerizable. . .

CLMEN. . .
of claim 1, wherein said **block copolymer** has a molecular weight from
30,000 to 500,000 g/mol.

CLMEN. . .
claim 1, wherein said **block copolymer** has a molecu
50,000 to 200,000 g/mol.

In these answers, the
required information
is spread out over
several claims.

L3 ANSWER 6 OF 13 PCTFULL COPYRIGHT 2011 LNU on STN

CLMEN
6. Thermoplastic **elastomeric** material according to claim 5, wherein the
styrene-based thermoplastic **elastomer** (a) is selected from **block
copolymers** having the following formulae: A(BA)_m. . .

CLMEN. . .
material according to any one of claims 6 to 11, wherein the **blocks** A and
A' have a. . . in the range of from 3,000 g/mol to 125,000 g/mol.

STN STN is available through FIZ Karlsruhe, Germany
and Chemical Abstracts Service, U.S.A.

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

Search Example: Alloys

=> s alloy (2s) (bismuth (1a) 40-60/percent (s) lead (1a) percent>20
(s) tin (s) cadmium)

L4 17 ALLOY (2S) (BISMUTH (1A) 40 PERCENT - 60 PERCENT /PERCENT (S)
LEAD (1A) PERCENT>20 PERCENT (S) TIN (S) CADMIUM)

=> d kwic

Percentages are also searchable.

L4 ANSWER 1 OF 17 PCTFULL COPYRIGHT 2011 LNU on STN

DETDEN . . .
temperature curing processes is used to cover the fasteners. One example
of such an **alloy** is a eutectic **alloy**, made of about 50% **bismuth**, about
26.7% **lead**, about 13.3% **tin**, and about 10% **cadmium** by weight, and with a
melting point of approximately 70 C (158 F). During the curing. . .

L4 ANSWER 2 OF 17 PCTFULL COPYRIGHT 2011 LNU on STN

DETDEN . . .
lowered clearly, if a such **alloy** contains approx. 14% to 60% **bismuth**,
20% to 30% **lead** or up to 45% **tin** or also antimony, **cadmium**, Indium,
zinc, tellurium, mercury or thallium. In particular with initially the
. . .

STN STN is available through FIZ Karlsruhe, Germany
and Chemical Abstracts Service, U.S.A.

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

Search Example: Formulations

```
=> s C11D0001-65+NT/IPC and (?phosphate? (5a) percent<0.5)
L5      4 C11D0001-65+NT/IPC AND (?PHOSPATE? (5A) PERCENT<0.5 PERCENT)
```

Percentages are very useful for searching formulations of pharmaceuticals, pesticides or household chemicals.

```
=> d kwic
```

```
L5      ANSWER 1 OF 4 PCTFULL COPYRIGHT 2011 LNU on STN
CLMEN. . .
```

STDC0076 (b) from 1wt% to 25wt% of an non-alkoxylated anionic deterative surfactant; STDC0040 (d) from 0wt% to 4wt% zeolite builder; STDC0046 (e) from 0wt% to 4wt% phosphate builder; and STDC0040 (f) from 0wt% to 10wt% silicate salt.

```
L5      ANSWER 3 OF 4 PCTFULL COPYRIGHT 2011 LNU on STN
DETDEN
```

The alkaline agent to provide a pH (when diluted to 0.25 to 5% by weight phosphates, metal carbonates, and metal hydroxides, particularly alkali or alkaline metal hydroxides such as NaOH; and KOH. The weight percentage. .

Numeric Property Search in PCTFULL Open Ranges

```
=> s volt>0.5
L6      116530 VOLT>0.5 V
```

```
=> d kwic
```

```
L6      ANSWER 1 OF 116530 PCTFULL COPYRIGHT 2011 LNU on STN
```

```
DETDEN . . .
the battery may be at 3.7 volts and be able to. . .
```

```
DETDEN . . .
that may receive power (e.g., +5 volts) from a digital device. . .
```

```
DETDEN . . .
example, the battery may provide 3.7 volts and the buck/boost power. . .
step up the voltage to +5 volts.
```

```
L6      ANSWER 2 OF 116530 PCTFULL (
```

Open ranges are truly open – and will thus retrieve irrelevant results.

```
DETDEN . . .
voltage is typically from about 2,000 to 80,000 volts. The charge. . .
```

```
DETDEN . . .
Biax tube. A voltage of 9.2 kV was employed. PTFE was. . .
```

Numeric Property Search in PCTFULL Open Ranges

```

=> s iron (2A) particle (2A) 1E-7<siz
      132395 IRON
      198830 PARTICLE
TERM '1E-7 M <SIZ' EXCEEDED TRUNCATION LIMITS - SEARCH ENDED

=> s iron (2A) particle (2A) 1E-7<siz<1E-3
      132395 IRON
      198830 PARTICLE
      350796 1E-7 M <SIZ<1E-3 M
L7      19 IRON (2A) PARTICLE

=> d kwic
L7      ANSWER 1 OF 19 PCTFULL COPYRIGHT 2011 LNU on STN
DETDEN . . .
      the red and the black iron oxide the particle size is 90 - 100 nm. . .

L7      ANSWER 5 OF 19 PCTFULL COPYRIGHT 2011 LNU on STN
CLMEN. . .
      the preceding claims wherein the iron containing colloidal particle is
      from 5 nm to 1 mm in. . .

```

Open ranges will often exceed truncation limits.

Restricting the range will allow the search to complete.

Technically, this is not a correct hit. In the Numeric Property Search feature, 'greater than' will always be interpreted as 'greater or equal'.

This range overlaps with the range searched.

STN STN is available through FIZ Karlsruhe, Germany
and Chemical Abstracts Service, U.S.A.

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

Numeric Property Search in PCTFULL Tolerances

```

=> s 100/temp
L8      5658 100 K /TEMP

=> s 100+-1/temp
L9      5780 100+-1 K /TEM

=> s 19 not 18
L10     122 L9 NOT L8

=> d kwic
L10     ANSWER 2 OF 122 PCTFULL COPYRIGHT 2011 LNU on STN
. . .171 -173 C. . .

L10     ANSWER 9 OF 122 PCTFULL COPYRIGHT 2011 LNU on STN
. . .ΔT= 99.78K

=> set tol temp=1
SET COMMAND COMPLETED

=> s 100/temp
L11     3252 100 +-1 K /TEMP

=> s 19 not 111
L12     0 L9 NOT L11

```

You can define tolerances for numeric properties directly in the search command.

Use the SET TOL command to automatically turn exact values into predefined ranges. Use SET TOL TEMP=1 PERM to save this setting for your Login ID.

STN STN is available through FIZ Karlsruhe, Germany
and Chemical Abstracts Service, U.S.A.

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

Numeric Property Search in PCTFULL Tolerances

```
=> set tol temp=1%
SET COMMAND COMPLETED

=> s 100/temp
L13      5780 100 +-1% K /TEMP

=> s l13 not l11
L14      0 L13 NOT L11

=> s 200/temp
L15      6831 200 +-1% K /TEMP
```

You can also define percent-based tolerances.

```
L15      ANSWER 1 OF 6831 PCTFULL C
DETDEN . . .
in 1,4- dioxane (0.4 mL) whilst
-75 C. The ice bath was. . . low

L15      ANSWER 2 OF 491 PCTFULL COPYRIGHT 2011 LNU ON STN
DETDEN . . .
mL of anhydrous tetrahydrofuran (THF) and cooled to -75 C. A
solution of 3-bromosalicyl alcohol isopropylidene acetal. . .
```

-75° Centigrade equals 198.15 Kelvin. These documents would not have been retrieved with an absolute tolerance of 1.

Numeric Property Search in PCTFULL Non-SI Search Units

```
=> s 1-5 inch/siz
L16      251985 1-5 INCH/SIZ

=> d kwic

L16      ANSWER 1 OF 251985 PCTFULL COPYRIGHT 2011 LNU ON STN
DETDEN . . .
surrounding that value. For example, a dimension disclosed as "40 mm" is
intended to mean "about 40 mm".

=> s 100 fahrenheit/temp
L17      43153 100 FAHRENHEIT +-1 /TEMP

=> d kwic
```

Some imperial units can also be used for searching.

```
L17      ANSWER 1 OF 43153 PCTFULL COPYE
DETDEN . . .
typically range from about 5 C to about 90 C. The water to fabric ratio
is typically. . .
```

The SET TOL TEMP setting remains active regardless of the unit chosen!

100° F ≈ 37.77°C, this is a correct hit!

Numeric Property Search in PCTFULL Search Options

=> S 50/VOL

L1 454 50 M**3 /VOL

searching with *default units*

=> S 50-60/VOL

L2 599 50 M**3 - 60 M**3 /VOL

searching with *intervals*

=> S 10-30 ML/VOL

L3 114646 10-30 ML/VOL

searching with *other units*

=> S SIZ < 5 MM

L4 463893 SIZ < 5 MM

searching with *open ranges*

=> S 5 MM +-1/SIZ

L5 160136 5 MM +-1/SIZ

searching with *tolerances*

=> S 5 MM +-5%/SIZ

L6 113784 5 MM +- 5%/SIZ

searching with *tolerances in %*

STN STN is available through FIZ Karlsruhe, Germany
and Chemical Abstracts Service, U.S.A.

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

Numeric Property Search in PCTFULL Availability

=> s mw/php (5a) (?block ?polym? or (?block? (T) ?polym?))

L18 163 MW/PHP (5A) (?BLOCK ?POLYM? OR (?BLOCK? (T) ?POLYM?))

=> d kwic 1-5

The availability of physical properties
can be researched in the /PHP-field.

DETDEN . . .

After quenching, the resultant iPF-block-PEP **diblock copolymer** had an Mn
= **122** kg/mol and MJMn = 1.20. . . .

DETDEN . . .

molecular weight of the Polyoxyalkylen **Blockcopolymers** **200000** g/mol,
prefers **100000** g/mol and particularly prefers 50000. . . .

=> s siz/php (s) (liposom? or (lipid? (W) vesicl?))

L19 9531 SIZ/PHP (S) (LIPOSOM? OR (LIPID? (W) VESICL?))

=> d kwic

DETDEN . . .

by entrapping the drug in **liposomes** or microemulsions which are
compatible. . . size in the range of **0.01** to 10 micrometers.

STN STN is available through FIZ Karlsruhe, Germany
and Chemical Abstracts Service, U.S.A.

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

The New PCTFULL Summary

- Higher quality of full-text
 - Better identification of claims and descriptions
 - Improved bibliographic data incl. priority numbers
 - More machine translations into English
 - Searchable address data => HELP RLO
- Numeric property searching in all English text fields => HELP NPS
 - 34 chemical and physical properties
 - Automatic unit conversion
 - Flexibility in data input (exact values, intervals, tolerances, units, orders of magnitude)

STN STN is available through FIZ Karlsruhe, Germany
and Chemical Abstracts Service, U.S.A.

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

Thank you for your attention!

STN[®]

CAS
E-mail: help@cas.org
Support and Training:
www.cas.org

FIZ Karlsruhe
helpdesk@fiz-karlsruhe.de
Support and Training:
www.stn-international.de

STN STN is available through FIZ Karlsruhe, Germany
and Chemical Abstracts Service, U.S.A.