

- IHTSDO delivering
- SNOMED CT®
- the global clinical terminology



# SNOMED CT Search & Data Entry

Jeremy Rogers

Fadi El-Turk

Anne Randorff Højen

# ***SNOMED CT***

## **Search candidate generation**

Drs Jeremy Rogers and Ed Cheetham

*IHTSDO Consultant Terminologists*

*Principal Terminology Specialists NHS HSCIC*

IHTSDO Showcase, Washington DC, October 11<sup>th</sup> 2013

# *Outline*

Symptoms

Pathology

Treatment

Evidence

An Experiment:

# 16 clinical search phrases into 3 SNOMED CT browser implementations

pain lower back

low back pain

traveller diarrhoea

acute asthma

breast lump

breast mass

medication review

vaccine given

enlarged node

morning headache

staging scan

urge incontinence

excision biopsy

post operative analgesia

transoesophageal ultrasound

emergency caesarean

# 'low back pain'

SNOMED\_2013a - SNOMED Browser

File Options Tools Help

Home F2 Centre F3 Back F4 Forward F5 Spawn F6 Subsets F7 ICD OPCS Mapped

low back pain More >>>

- PROCEDURE (2)
  - [UK] Low back pain clinical pathway
  - [UK] Lower back pain assessment
- CLINICAL FINDING (2)
  - Headache following lumbar puncture
  - Low back pain

20130925

CliniClue Xplore: Sctuk\_20130401 [Registered user: DemoUser060718@Clini

File Edit Language Layout Tools Help

Concept Id 138875005 SNOMED CT Concept (Type: = SNOMED CT Cond

Description Id 220309016

Words - any order 12

Find low back pain

- Clinical finding
  - low back pain
- Procedure
  - lower back pain assessment
  - low back pain clinical pathway

Hierarchy Subtype hierarchy

- SNOMED CT Concept
  - body structure

Apelon Mycroft - SNOMED CT [UK Edition] (SNOMED CT [UK Edition])

File Favorites Help

Tree Walker Search

Max Results: 100 Sort: Ascending Search

Search For: low back pain

Beneath: <ALL>

Search By: Synonym

Search On:

Search Results: 1 matches found

Low back pain (finding)

# 'pain lower back'

SNOMED\_2013a - SNOMED Browser

File Options Tools Help

Home F2 Centre F3 Back F4 Forward F5 Spawn F6 Subsets F7 ICD OPCS Maps

Search: pain lower back

- PROCEDURE (2)
  - [X] [UK] Low back pain clinical pathway
  - [X] [UK] Lower back pain assessment
- CLINICAL FINDING (6)
  - [X] Genitofemoral nerve neuralgia
  - [X] Iliohipogastric nerve neuralgia
  - [X] Ilioinguinal nerve neuralgia
  - [X] Intermittent cauda equina claudication
  - + Low back pain
  - [X] Lumbago-sciatica due to displacement of lumbar

20130925

CliniClue Xplore: Sctuk\_20130401 [Registered user: DemoUser060718@Clini

File Edit Language Layout Tools Help

Concept Id: 138875005  
Description Id: 220309016

SNOMED CT Concept (Type: = SNOMED CT Cond

Words - any order 1

Find: pain lower back

- Procedure
  - lower back pain assessment

Hierarchy: Subtype hierarchy

- SNOMED CT Concept
  - body structure

Apelon Mycroft - SNOMED CT [UK Edition] (SNOMED CT [UK Edition])

File Favorites Help

Tree Walker Search

Max Results: 100 Sort: Ascending Search

Search For: pain lower back

Beneath: <ALL>

Search By: Synonym

Search On:

Search Results: 0 matches found

# 'traveller diarrhoea'

The image displays three overlapping windows from SNOMED CT browsers, illustrating a search for the term 'traveller diarrhoea'.

- Left Window (SNOMED\_2013a - SNOMED Browser):** Shows a hierarchical tree view. The search term 'traveller diarrhoea' is entered in the search bar. The results list includes 'CLINICAL FINDING (1)' and 'Traveler's diarrhoea'.
- Middle Window (CliniClue Xplore):** Shows a detailed view of a SNOMED CT Concept. The Concept Id is 138875005 and the Description Id is 220309016. The search bar contains 'traveller diarrhoea'. The Hierarchy is set to 'Subtype hierarchy', showing 'SNOMED CT Concept' and 'body\_structure'.
- Right Window (Apelon Mycroft - SNOMED CT [UK Edition]):** Shows a search interface with the search term 'traveller diarrhoea'. The search results section indicates 'Search Results: 0 matches found'.

# 'traveler diarrhoea

The image displays three overlapping windows from SNOMED CT browsers, illustrating a search for 'traveler diarrhoea'.

- Left Window (SNOMED\_2013a - SNOMED Browser):** Shows a search for 'traveler diarrhoea'. The results list 'CLINICAL FINDING (1)' and 'Traveler's diarrhoea'.
- Middle Window (CliniClue Xplore):** Shows details for Concept Id 138875005 and Description Id 220309016. The search term 'traveler diarrhoea' is entered in the 'Find' field. The results list 'Clinical finding' and 'Traveler's diarrhoea'. The 'Hierarchy' is set to 'Subtype hierarchy'.
- Right Window (Apelon Mycroft - SNOMED CT [UK Edition]):** Shows a search for 'traveler diarrhoea' with 0 matches found. The search parameters are: Max Results: 100, Sort: Ascending, Search For: traveler diarrhoea, Beneath: <ALL>, Search By: Synonym, Search On: (empty).



# 'acute asthma'

SNOMED\_2013a - SNOMED Browser

File Options Tools Help

Home F2 Centre F3 Back F4 Forward F5 Spawn F6 Subsets F7 ICD OPCS Maps

acute asthma More >>>

- CLINICAL FINDING (6)
  - Acute asthma
  - Acute exacerbation of chronic asthmatic bronchi
  - Cardiac asthma
  - Exacerbation of asthma
  - Silo-fillers' disease
  - Simple pulmonary eosinophilia
- PROCEDURE (1)
  - [UK] Acute asthma control management

20130925

CliniClue Xplore: Sctuk\_20130401 [Registered user: DemoUser060718@Clini

File Edit Language Layout Tools Help

Concept Id 138875005  
Description Id 220309016

SNOMED CT Concept  
(Type: = SNOMED CT Cond

Words - any order 5

Find acute asthma

- Clinical finding
  - acute asthma
  - acute exacerbation of asthma
  - acute exacerbation of chronic asthmatic bronchitis
- Procedure
  - acute asthma control management

Hierarchy Subtype hierarchy

- SNOMED CT Concept
  - body\_structure

Apelon Mycroft - SNOMED CT [UK Edition] (SNOMED CT [UK Edition])

File Favorites Help

Tree Walker Search

Max Results: 100 Sort: Ascending Search

Search For: acute asthma

Beneath: <ALL>

Search By: Synonym

Search On:

Search Results: 1 matches found

Acute asthma (disorder)

# 'breast lump'

SNOMED\_2013a - SNOMED Browser

File Options Tools Help

Home F2 Centre F3 Back F4 Forward F5 Spawn F6 Subsets F7 ICD OPCS Maps

breast lump More >>>

- PROCEDURE (2)
  - Excisional biopsy of breast mass
  - Lumpectomy of breast
- SITUATION WITH EXPLICIT CONTEXT (3)
  - Breast lump present
  - No breast lump
  - O/E - breast lump not tethered
- CLINICAL FINDING (19)
  - Bovine ulcerative mammillitis
  - Breast haematoma
  - Breast lump
  - Cyst of breast
  - Infiltrating ductal carcinoma of breast, stage
  - Infiltrating ductal carcinoma of breast, stage
  - Infiltrating ductal carcinoma of breast, stage
  - Infiltrating ductal carcinoma of breast, stage
  - Malignant tumour of vertebral column
  - Melanocytic naevus of skin of breast
  - Neoplasm of breast
  - O/E - breast lump consistency
  - O/E - breast lump position
  - O/E - breast lump regularity
  - O/E - breast lump smooth
  - O/E - breast lump soft
  - O/E - breast lump tethering
  - Persistent breast nodularity
  - Sarcoma of bone and connective tissue
- BODY STRUCTURE (5)
  - Carcinoma resembling breast carcinoma
  - Juvenile carcinoma of the breast
  - Paget's disease and infiltrating duct carcinoma
  - Paget's disease and intraductal carcinoma of br
  - Paget's disease, mammary
- SPECIMEN (1)
  - Lumpectomy breast sample

20130925

CliniClue Xplore: Sctuk\_20130401 [Registered user: DemoUser060718@Clini

File Edit Language Layout Tools Help

Concept Id 138875005  
Description Id 220309016

SNOMED CT Concept  
(Type:= SNOMED CT Conc

Words - any order 98

Find breast lump

- Clinical finding
  - breast lump
    - O/E - breast lump cystic
    - O/E - breast lump consistency
    - O/E - breast lump smooth
    - O/E - breast lump soft
    - O/E - breast lump position
    - O/E - breast lump regularity
    - O/E - breast lump tethering
- Procedure
  - lumpectomy of breast
- Situation with explicit context
  - no breast lump
  - breast lump present
  - O/E - breast lump not tethered
- Specimen
  - lumpectomy breast sample

Hierarchy Subtype hierarchy

- SNOMED CT Concept
  - body structure

Apelon Mycroft - SNOMED CT [UK Edition] (SNOMED CT [UK Edition])

File Favorites Help

Tree Walker Search

Max Results: 100 Sort: Ascending Search

Search For: breast lump

Beneath: <ALL>

Search By: Synonym

Search On:

Search Results: 1 matches found

Breast lump (finding)

# 'breast mass'

SNOMED\_2013a - SNOMED Browser

File Options Tools Help

Home F2 Centre F3 Back F4 Forward F5 Spawn F6 Subsets F7 ICD OPCS Maps

breast mass More >>>

- PROCEDURE (5)
  - Excision of adenoma of thyroid
  - Excision of cyst of thyroid
  - Excisional biopsy of breast mass
  - Incisional biopsy of breast mass
  - Parathyroid adenoma excision
- CLINICAL FINDING (14)
  - Breast haematoma
  - Breast lump
  - Cyst of breast
  - Infiltrating ductal carcinoma of breast, stage
  - Infiltrating ductal carcinoma of breast, stage
  - Infiltrating ductal carcinoma of breast, stage
  - Infiltrating ductal carcinoma of breast, stage
  - Malignant tumour of vertebral column
  - Mammographic breast mass
  - Massive pubertal hypertrophy of breast
  - Melanocytic naevus of skin of breast
  - Neoplasm of breast
  - Persistent breast nodularity
  - Sarcoma of bone and connective tissue
- SPECIMEN (2)
  - Breast cyst fluid sample
  - Specimen from breast obtained by incisional biopsy
- SITUATION WITH EXPLICIT CONTEXT (1)
  - No breast lump
- BODY STRUCTURE (5)
  - Carcinoma resembling breast carcinoma
  - Juvenile carcinoma of the breast
  - Paget's disease and infiltrating duct carcinoma
  - Paget's disease and intraductal carcinoma of breast
  - Paget's disease, mammary

20130925

CliniClue Xplore: Sctuk\_20130401 [Registered user: DemoUser060718@CliniClue]

File Edit Language Layout Tools Help

Concept Id 138875005  
Description Id 220309016

SNOMED CT Concept (Type:= SNOMED CT Concept)

Words - any order 10

Find breast mass

- Clinical finding
  - breast mass
  - mammographic breast mass
  - massive pubertal hypertrophy of breast
- Procedure
  - excisional biopsy of breast mass
  - incisional biopsy of breast mass
- Specimen
  - specimen from breast obtained by incisional biopsy

Hierarchy Subtype hierarchy

- SNOMED CT Concept
  - body structure

Apelon Mycroft - SNOMED CT [UK Edition] (SNOMED CT [UK Edition])

File Favorites Help

Tree Walker Search

Max Results: 100 Sort: Ascending Search

Search For: breast mass

Beneath: <ALL>

Search By: Synonym

Search On:

Search Results: 1 matches found

Breast lump (finding)

An Experiment:

# 16 clinical search phrases into 3 SNOMED CT browser implementations

pain lower back	enlarged node	4+	5+	0
low back pain	morning headache	1	1	1
traveller diarrhoea	staging scan	7	0	0
acute asthma	urge incontinence	3	3	1
breast lump	excision biopsy	45+	22+	1
breast mass	post operative analgesia	1	0	0
medication review	transoesophageal ultrasound	2	1	0
vaccine given	emergency caesarean	3	3	0

# ***Outline***

Symptoms

**Pathology**

Treatment

Evidence

# The Problem:

## CodeForString(str) operation

### Many different strategies can be employed

All produce different false +ve and -ve results

At least one existing implementation for most elements

Very few implementations with all elements

Naïve approaches predominate

But clever strategies can be slow

and need extra lexical resources

Different strategies  $\Rightarrow$  different results

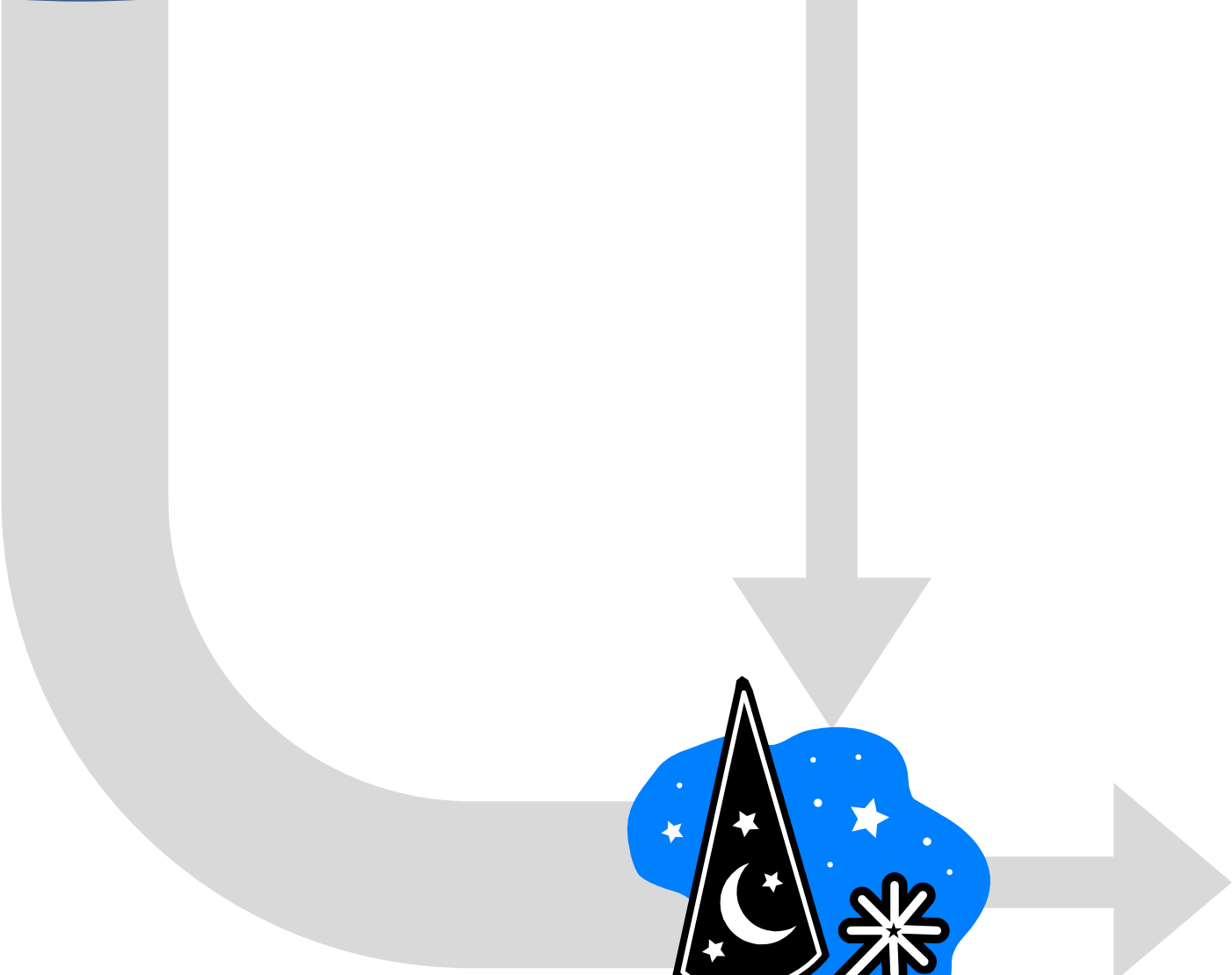
Different results  $\Rightarrow$  interrater variability

Interrater variability  $\Rightarrow$  reduced interoperability



**SNOMED CT**

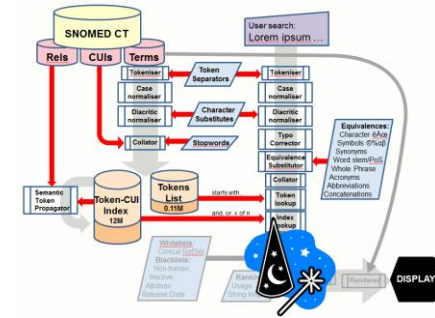
User search:  
Lorem ipsum ...



**DISPLAY**



# From term to concept...



## TOKENISATION

- Non WS token separators
- Character stripping

## TOKEN SET EXPANSION

- Word equivalence<sup>TIG 6.5.4</sup>
- Diacritic/Char substn
- Metaphone / SoundEx
- Levenshtein Distance
- Token concatenation

## TOKEN CONTRACTION

- Stopword lists
- Word stemming

## TOKEN LOOKUP

- Wildcard characters in corpus
- Exact matching
- Case sensitivity
- Force include/exclude tokens
- Query token proximity
- Query token (re-)ordering<sup>TIG</sup>
- Default wildcarding
- Multitoken semantics<sup>TIG</sup>

## Token inheritance

- Match scoring and ranking
- Match grouping & ordering

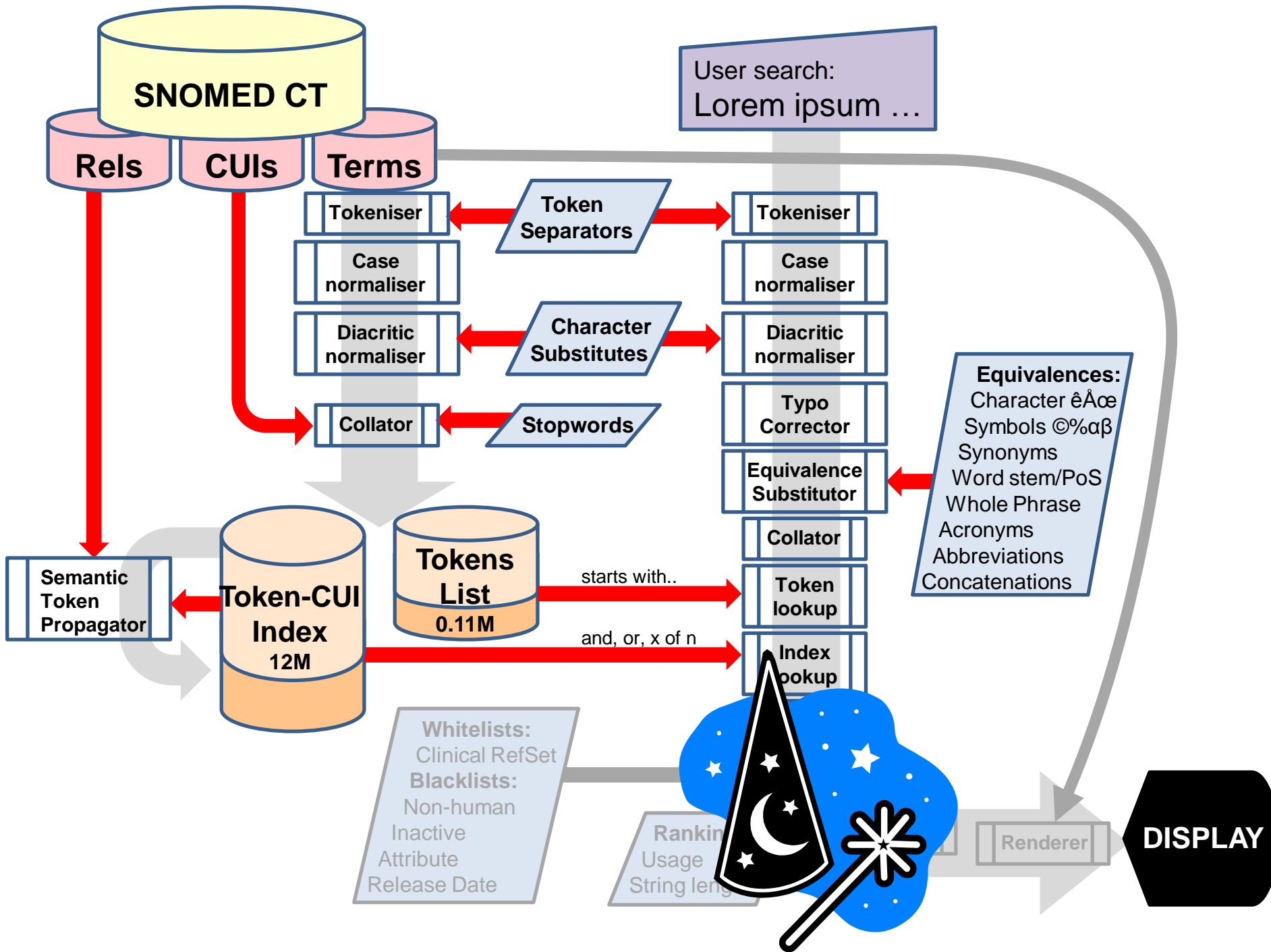
# ***Outline***

Symptoms

Pathology

**Treatment**

Evidence



# IHTSDO Lexical Resources

## Stopwords

- zres\_ExcludedWords\_en-US\_INT\_20070731.txt
- (n=53) Virtually unchanged since 2002

### IHTSDO Stopwords 2013

(2002 version)

ABOUT	CANNOT	MORE	SPECIFIC
ALONGSIDE	CHRONICALLY	MUST	THAN
AN	CONSISTS	NO	THAT
AND	COVERED	NOT	THE
ANYTHING	DOES	OF	THINGS
AROUND	DURING	ON	THIS
AS	EVERY	ONLY	THROUGHOUT
AT	FINDS	OR	TO
BECAUSE	FOR	PROPERLY	UP
BEFORE	FROM	SIDE	USING
BEING	IN	SIDED	USUALLY
BOTH	INSTEAD	SOME	WHEN
BY	INTO	SOMETHING	WHILE
			WITHOUT

DOWN?

WITH?

# mySQL 5.5 Stopword List

## mySQL Full Text Search Stopword List (2011)

a's	able	about	above	according	her	here	here's	hereafter	hereby	serious	seriously	seven	several	shall
accordingly	across	actually	after	afterwards	herein	hereupon	hers	herself	hi	she	should	shouldn't	since	six
again	against	ain't	all	allow	him	himself	his	hither	hopefully	so	some	somebody	somehow	someone
allows	almost	alone	along	already	how	howbeit	however	i'd	i'll	something	sometime	sometimes	somewhat	somewhere
also	although	always	am	among	i'm	i've	ie	if	ignored	soon	sorry	specified	specify	specifying
amongst	an	and	another	any	immediate	in	inasmuch	inc	indeed	still	sub	such	sup	sure
anybody	anyhow	anyone	anything	anyway	indicate	indicated	indicates	inner	insofar	t's	take	taken	tell	tends
anyways	anywhere	apart	appear	appreciate	instead	into	inward	is	isn't	th	than	thank	thanks	thanx
appropriate	are	aren't	around	as	it	it'd	it'll	it's	its	that	that's	thats	the	their
aside	ask	asking	associated	at	itself	just	keep	keeps	kept	theirs	them	themselves	then	thence
available	away	awfully	be	became	know	known	knows	last	lately	there	there's	thereafter	thereby	therefore
because	become	becomes	becoming	been	later	latter	latterly	least	less	therein	theres	thereupon	these	they
before	beforehand	behind	being	believe	lest	let	let's	like	liked	they'd	they'll	they're	they've	think
below	beside	besides	best	better	likely	little	look	looking	looks	third	this	thorough	thoroughly	those
between	beyond	both	brief	but	ltd	mainly	many	may	maybe	though	three	through	throughout	thru
by	c'mon	c's	came	can	me	mean	meanwhile	merely	might	thus	to	together	too	took
can't	cannot	cant	cause	causes	more	moreover	most	mostly	much	toward	towards	tried	tries	truly
certain	certainly	changes	clearly	co	must	my	myself	name	namely	try	trying	twice	two	un
com	come	comes	concerning	consequently	and	near	nearly	necessary	need	under	unfortunately	unless	unlikely	until
consider	considering	contain	containing						new	unto	up	upon	us	use
corresponding	could	couldn't	course						non	used	useful	uses	using	usually
definitely	described	despite	did						not	value	various	very	via	viz
different	do	does	doesn't						obviously	vs	want	wants	was	wasn't
don't	done	down	downwards						ok	way	we	we'd	we'll	we're
each	edu	eg	eight	either	okay	old	on	once	one	we've	welcome	well	went	were
else	elsewhere	enough	entirely	especially	ones	only	onto	or	other	weren't	what	what's	whatever	when
et	etc	even	ever	every	others	otherwise	ought	our	ours	whence	whenever	where	where's	whereafter
everybody	everyone	everything	everywhere	ex	ourselves	out	outside	over	overall	whereas	whereby	wherein	whereupon	wherever
exactly	example	except	far	few	own	particular	particularly	per	perhaps	whether	which	while	whither	who
fifth	first	five	followed	following	placed	please	plus	possible	presumably	who's	whoever	whole	whom	whose
follows	for	former	formerly	forth	probably	provides	que	quite	qv	why	will	willing	wish	with
four	from	further	furthermore	get	rather	rd	re	really	reasonably	within	without	won't	wonder	would
gets	getting	given	gives	go	regarding	regardless	regards	relatively	respectively	wouldn't	yes	yet	you	you'd
goes	going	gone	got	gotten	right	said	same	saw	say	you'll	you're	you've	your	yours
greetings	had	hadn't	happens	hardly	saying	says	second	secondly	see	yourself	yourselves	zero		
has	hasn't	have	haven't	having	seeing	seem	seemed	seeming	seems					
he	he's	hello	help	hence	seen	self	selves	sensible	sent					

1st, 2nd, 3rd Trimester ?

1st, 2nd 3rd degree burn ?

<b>ALONGSIDE</b>	<b>0</b>	<b>PROPERLY</b>	<b>0</b>
<b>CHRONICALLY</b>	<b>9</b>	<b>SIDE</b>	<b>238</b>
<b>CONSISTS</b>	<b>0</b>	<b>SIDED</b>	<b>73</b>
<b>COVERED</b>	<b>0</b>	<b>SPECIFIC</b>	<b>1356</b>
<b>FINDS</b>	<b>0</b>	<b>THINGS</b>	<b>26</b>

10 tokens in IHTSDO list, but not in mySQL list...

# Word Equivalence: 64 'tranquilisers'

## 206129006

- FSN:** Fetus or neonate affected by maternal tranquilizers during labor and delivery (disorder)  
**PREF:** Fetus or neonate affected by maternal tranquilizers during labor and delivery *enUS*  
**PREF:** Foetus or neonate affected by maternal tranquillisers during labour and delivery *enGB*  
**SYN:** Fetus or neonate affected by maternal tranquilizers during labor and delivery  
**SYN:** Fetus or neonate affected by maternal tranquillisers during labour and delivery

## 111766000

- FSN:** Poisoning by tranquilizer (disorder)  
**PREF:** Poisoning by tranquilizer *enUS*  
**PREF:** Poisoning by tranquilliser *enGB*  
*??? Tranquillizer ???*

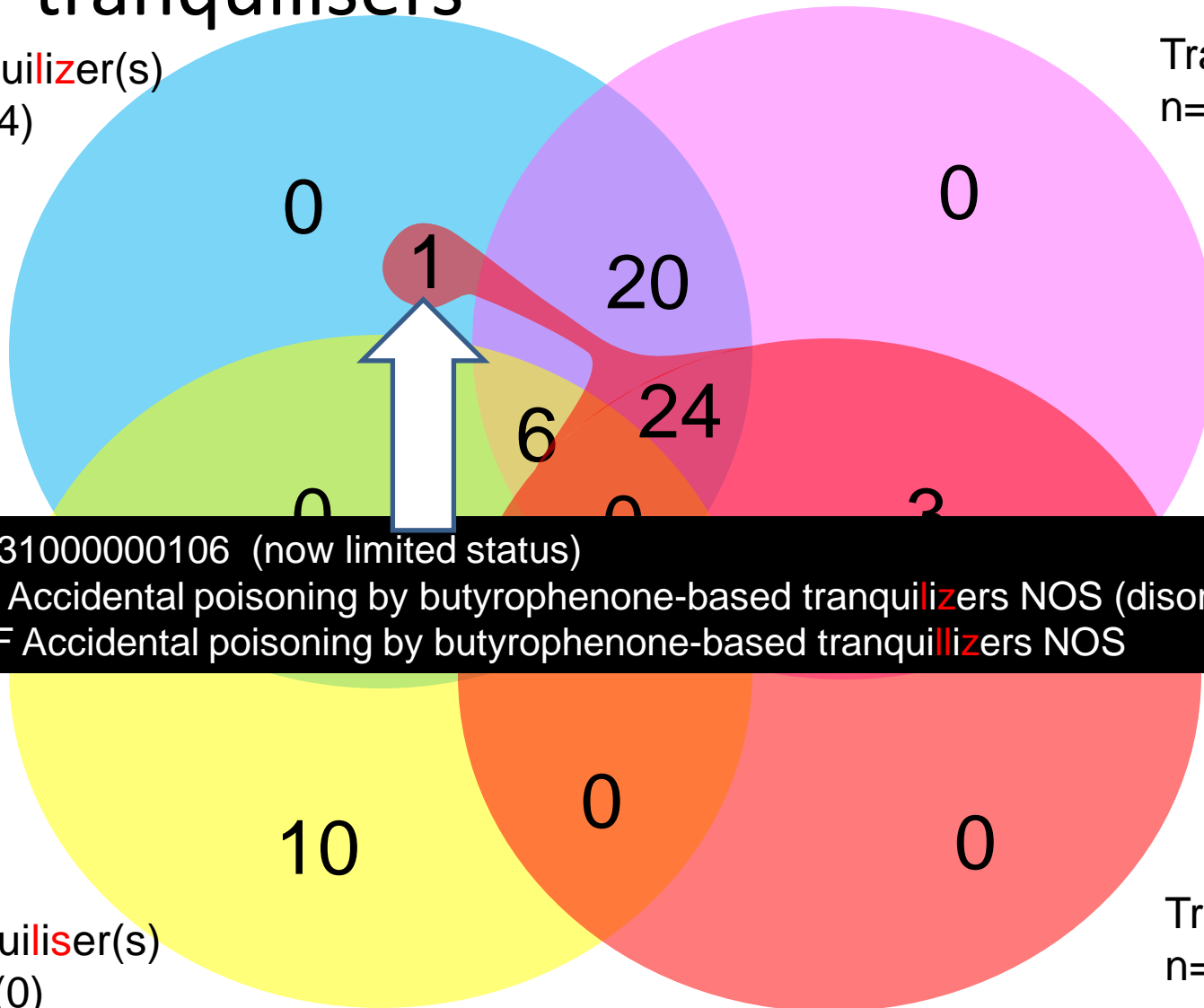
## 292343009

- FSN:** Benzodiazepine sedative adverse reaction (disorder)  
**PREF:** Benzodiazepine sedative adverse reaction  
**SYN:** Adverse reaction to benzodiazepine-based tranquilliser  
**SYN:** Adverse reaction to benzodiazepine-based tranquilizer  
*??? Tranquillizer ???*

# Word Equivalence: 64 'tranquilisers'

Tranquilizer(s)  
N=9 (4)

Tranquilliser(s)  
n=34 (26)



699731000000106 (now limited status)

FSN: Accidental poisoning by butyrophenone-based tranquilizers NOS (disorder)

PREF Accidental poisoning by butyrophenone-based tranquillizers NOS

Tranquilliser(s)  
n=11 (0)

Tranquilizer(s)  
n=16 (17)

# IHTSDO Lexical Resources

## Word and Phrase Equivalence

zres\_WordEquivalents\_en-US\_INT\_20020731.txt (n=4666)

res\_WordEquivalents\_en-GB\_GB1000000\_20060401.txt (n=71)

cicatrix	rupture	perineal	laryngeal	gastric	scrotal
scar	tear	perineum	larynx	stomach	scrotum
erythrocyte	bleeding	scapula	bronchial	pancreas	penile
red blood cell	haemorrhage	shoulder blade	bronchus	pancreatic	penis
gangrene	bruise	leucocyte	mediastinal	ileal	prostate
gangrenous	contusion	white blood cell	mediastinum	ileum	prostatic
avulsed	telangiectasia	platelet	mouth	duodenal	ovarian
avulsion	telangiectasis	thrombocyte	oral cavity	duodenum	ovaries
neurapraxia	cancer	thymic	palatal	rectal	ovary
neuropraxia	malignant neoplasm	thymus	palate	rectum	uterine
injury	malignant tumour	cardiac	roof of mouth	caecal	uterus
trauma	spinal	heart	teeth	caecum	adenohypophysis
agenesis	spine	spleen	tooth	anal	anterior pituitary
aplasia	vertebra	splenic	lingual	anus	neurohypophysis
neoplasm	arm	aorta	tongue	kidney	posterior pituitary
tumour	upper limb	aortic	pharyngeal	renal	adrenal
verruca	carpal	arterial	pharynx	hepatic	suprarenal
wart	carpus	artery	peritoneal	liver	ear drum
boil	wrist	vein	peritoneum	gall bladder	eardrum
furuncle	leg	venous	oesophageal	gallbladder	tympanic membrane
inflammation	lower limb	cranial	oesophagus	testes	bone
inflammatory	thoracic	cranium	bowel	testicle	bony
lump	thorax	skull	intestinal	testicular	osseous
mass	abdomen	nasal	intestine	testis	cartilage
calculus	abdominal	nose	lung		cartilaginous
stone	pelvic		pulmonary		
	pelvis				



# IHTSDO Lexical Resources

## Character Substitutions (1)

DeveloperToolkit\Indexes\tls1\_Index\_Generator\_INT\_20100131.zip

\conf\greek\_character\_conversion\_table.txt (2004)

\conf\char\_conversion\_table.txt (2004)

ALPHA	char=A	0391	char=α	03B1	NU	char=N	039D	char=ν	03BD
BETA	char=B	0392	char=β	03B2	XI	char=Ξ	039E	char=ξ	03BE
GAMMA	char=Γ	0393	char=γ	03B3	OMICRON	char=Ο	039F	char=ο	03BF
DELTA	char=Δ	0394	char=δ	03B4	PI	char=Π	03A0	char=π	03C0
EPSILON	char=Ε	0395	char=ε	03B5	RHO	char=Ρ	03A1	char=ρ	03C1
ZETA	char=Ζ	0396	char=ζ	03B6	SIGMA	char=Σ	03A3	char=σ	03C3
ETA	char=Η	0397	char=η	03B7	TAU	char=Τ	03A4	char=τ	03C4
THETA	char=Θ	0398	char=θ	03B8	UPSILON	char=Υ	03A5	char=υ	03C5
IOTA	char=Ι	0399	char=ι	03B9	PHI	char=Φ	03A6	char=φ	03C6
KAPPA	char=Κ	039A	char=κ	03BA	CHI	char=Χ	03A7	char=χ	03C7
LAMBDA	char=Λ	039B	char=λ	03BB	PSI	char=Ψ	03A8	char=ψ	03C8
MU	char=Μ	039C	char=μ	03BC	OMEGA	char=Ω	03A9	char=ω	03C9



# Other issues

## Case sensitivity

'AIDS procedure' ⇨

## Minimum token length

3 char: HIV, WCC etc

2 char: MI, Hb, Fe etc

1 char: Haemoglobin A  
vs Haemoglobin S

AIDS - Acquired immunodeficiency syndrome antibody test

AIDS (HTLV-III) screening

Aids for severely handicapped

Aids for severely handicapped (procedure)

External memory aids training

External memory aids training (procedure)

External memory aids training (regime/therapy)

Handicapped aids

Handicapped aids (procedure)

Instruction in the use of aids

Notification of AIDS

Notification of AIDS (procedure)

Provision of aids to daily living

Provision of aids to daily living (procedure)

Use of aids for daily life

Use of aids for daily life (procedure)

Use of aids for daily life (regime/therapy)

Use of bathroom aids

Use of bathroom aids (procedure)

Use of bathroom aids (regime/therapy)

Use of indoor mobility aids

Use of indoor mobility aids (procedure)

Use of indoor mobility aids (regime/therapy)

Use of kitchen aids

Use of kitchen aids (procedure)

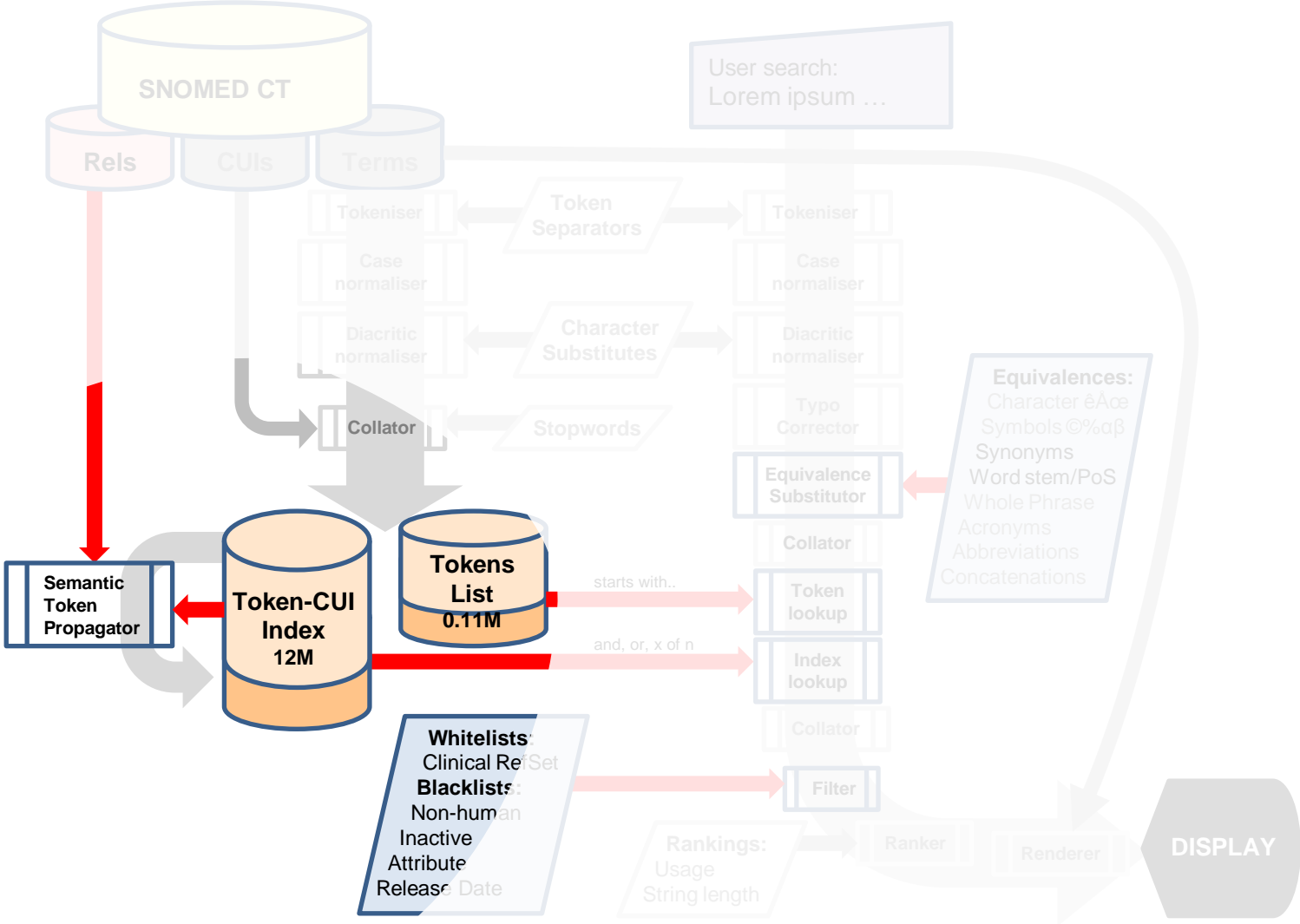
Use of kitchen aids (regime/therapy)

Use of work aids

Use of work aids (procedure)

Use of work aids (regime/therapy)

# Semantic Token Propagation



# ***Outline***

Symptoms

Pathology

Treatment

**Evidence**

# Demonstrator 1 : EdBrowse

Mainly based on supporting tables in RF1

Index tables

- Normalised to capitals

- Published token separators

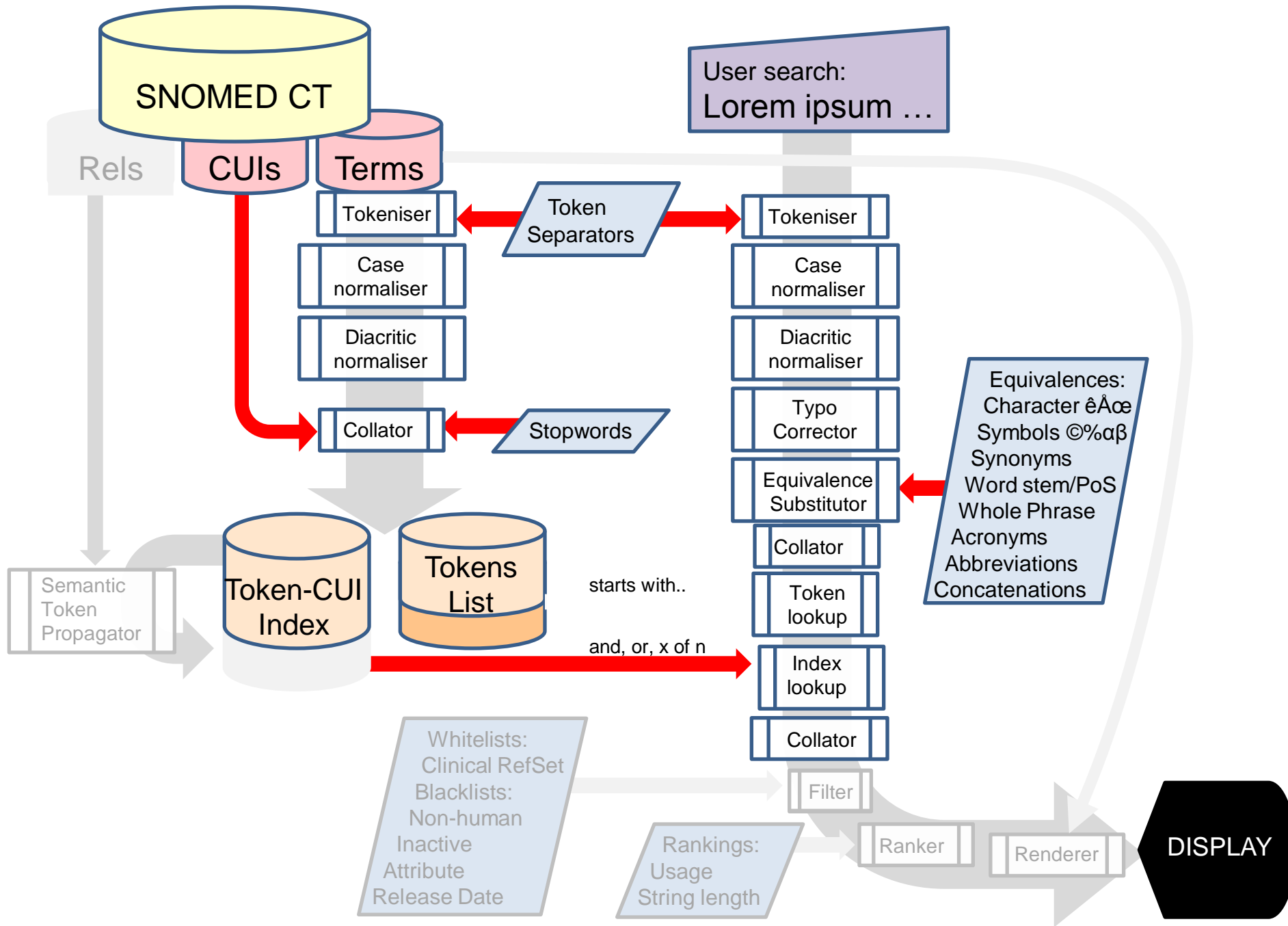
- Diacritic normalisation & character substitution tables

Equivalence table(s)

CUI<->Token variation

(Typo correction)

(Result ranking)

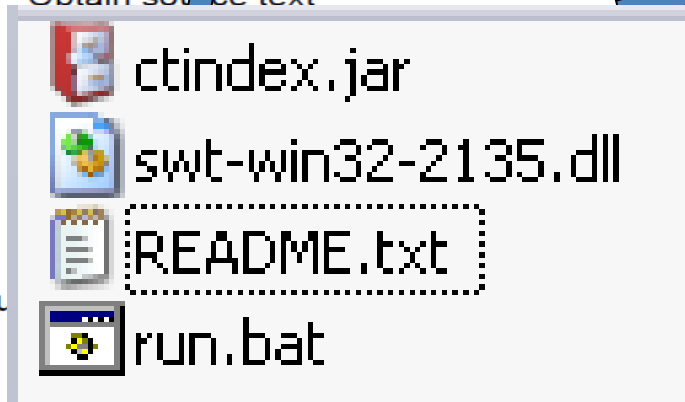


# Generate reference keys DTK – 2.1.2

The keyword generation process consists of the following steps:

- ❖ Single keyword generation

- ❖ Obtain source text



- ❖ Du

- ❖ run.bat

- ❖ Create word pair permutations

- ❖ Result → dual key list

Character  
sub'n/break  
rules

Excluded  
Words  
Table

157792019 | Sézary's disease of spleen

Descriptions  
Table

Single Key:  
SEZARYS  
DISEASE  
SPLEEN

Dual Key:  
DISSEZ  
DISSPL  
SEZSPL

Concepts and  
Descriptions



# RF1 Developer ToolKit Index tables

ConSKeys	
Keyword	ConceptId
AGENTS	102442003
OCCUPATI	102442003
PHYSICAL	102442003
FORCE	102442003
TOXIC	102443008
EXPOSURE	102443008
EVENT	102443008
NONOCCUP	102443008
AGRICULT	102443008
AGENTS	102443008
OCCUPATI	102443008
PHYSICAL	102443008
FORCE	102443008
TOXIC	102444002
EXPOSURE	102444002
EVENT	102444002
DUST	102444002
NOS	102444002
PHYSICAL	102444002
FORCE	102444002
TOXIC	102445001
EXPOSURE	102445001
EVENT	102445001

ConDKeys	
Dualkey	ConceptId
ENTOBS	102401001
WATWHI	102401001
OBSVAL	102401001
OBSWHI	102401001
OBSWAT	102401001
OBSQUA	102401001
OBSRAF	102401001
FUNQUA	102401001
FUNRAF	102401001
FUNOBS	102401001
OBJPHY	102402008
DEVOBJ	102402008
DEVPHY	102402008
DEVSNO	102402008
OBJSNO	102402008
PHYSNO	102402008
OBJPHY	102403003
DEVPHY	102403003
DEWAT	102403003
DEVOBJ	102403003
DEVSKI	102403003
OBJWAT	102403003
OBJSKI	102403003
PHYWAT	102403003

DescSKeys	
Keyword	Description
SUPERVIS	101246016
DEGENERA	101247013
DISC	101247013
INTERVER	101247013
LUMBOSAC	101247013
ALCALIGE	101249011
LATUS	101249011
ACANTHOM	101250011
EPULIS	101250011
ORIGIN	101250011
PERIODON	101250011
AVITELLI	10125015
CHALMERS	10125015
EXCISION	101251010
VEIN	101251010
PHLEBECT	101255018
VENECTOM	101256017
DIFFUSE	101257014
ORGAN	101257014
TISSUE	101257014
UPTAKE	101257014
DEHYDROG	101258016
GLUTAMAT	101258016
BETAGLUT	101259012
GLUTAMYL	101259012

DescDKeys	
DualKey	Description
ANTHLA	100018018
CRIDIF	100019014
CRIGOI	100019014
CRITHY	100019014
CRITOX	100019014
CRIWIT	100019014
DIFGOI	100019014
DIFTHY	100019014
DIFTOX	100019014
DIFWIT	100019014
GOITHY	100019014
GOITOX	100019014
GOIWIT	100019014
THYTOX	100019014
THYWIT	100019014
TOXWIT	100019014
DIFGOI	100020015
DIFSTO	100020015
DIFTHY	100020015
DIFTOX	100020015
DIFWIT	100020015
GOISTO	100020015
GOITHY	100020015
GOITOX	100020015

# Thus we find for:

## 157792019 | Sézary's disease of spleen...

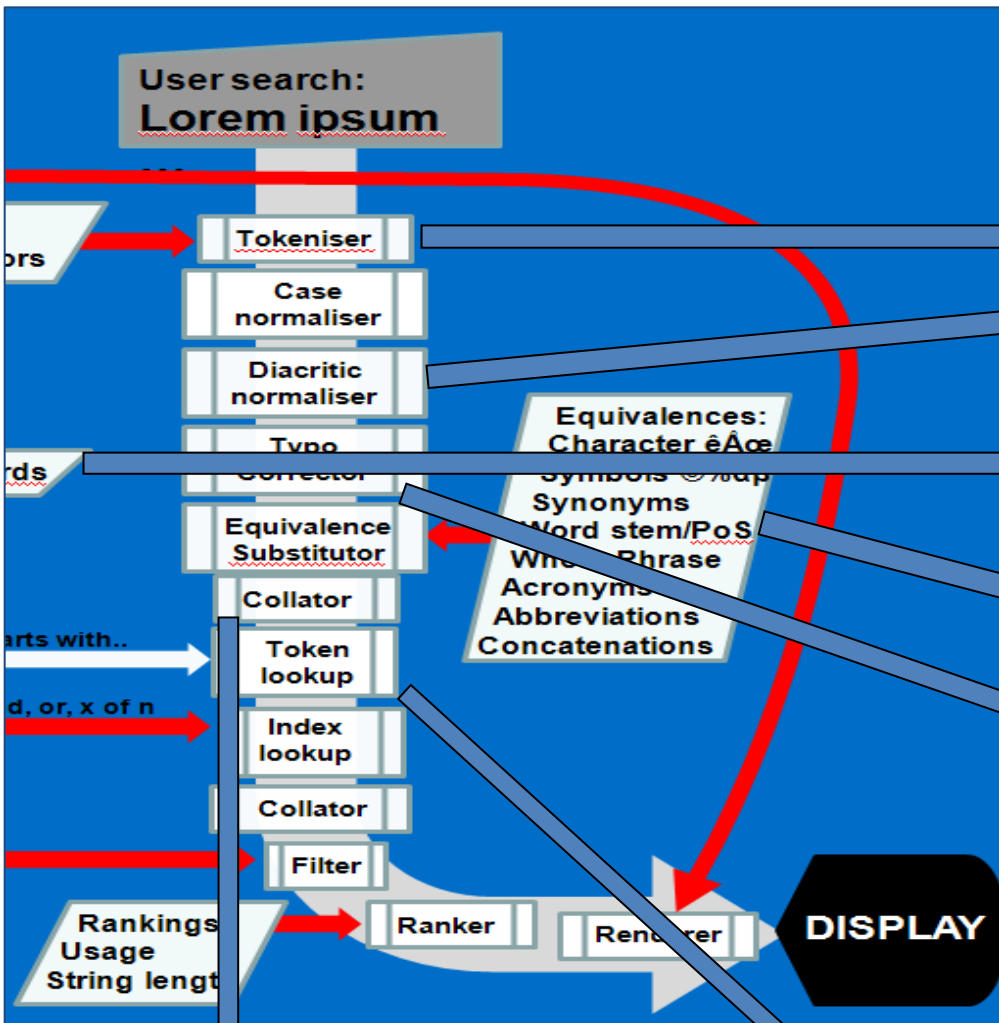
DescSKeys	
Keyword	Description
DISEASE	157792019
SEZARYS	157792019
SPLEEN	157792019
*	

DescDKeys	
DualKey	Description
DISSEZ	157792019
DISSPL	157792019
SEZSPL	157792019
*	

ConSKeys	
Keyword	ConceptId
DISEASE	95263006
DISORDER	95263006
SEZARYS	95263006
SPLEEN	95263006
*	

ConDKeys	
Dualkey	ConceptId
DISSEZ	95263006
DISSPL	95263006
SEZSPL	95263006
*	

# Generate user search tokens



- Character sub'n/break rules
- Excluded Words Table
- Word Equivalents
- Typo – via MS Office

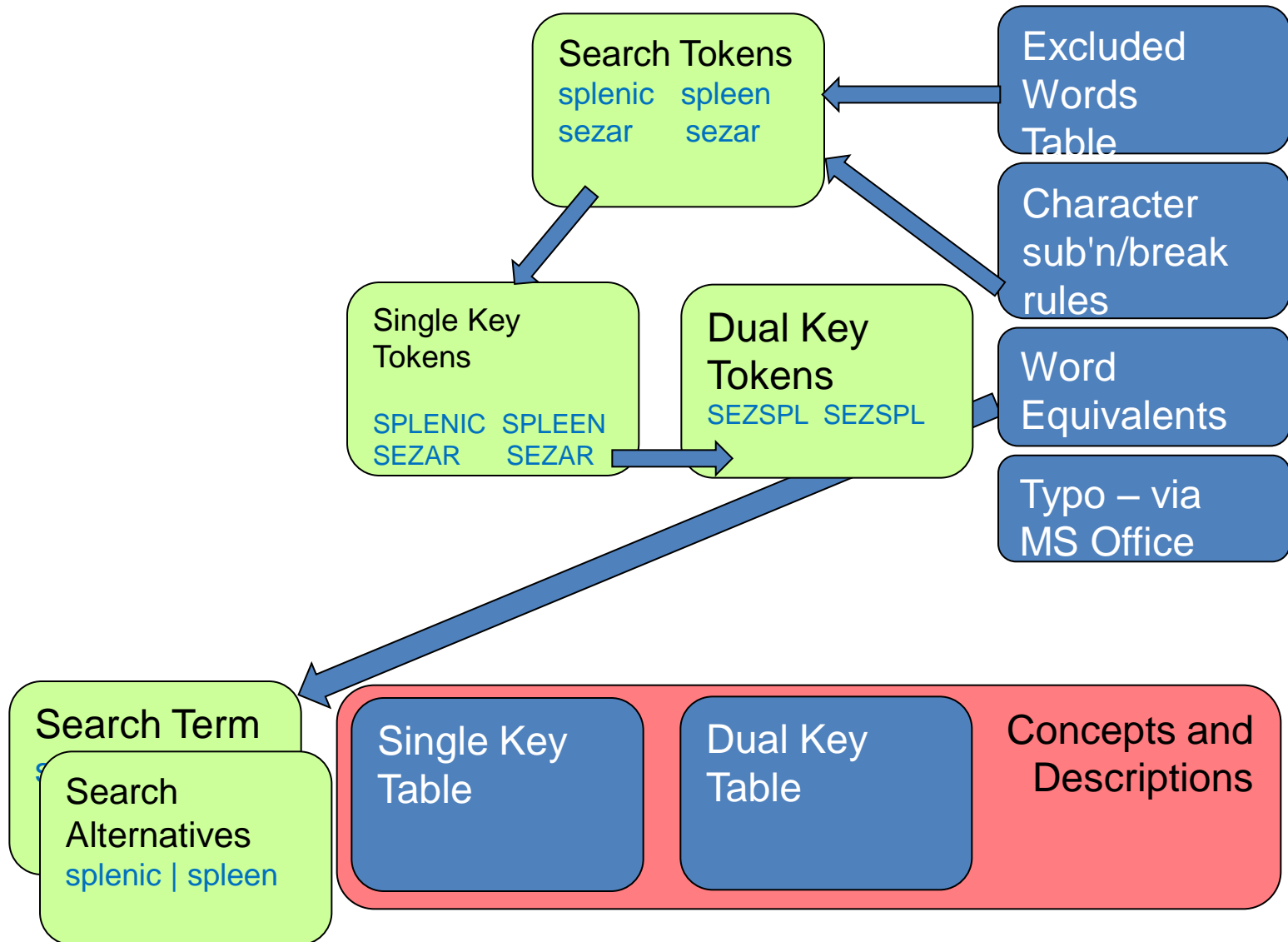
Search Tokens

Single Key Table

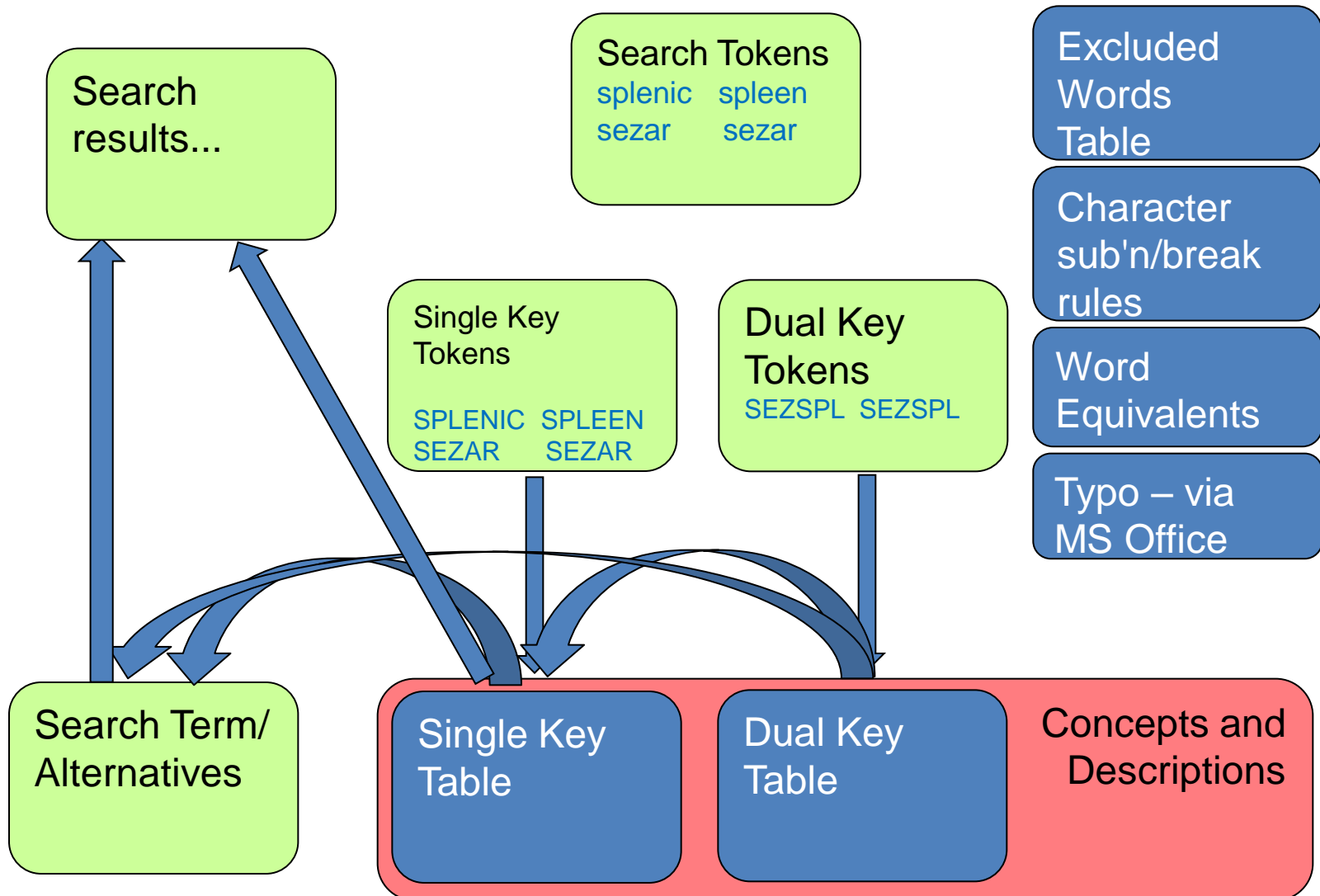
Dual Key Table

Concepts and Descriptions

# Generate user search tokens...



# ..compare token and reference keys and refine



Search Term  
splenic sezar

Strong match between normalised search and index tokens...

SEZAR + SEZARY(S)  
SPLEEN + SPLEEN

SPLEEN ...enables softer matches between lexically variable search and reference terms:

sezar + Sézary's  
splenic + spleen

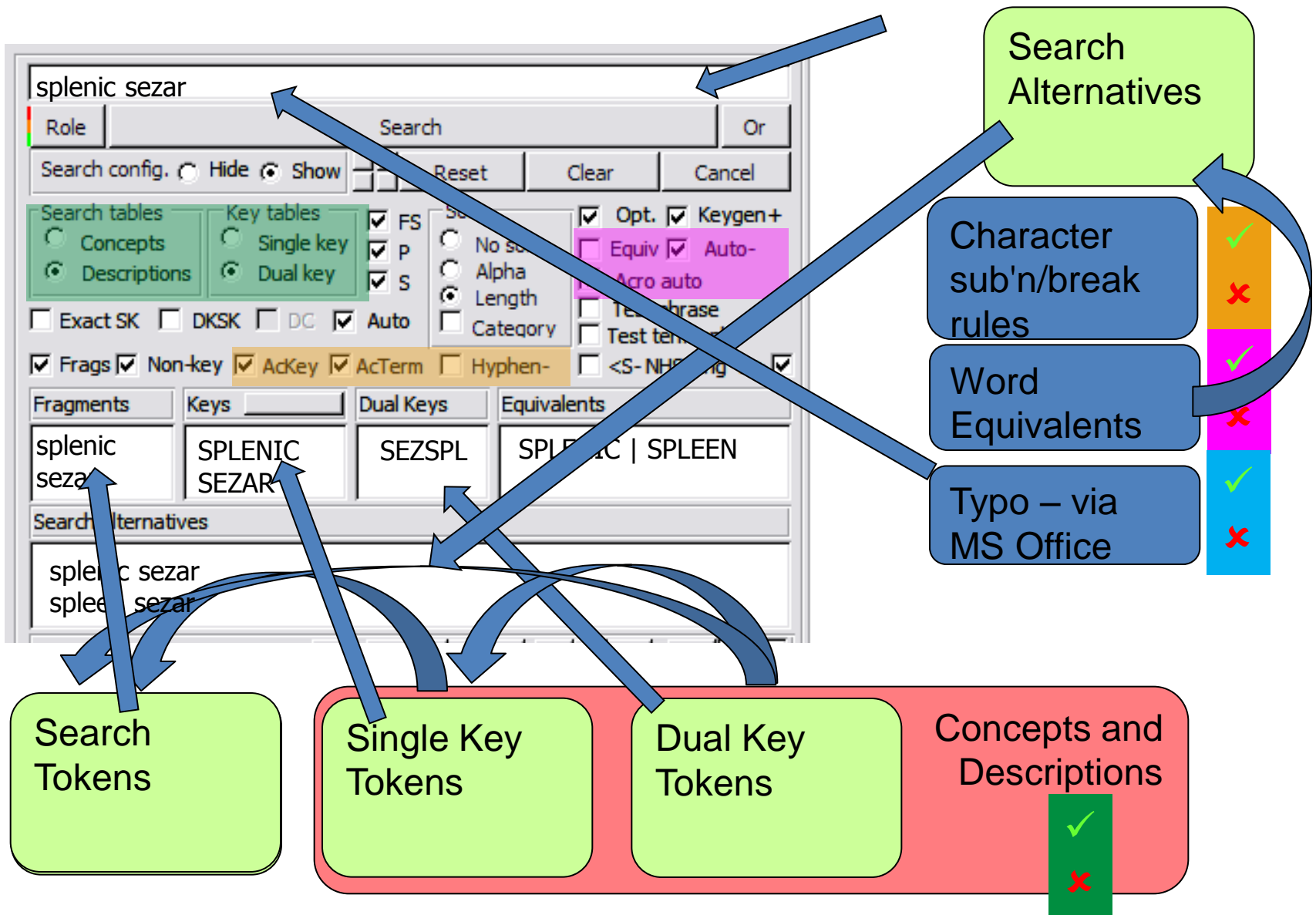
SEZAR	SPLEEN	DescriptionId	TERM				
SEZARY	28208	SPLENIC	731	SPLEEN	625635015	Liver and spleen (combined site) (body structure)	

SEZAR.KeyWord	SPLEEN.KeyWord	DescriptionId	TERM
SEZARYS	SPLEEN	157792019	Sézary's disease of spleen
SEZARYS	SPLEEN	2817094011	Sezary's disease of spleen
SEZARY	SPLEEN	2837759015	Sézary disease of spleen
SEZARYS	SPLEEN	839380010	Sézary's disease of spleen (disorder)

SEZAR\* AND ( SPLENIC\* OR SPLEEN\* )

SEZARY	28377	SPLENIC	241	SPLEEN	184334015	Spleen injection
SEZARY	28419	SPLENIC	740	SPLEEN	460786010	Injection of spleen
SEZARYS	17703	SPLENIC	246	SPLEEN	713163017	Spleen injection (procedure)
SEZARYS	25504	SPLENIC	476397019			Deep vein thrombosis of splenic vein

# Search examples - guide



Character  
sub'n/break  
rules

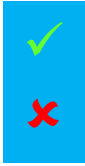


Sézary disease of  
spleen

δ thalassaemia



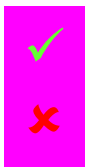
Typo – via  
MS Office



deja vu

diphtheria

Word  
equivalents  
table

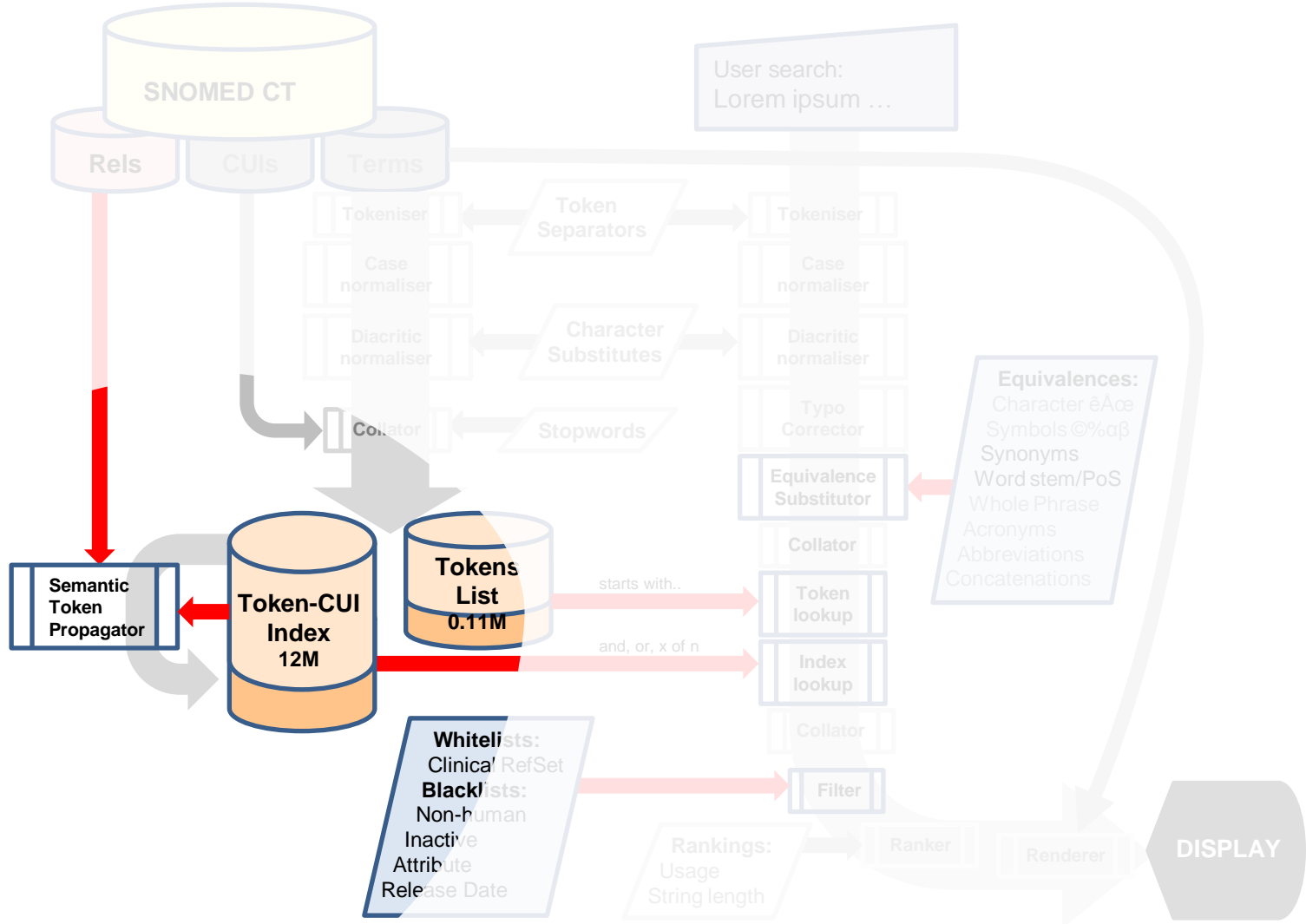


pharyngeal cancer

tranquiliser

traveller diarrhoea

# Demonstrator 2 : Data Migration Workbench



# Demonstrator 2 : Data Migration Workbench

Search expression

Search options

The screenshot displays the 'Cluster Definition Editor' window. The search expression 'tranquillizer' is entered in the search bar. The search options are set to 'Word Equiv', 'Strategy', and 'Filters'. The results pane shows a list of 69 SNOMED CT concepts, with 50 inactive terms remaining in the result set. The results are displayed in a table with columns for the concept ID and the concept name.

Concept ID	Concept Name
231460003	x Tranquilliser abuse
45604009	Tranquilizer (substance)
64386003	Sedative abuse (disorder)
111766000	Poisoning by tranquilizer (disorder)
45335007	x Poisoning by tranquilizer, NEC -RETIRED-
37750001	= Administration of tranquilizer (procedure)
216518007	x Accidental poisoning by tranquilizers (event)
218789000	x Adverse reaction to tranquilizers NOS (finding)
216518007	x Accidental poisoning by tranquilizers (disorder)
27700100000103	History of major tranquilizer misuse (situation)
218784005	x Adverse reaction to other tranquilizers (finding)
27700100000103	History of major tranquilliser misuse (situation)
216543001	x Accidental poisoning by tranquilizers NOS (event)
212653007	x Poisoning: [tranquilliser] or [psychotropic agent]
292343009	Benzodiazepine sedative adverse reaction (disorder)
216538008	x Accidental poisoning by other tranquilizers (event)
218767004	x Adverse reaction to other major tranquilizer (finding)
27770100000101	History of daily major tranquiliser misuse (situation)
216542006	x Accidental poisoning by other tranquilizers NOS (event)
27770100000101	History of daily major tranquilliser misuse (situation)
27773100000107	History of weekly major tranquiliser misuse (situation)
16248006	Poisoning by butyrophenone-based tranquilizer (disorder)
27773100000107	History of weekly major tranquilliser misuse (situation)
48534006	Poisoning by phenothiazine-based tranquilizer (disorder)
27779100000108	Previous history of major tranquilizer misuse (situation)
81914009	Poisoning by benzodiazepine-based tranquilizer (disorder)
27779100000108	Previous history of major tranquilliser misuse (situation)
27776100000102	History of infrequent major tranquiliser misuse (situation)
27776100000102	History of infrequent major tranquilliser misuse (situation)
216519004	x Accidental poisoning by phenothiazine-based tranquilizers (event)
216525000	x Accidental poisoning by butyrophenone-based tranquilizers (event)
216530001	x Accidental poisoning by benzodiazepine-based tranquilizers (event)
218761003	x Adverse reaction to phenothiazine-based tranquilizers NOS (finding)
218766008	x Adverse reaction to butyrophenone-based tranquilizers NOS (finding)
218783004	x Adverse reaction to benzodiazepine-based tranquilizers NOS (finding)
216525000	x Accidental poisoning by butyrophenone-based tranquilizers (disorder)
216519004	x Accidental poisoning by phenothiazine-based tranquilizers (disorder)
216524001	x Accidental poisoning by phenothiazine-based tranquilizers NOS (event)
216530001	x Accidental poisoning by benzodiazepine-based tranquilizers (disorder)
216529006	x Accidental poisoning by butyrophenone-based tranquilizers NOS (event)
216537003	x Accidental poisoning by benzodiazepine-based tranquilizers NOS (event)
82719008	Fetus OR newborn affected by maternal anesthesia AND/OR analgesia (disorder)
206129006	Fetus or neonate affected by maternal tranquilizers during labor and deliv
219118000	x Suicide and selfinflicted poisoning by tranquilizers and other psychotropi
191903008	x Psychoactive substance abuse: [hypnotic (including non-dependent)] or [an-

Concept hierarchy

Results

# Demonstrator 2 : Data Migration Workbench

## **Word Equivalence**

traveller diarrhoea

tranquiliser abuse

flu jab

## **Semantic propagation**

acute angina

remove gallbladder

remove ovary fallopian

# ThankYou!

[jeremy.rogers@hscic.gov.uk](mailto:jeremy.rogers@hscic.gov.uk)

[ed.cheetham@hscic.gov.uk](mailto:ed.cheetham@hscic.gov.uk)

- IHTSDO delivering
- SNOMED CT<sup>®</sup>
- the global clinical terminology



# SNOMED CT Search & Data Entry

**Jeremy Rogers**

**Fadi El-Turk**

**Anne Randorff Højen**

# Agenda

---

- Welcome and introduction
- An Overview of Search and Data Entry techniques
- SNOMED CT Search Candidate Generation Pipeline



Fadi El-Turk and Anne Randorff Højen

# AN OVERVIEW OF SEARCH AND DATA ENTRY TECHNIQUES

# SNOMED CT = a list of words?

## 'Ontology-driven' data quality issues

(disorder)	234693	(morphologic abnormality)	17024
(finding)	123430	(qualifier value)	4544
(procedure)	6583	(body structure)	3517
(situation)	6379	(substance)	2340
(event)	4592	(attribute)	1293
(regime/therapy)	1344	(observable entity)	1139
		(physical object)	746
		(product)	233
		(cell)	226
		(navigational concept)	206
		(organism)	199
		(physical force)	152
		(record artifact)	97
		(ethnic group)	17
		(environment)	15
		(assessment scale)	14
		(person)	10
		(specimen)	9
		(administrative concept)	6
		(tumor staging)	5
		(social concept)	4
		(cell structure)	3
		(occupation)	2
		(inactive concept)	1
<b>TOTALS</b>	<b>377021</b>		<b>31802</b>

Types of SNOMED CT used to to code main reason for the patients attend A&E Department

# SNOMED CT = a list of words?

## Miscoding examples

1097	Temperature	246508008 Temperature (attribute)
17	Drug used	246488008 Drug used (attribute)
373	ETOH - Alcohol intake	160573003 Alcohol intake (observable entity)
136	Nasogastric tube	17102003 Nasogastric tube, device (physical object)
82	Catheter	19923001 Catheter, device (physical object)
78	Dressing	37898001 Dressing, device (physical object)
110		53570002 Removal of foreign body from eye (procedure)
83		172828005 Removal of foreign body from nose (procedure)
43		172278002 Removal of foreign body from eyelid (procedure)
293		82576008 Retained foreign body in eye (disorder)
166		74699008 Foreign body in nose (disorder)
7		25012008 Retained foreign body of eyelid (disorder)

## SNOMED CT = a list of words? Miscoding examples

- 23% of 74 abdominal aortic aneurysms miscoded as a Drug Trade Family (9192101000001100 AAA (product)); AAA make sore throat spray (and not much else)
- 49% of 5,062 probably alcohol-related attendances were miscoded as either the substance (alcohol, ethyl alcohol) or as the word 'intoxicated', even though (in SNOMED) the latter actually means being inappropriately elated, possibly but not necessarily as a result of psychoactive substance intake and so not necessarily involving alcohol intake at all.

# SNOMED CT is NOT a long list of clinical terms

---

## SNOMED CT is complex and comprehensive results!

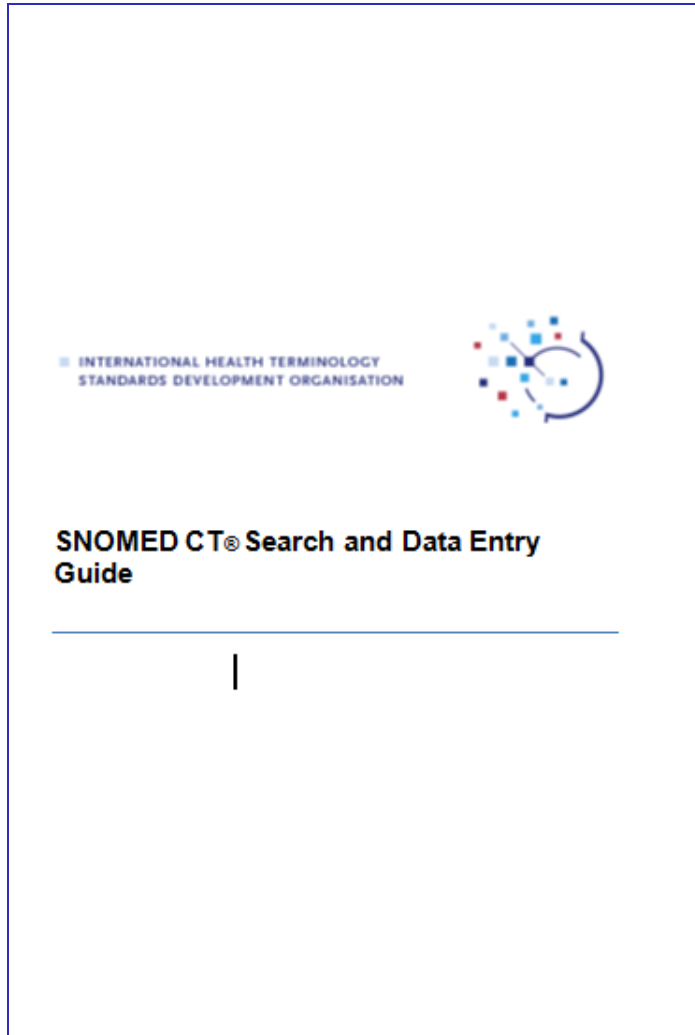
- Nutritional finding
- Reproductive finding
- Finding relating to sexuality and sexual activity
- Communication, speech and language finding
- Finding relating to psychosocial functioning
- Cognitive function finding
- Sleep rest pattern finding
- Elimination pattern
- Value belief finding
- Pain
- Observation of sensation
- Finding of functional performance and activity
- Mucosal Finding
- Skin Finding
- Finding of respiration
- Cardiovascular finding



Treating it as a long list will probably give you a long list of search results!

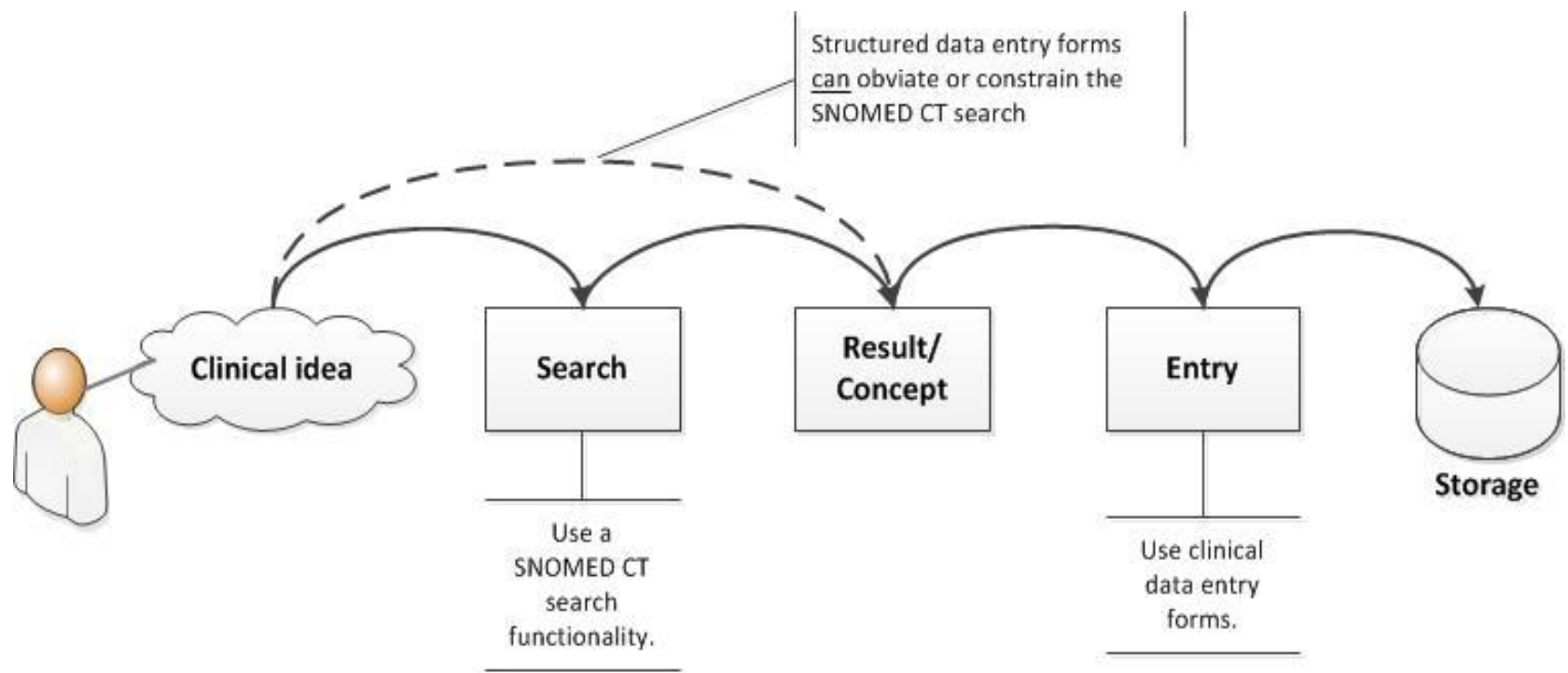
# The new IHTSDO 'Search and Data entry Guide'

---



- Provides recommendations on Search and Data Entry Techniques
- Developed as part of the 'SNOMED CT Implementation Advisor Scheme 2013'
- Currently being reviewed by the I&I committee
- To be published in January 2014 release

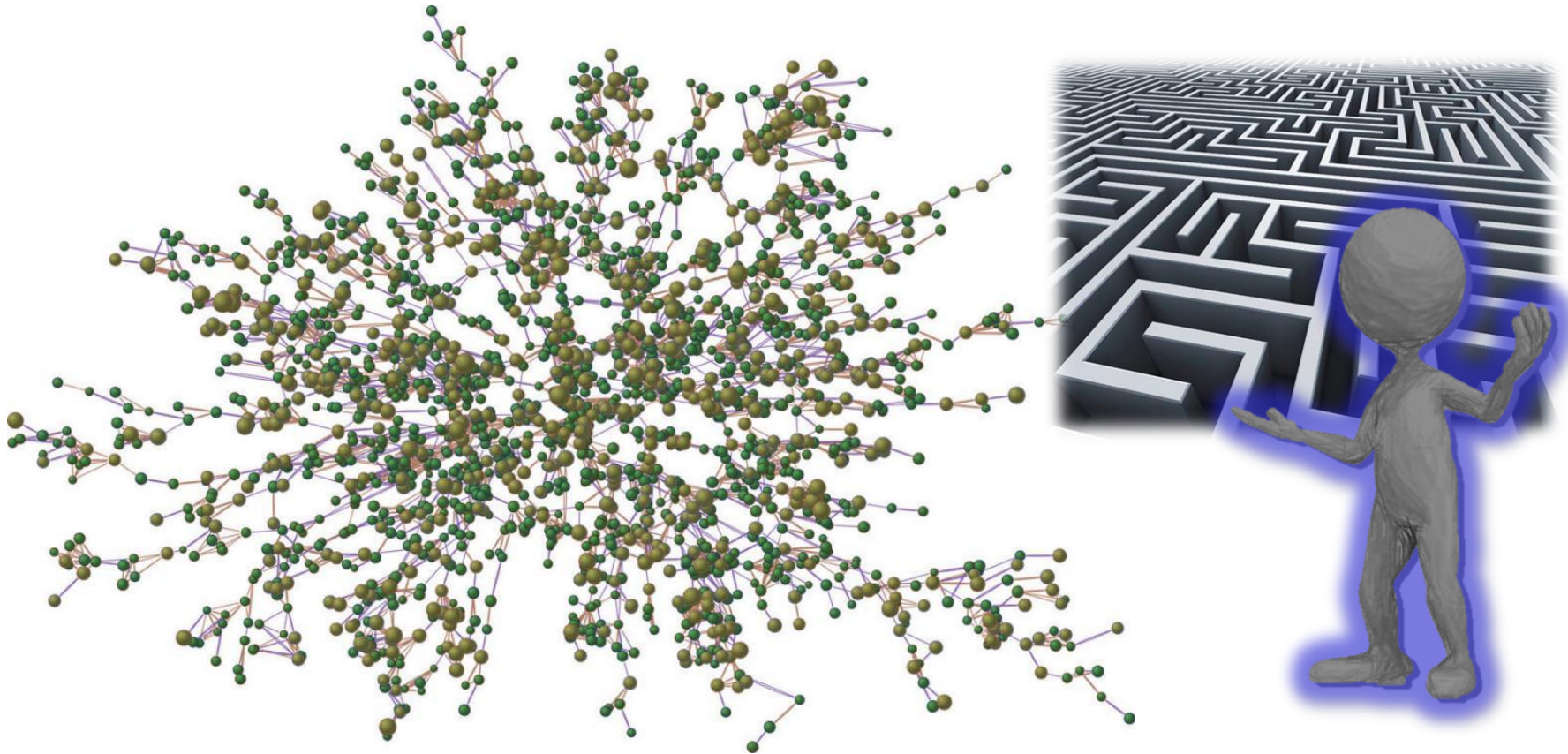
# Search, browse, data entry – what does it mean?



## Why are search and data entry techniques important?

---

- SNOMED CT is comprehensive and complex





# Use cases for searches

---

- Writing an entry into a clinical record at or near the point of care
- Designing a data entry template
- Designing a query or report
- Binding SNOMED CT to knowledge artefacts
- Evaluation of the extent of coverage in a particular domain
- Mapping between SNOMED CT and another code scheme
- Developing a reference set

Other?

## Search by identifiers

---

- The default behavior of the SCTID-to ConceptID search operation **must** include both SNOMED CT ConceptIDs and DescriptionIDs as possible input.
- Diacritic and accented characters **must** be properly displayed and searched over e.g. Sjögren, Ménière.

## Search by text

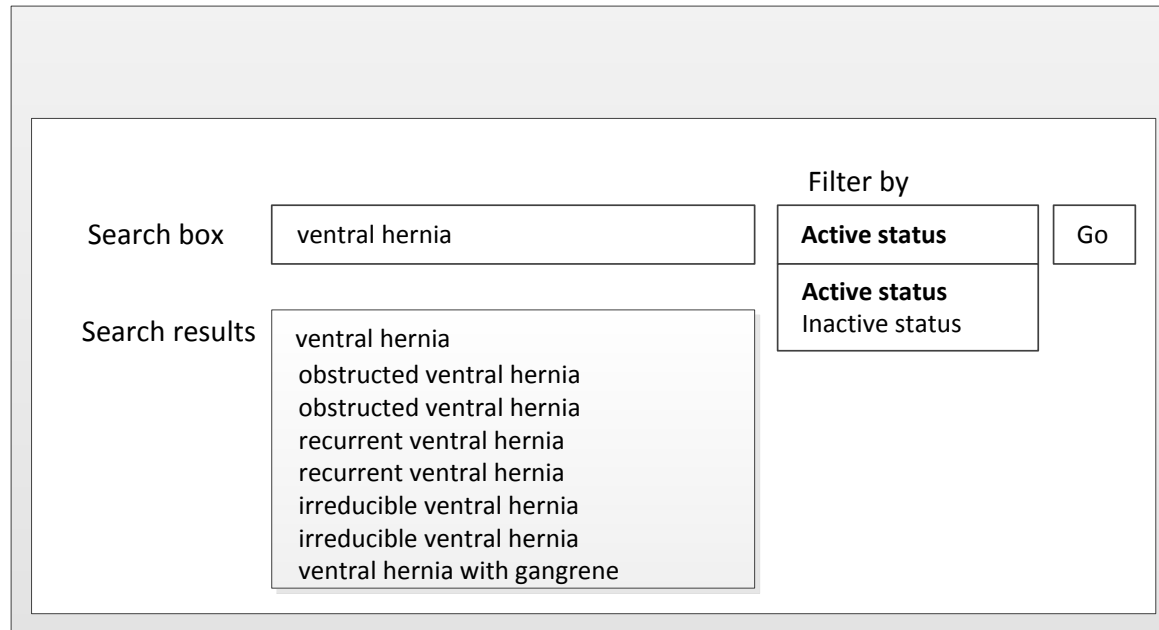
---

- String searches must be user configurable to search token as:
  - words any order
  - phrase match
  - identical term
  - starts with
  - ends with
- Support the display and search over:
  - Diacritic and accented characters e.g. Sjögren, Ménière.
  - Superscript (^ ^) and subscript (> <) (e.g. for chemical formulae or radionuclides <sup>186</sup>Platinum).
  - Search over descriptions involving % or \* symbols.
    - Must not interact with any wildcarding system in the query syntax.

## Constrain searches (1)

---

- Constrain searches by Concept and Description status



The screenshot shows a search interface with the following components:

- Search box:** Contains the text "ventral hernia".
- Search results:** A list of search results including:
  - ventral hernia
  - obstructed ventral hernia
  - obstructed ventral hernia
  - recurrent ventral hernia
  - recurrent ventral hernia
  - irreducible ventral hernia
  - irreducible ventral hernia
  - ventral hernia with gangrene
- Filter by:** A section with a "Go" button and a dropdown menu currently set to "Active status".

- Recommended for data entry use cases
- There are a few use cases where a user may legitimately wish to search *Inactive Concepts* and *Descriptions* (e.g. querying a clinical finding in a research database for retrospective analysis).

## Constrain searches (2)

- Constrain searches by supertype ancestors (e.g. Disease)

1

Search box: renal calculus

Search results:

- renal calculus
- calyceal renal calculus
- calcium renal calculus
- O/E: renal calculus
- history of renal calculus
- uric acid renal calculus
- calculus in renal pelvis
- on examination - renal calculus
- O/E: cystine renal calculus

Filter by: **Disease**

Procedure  
Clinical Finding

Go

Results found: 30

2

Search box: renal calculus

Search results:

- renal calculus
- calyceal renal calculus
- calcium renal calculus
- O/E: renal calculus
- uric acid renal calculus
- calculus in renal pelvis
- on examination - renal calculus
- O/E: cystine renal calculus
- O/E: oxalate renal calculus

Filter by: **Disease**

Go

Results found: 13

- Recommended as a user configurable option for most use cases

## Constrain searches (3)

---

- **Constrain search by ref sets**
  - Language Subsets to avoid uncommon or foreign terms
  - Realm Subsets to simplify or encourage selection of concepts or used in a particular country, organization, or specialty
  - Context Subsets to specify or order the valid Concepts for entry in a particular field

## Constrain searches (4)

---

- Constrain search to avoid multiple hits on the same concept

Search box	<input type="text" value="hernia"/>	Display	
Search results	<ul style="list-style-type: none"><li>herniated structure</li><li>hernia</li><li>herniated tissue</li><li>herniated structue</li><li>herniation</li></ul>	<input type="text" value="Fully specified name"/>	<input type="button" value="Go"/>
		<input type="text" value="Fully specified name"/>	<input type="text" value="Synonym"/>

## Extend searches (1)

---

- Use Word Equivalents



## Extend searches (2)

---

- Extend by post-coordinated searching

## Improve search speeds

---

- Enable real time searching
- Show an indication of estimated number of matches before starting a search
- Allow slow searches to be paused or cancelled
- Optimize indexing
- Enable background encoding
  - Caution: Anyone implementing this approach should take care to undertake appropriate quality assurance of the results for a given data entry use-case

# OPTIMISING THE DISPLAY OF SEARCH RESULTS

# Order Search Results Rationally (1)

- **Order shortest matching terms first**

The screenshot shows a search interface with a search box containing the text 'hernia' and a 'Go' button. Below the search box is a list of search results: 'hernia', 'hernia', 'hernia sac', 'hernia truss', 'hernia belt', 'cecal hernia', 'hernia repair', 'pudental hernia', and 'sciatic hernia'. To the right of the list, a large downward-pointing arrow indicates the sorting order, with 'Shortest description' at the top and 'Longer description' at the bottom.

Search results	Description Length
hernia	Shortest description
hernia	Shortest description
hernia sac	Longer description
hernia truss	Longer description
hernia belt	Longer description
cecal hernia	Longer description
hernia repair	Longer description
pudental hernia	Longer description
sciatic hernia	Longer description

## Order Search Results Rationally (2)

- Order preferred term matches before synonyms

The screenshot shows a search interface with a search box containing the text 'hernia' and a 'Go' button. Below the search box is a list of search results. The results are ordered as follows: Preferred term (P) hernia, Synonym (S) hernia, Preferred term (P) hernia sac, Preferred term (P) hernia belt, Synonym (S) hernia truss, Preferred term (P) hernia repair, Synonym (S) hernia cerebri, Preferred term (P) hernia adiposa, and Preferred term (P) hernia of ovary. Two callout boxes on the left identify the first result as a 'Preferred term' and the second result as a 'Synonym'.

Type	Term
Preferred term	P hernia
Synonym	S hernia
Preferred term	P hernia sac
Preferred term	P hernia belt
Synonym	S hernia truss
Preferred term	P hernia repair
Synonym	S hernia cerebri
Preferred term	P hernia adiposa
Preferred term	P hernia of ovary

## Order Search Results Rationally (3)

---

- Order user preferred language matches first in multilingual environments

## Order Search Results Rationally (4)

---

- Order according to priority in any active reference sets

## Order Search Results Rationally (5)

- Display search results a few at a time, with most frequently used descriptions listed first.

1

Search box

Search results

- P hernia
- S hernia
- P hernia sac
- P hernia belt
- S hernia truss
- P hernia repair
- S hernia cerebri
- P hernia adipoc...
- P hernia of ovary

2

Search box

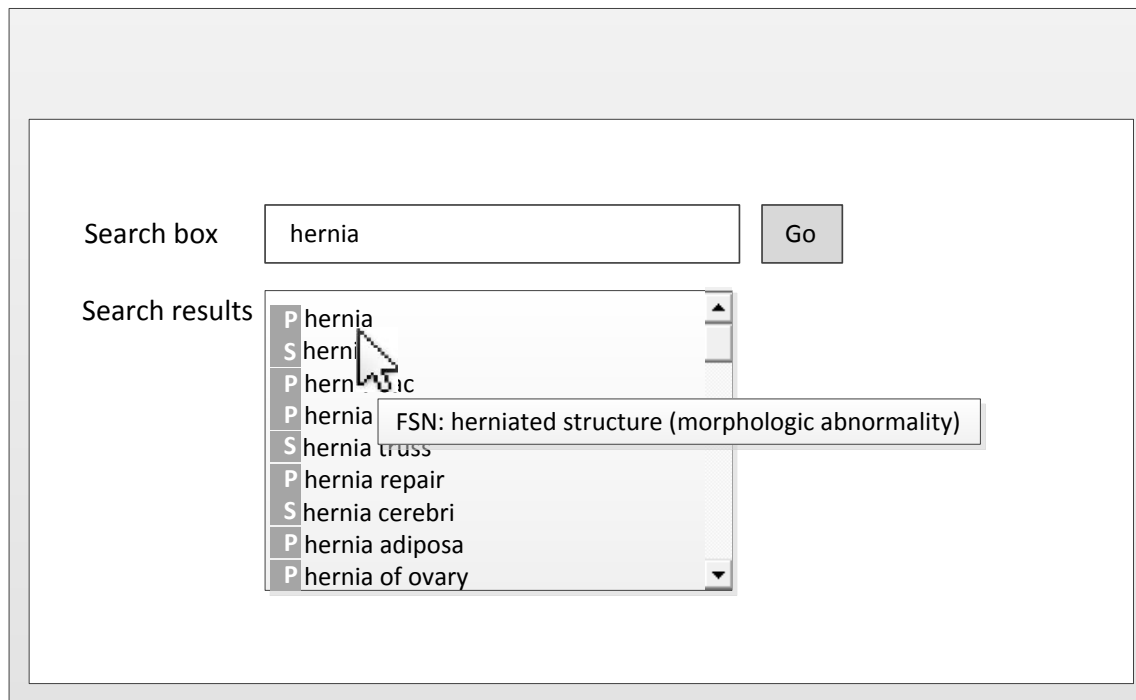
Search results

*Frequently searched terms*

- S hernia cerebri
- 
- S hernia
- P hernia
- S hernia sac
- P hernia belt
- P hernia truss



# Distinguish identical terms of different concepts (1)



## Distinguish identical terms of different concepts (2)

Search box

Search results

- P hernia
- S hernia
- P hernia
- P hernia
- S hernia t FSN: hernia of abdominal cavity (disorder)
- P hernia repair
- S hernia cerebri
- P hernia adiposa
- P hernia of ovary

# Rationalize search results by subsumption checking

Before

Search box

Search results

- P hernia
- S hernia
- P hernia sac
- P hernia belt
- S hernia truss
- P hernia repair
- S hernia cerebri
- P hernia adiposa
- P hernia of ovary

Results found: 413

After

Search box

Search results

- S hernia

Results found: 1

# Display navigation results effectively (1)

## Using the subtype hierarchy

The screenshot displays a search interface with a search box containing the text 'hernia' and a 'Go' button. Below the search box, the search results are presented as a hierarchical tree structure. The root node is 'SNOMED CT Concept', which is expanded to show 'body structure'. This is further expanded to show 'morphologically altered structure', which is expanded to show 'morphologically abnormal structure', which is expanded to show 'mechanical abnormality', which is expanded to show 'protrusion', which is expanded to show 'hernia'. The 'hernia' node is expanded to show three sub-nodes: 'complete hernia', 'hernia adiposa', and 'hernia, reduction en masse'. Each node in the hierarchy has a small square icon with a minus sign to its left, indicating it is expanded. The 'hernia' node has a plus sign icon to its left, indicating it is expanded. The sub-nodes under 'hernia' have plus sign icons to their left, indicating they are collapsed.

Search box: hernia [Go]

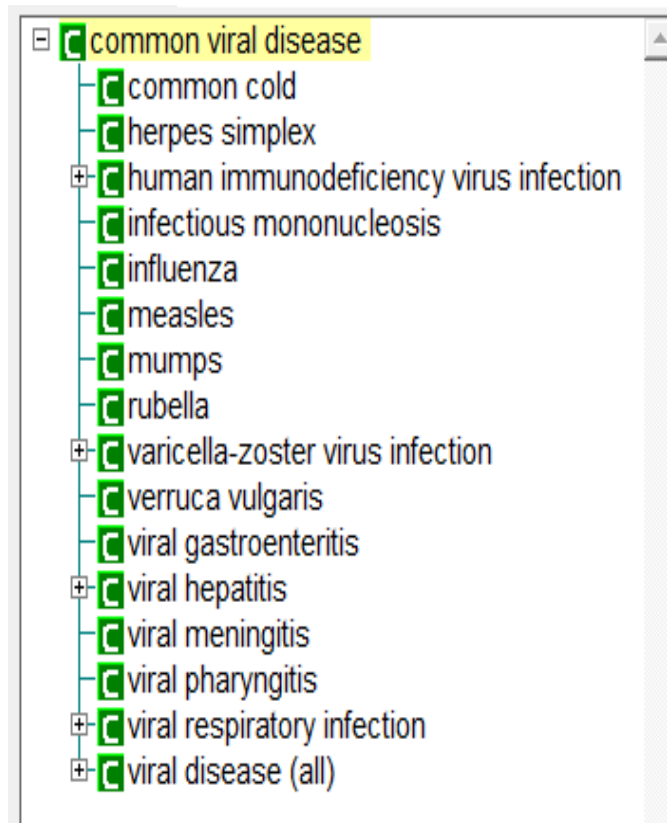
Search results:

- [-] SNOMED CT Concept
  - [-] body structure
    - [-] morphologically altered structure
      - [-] morphologically abnormal structure
        - [-] mechanical abnormality
          - [-] protrusion
            - [-] **hernia**
              - [+] complete hernia
              - [+] hernia adiposa
              - [+] hernia, reduction en masse

## Display navigation results effectively (2)

---

Using the navigation hierarchy (hand-crafted)



# SNOMED CT data entry

---

- Benefits of SNOMED CT data entry
- SNOMED CT and structured records
- Requirements for entry and display
- Constraining data entry
- Entering refinements for postcoordinated expressions

Jeremy Rogers

# **SNOMED SEARCH CANDIDATE GENERATION PIPELINE**

- IHTSDO delivering
- SNOMED CT<sup>®</sup>
- the global clinical terminology



# Thank you for your attention

## Questions?

- Contact IHTSDO: [info@ihtsdo.org](mailto:info@ihtsdo.org)
- Web site: [www.ihtsdo.org](http://www.ihtsdo.org)