
PHOEBE 2.0: CREATING COMPREHENSIVE CONCEPT SETS

Now with:

- enhanced lexical match (with lemmatization, bigrams, conversion to a common part of speech, tf-idf and pairwise cosine similarity)
- patient context (pairwise cosine similarity based on the vectors of the concept co-occurring with a given concept)

Use with:

- Counts from the network
- Your own counts

Anna, Patrick, George and Chris

Anna Ostropolets: ao2671@cumc.columbia.edu

Go to atlas-demo.ohdsi.org.

Option 1: start with a string

Type your string into Search tab and click on the shopping card

You will be directed to a new concept set (don't forget to name it!) with the recommendations. But first, click on the Concept Set Expression

[illegible]

This is your starting point:

🛒

Concept Set #1871400

created by anonymous on 2022-10-12 23:46, modified by anonymous on 2022-10-12 23:46

Prostate Cancer [PHOEBE]

📄

✕

📄

🔍

Optimize

🗑️

Concept Set Expression

Included Concepts174

Included Source Codes

Recommend

Explore Evidence

Export

Import

Compare

Versions

Messages

Enter the concept set description here

Show50entries

Filter:Search...

Showing 1 to 1 of 1 entries

Previous1Next

<input type="checkbox"/>	Concept Id	Concept Code	Concept Name	Domain	Standard Concept Caption	<input type="checkbox"/>	Exclude	<input checked="" type="checkbox"/>	Descendants	<input type="checkbox"/>	Mapped
<input checked="" type="checkbox"/>	4163261	399068003	Malignant tumor of prostate	Condition	Standard	<input type="checkbox"/>		<input checked="" type="checkbox"/>		<input type="checkbox"/>	

Remove Selected

Add Concepts

Classification

Non-Standard

Standard

Now, let's check out the recommendations.

We have different types: ontology, lexical and patient context.

🛒

Concept Set #1871400

created by anonymous on 2022-10-12 23:46, modified by anonymous on 2022-10-12 23:46

Prostate Cancer [PHOEBE]

📄

✕

📄

🔍

Optimize

🗑️

Concept Set Expression

Included Concepts174

Included Source Codes

Recommend

Explore Evidence

Export

Import

Compare

Versions

Messages

Show columns

Copy

CSV

Show50entries

Filter:Search...

Showing 1 to 50 of 66 entries

Previous12Next

▼ Vocabulary

SNOMED (37)

HCPSCS (13)

CPT4 (12)

ICDO3 (3)

ICDO3 (3)

Invalid Reason

Valid (66)

Class

Clinical Finding (33)

HCPSCS (13)

CPT4 (12)

ICDO Condition (3)

ICDO Condition (3)

Domain

Condition (36)

Observation (24)

Measurement (5)

Procedure (4)

Relationship

Lexical via standard (37)

Ontology-parent (24)

Patient context (4)

Ontology-descendant (3)

Relationship via standard (3)

<input type="checkbox"/>	Id	Code	Name	Class	RC	DRC	Domain	Vocabulary	Count
<input checked="" type="checkbox"/>	439392	372087000	Primary malignant neoplasm	Clinical Finding	5,737,926	1,039,285,652	Condition	SNOMED	1
<input checked="" type="checkbox"/>	4134595	128290005	Chronic disease of genitourinary system	Clinical Finding	60,964	868,536,150	Condition	SNOMED	1
<input checked="" type="checkbox"/>	40492037	448882009	Malignant neoplasm of intraabdominal organ	Clinical Finding	200	416,035,469	Condition	SNOMED	1
<input checked="" type="checkbox"/>	197804	93839008	Primary malignant neoplasm of intra-abdominal organs	Clinical Finding	268,712	360,107,907	Condition	SNOMED	1
<input checked="" type="checkbox"/>	4177236	363484005	Malignant tumor of pelvis	Clinical Finding	4,646	255,290,575	Condition	SNOMED	1
<input checked="" type="checkbox"/>	4272032	63476009	Prostate specific antigen measurement	Procedure	40,999,567	242,850,865	Measurement	SNOMED	1
<input checked="" type="checkbox"/>	198988	93953009	Primary malignant neoplasm of pelvis	Clinical Finding	429,115	231,286,010	Condition	SNOMED	1
<input checked="" type="checkbox"/>	432851	128462008	Secondary malignant neoplastic disease	Clinical Finding	6,627,602	159,384,332	Condition	SNOMED	1
<input checked="" type="checkbox"/>	4181487	363515000	Malignant tumor of male genital organ	Clinical Finding	3,585	135,139,145	Condition	SNOMED	1
<input checked="" type="checkbox"/>	4129902	126906006	Neoplasm of prostate	Clinical Finding	13,333	134,284,274	Condition	SNOMED	2

You can examine them separately or together. The concepts are ordered by the record count with descendants (counts are coming from the 22 data sources in the OHDSI network, but you can also change it to your data source in Configuration in your Atlas instance).

For example, let's look at the recommendations coming from the patient context.
Let's add some of them:

🛒

Concept Set #1871400

created by anonymous on 2022-10-12 23:46, modified by anonymous on 2022-10-12 23:46

Prostate Cancer [PHOEBE]

📄

✕

📄

📄

Optimize

🗑️

Concept Set Expression

Included Concepts 174

Included Source Codes

Recommend

Explore Evidence

Export

Import

Compare

Versions

Messages

Show columnsCopyCSVShow 50 entriesFilter: Search...

Showing 1 to 4 of 4 entriesPrevious 1 Next

	Id	Code	Name	Class	RC	DRC	Domain	Vocabulary	Count
<input checked="" type="checkbox"/>	200970	92691004	Carcinoma in situ of prostate	Clinical Finding	1,326,998	1,327,098	Condition	SNOMED	1
<input type="checkbox"/>	40482030	445068007	Dysplasia of prostate	Clinical Finding	396,561	396,561	Condition	SNOMED	1
<input checked="" type="checkbox"/>	192681	254901000	Prostatic intraepithelial neoplasia	Clinical Finding	86,584	89,466	Condition	SNOMED	1
<input checked="" type="checkbox"/>	37209512	16294321000119104	Atypical small acinar proliferation of prostate	Clinical Finding	0	0	Condition	SNOMED	1

Showing 1 to 4 of 4 entriesPrevious 1 Next

Vocabulary

SNOMED (37)

HCPSC (13)

CPT4 (12)

ICDO3 (3)

ICDO3 (3)

Invalid Reason

Valid (66)

Class

Clinical Finding (33)

HCPSC (13)

CPT4 (12)

ICDO Condition (3)

Context-dependent (2)

Domain

Condition (36)

Observation (24)

Measurement (5)

Procedure (1)

Relationship

Excluded via standard (37)

Ontology-parent (24)

Patient context (4)

Ontology-descendant (3)

Lexical via source (2)

☐ Exclude

☒ Descendants

☐ Mapped

Add To Concept Set

Now our concept set has new concepts:

Concept Set #1871400
created by anonymous on 2022-10-12 23:46, modified by anonymous on 2022-10-12 23:46

Prostate Cancer [PHOEBE]

Concept Set Expression | Included Concepts (194) | Included Source Codes | Recommend | Explore Evidence | Export | Import | Compare | Versions | Messages

Enter the concept set description here

Show 50 entries Filter: Search...

Showing 1 to 4 of 4 entries

<input type="checkbox"/>	Concept Id	Concept Code	Concept Name	Domain	Standard Concept Caption	<input type="checkbox"/> Exclude	<input checked="" type="checkbox"/> Descendants	<input type="checkbox"/> Mapped
<input checked="" type="checkbox"/>	192681	254901000	Prostatic intraepithelial neoplasia	Condition	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/>	4163261	399068003	Malignant tumor of prostate	Condition	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/>	200970	92691004	Carcinoma in situ of prostate	Condition	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/>	37209512	16294321000119104	Atypical small acinar proliferation of prostate	Condition	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Classification Non-Standard Standard

And we have more recommendations for updated concept set. For example, these lexical recommendations:

Concept Set #1871400
created by anonymous on 2022-10-12 23:46, modified by anonymous on 2022-10-12 23:46

Prostate Cancer [PHOEBE]

Concept Set Expression | Included Concepts (194) | Included Source Codes | Recommend | Explore Evidence | Export | Import | Compare | Versions | Messages | Delete

Show columns | Copy | CSV | Show 50 entries Filter: Search...

Showing 1 to 24 of 24 entries

Vocabulary
SNOMED (36)
HCPCS (13)
CPT4 (12)
ICDO3 (3)
ICD9CM (3)
Invalid Reason
Valid (65)
Class
Clinical Finding (32)
HCPCS (13)
CPT4 (12)
ICDO Condition (3)
Context-dependent (2)
Domain
Condition (35)
Observation (24)
Measurement (5)
Procedure (1)
Relationship
Lexical via standard (37)
Ontology-parent (26)
Ontology-descendant (3)
Lexical via source (2)
Patient context (1)

<input type="checkbox"/>	Id	Code	Name	Class	RC	DRC	Domain	Vocabulary	Count
<input checked="" type="checkbox"/>	4213085	414205003	Family history of prostate cancer	Context-dependent	75,617	75,617	Observation	SNOMED	1
<input checked="" type="checkbox"/>	42628527	81539	Oncology (high-grade prostate cancer), biochemical assay of four proteins (Total PSA, Free PSA, Intact PSA, and human kallikrein-2 [hK2]), utilizing plasma or serum, prognostic algorithm reported as a probability score	CPT4	66,108	66,108	Observation	CPT4	1
<input checked="" type="checkbox"/>	2106847	3271F	Low risk of recurrence, prostate cancer (PRCA)	CPT4	52,531	52,531	Observation	CPT4	2
<input checked="" type="checkbox"/>	2106848	3272F	Intermediate risk of recurrence, prostate cancer (PRCA)	CPT4	36,220	36,220	Observation	CPT4	1
<input checked="" type="checkbox"/>	2617863	G8464	Clinician documented that prostate cancer patient is not an eligible candidate for adjuvant hormonal therapy; low or intermediate risk of recurrence or risk of recurrence not determined	HCPCS	35,051	35,051	Observation	HCPCS	1
<input checked="" type="checkbox"/>	2108678	4164F	Adjuvant (ie, in combination with external beam radiotherapy to the prostate for prostate cancer) hormonal therapy (gonadotropin-releasing hormone [GnRH] agonist or antagonist) prescribed/administered (PRCA)	CPT4	32,457	32,457	Observation	CPT4	1
<input checked="" type="checkbox"/>	2106846	3270F	Bone scan not performed prior to initiation of treatment nor at any time since diagnosis of prostate	CPT4	29,439	29,439	Observation	CPT4	1

For example, it can look like this (malignant neoplasms and borderline cancer):

Option 2: start with an existing concept set

You can follow the same process with each concept set that you already created.

Note that the recommendations are only generated for standard concepts.

For example, this randomly selected Atlas concept set:

🛒 Concept Set #1871394

created by anonymous on 2022-10-12 8:18, modified by anonymous on 2022-10-12 8:19

EXER Hypoglycemia

🔒

✕

📄

🏷️

Optimize

🗑️

Concept Set Expression

Included Concepts 11

Included Source Codes

Recommend

Explore Evidence

Export

Import

Compare

Versions

Messages

Enter the concept set description here

Show 50 entries

Filter: Search...

Showing 1 to 1 of 1 entries

Previous 1 Next

<input type="checkbox"/>	Concept Id	Concept Code	Concept Name	Domain	Standard Concept Caption	<input type="checkbox"/> Exclude	<input checked="" type="checkbox"/> Descendants	<input checked="" type="checkbox"/> Mapped
<input checked="" type="checkbox"/>	24609	302866003	Hypoglycemia	Condition	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Remove Selected

Add Concepts

Classification

Non-Standard

Standard

Click on Recommend:

🛒 Concept Set #1871394

created by anonymous on 2022-10-12 8:18, modified by anonymous on 2022-10-12 8:19

EXER Hypoglycemia

🔒

✕

📄

🏷️

Optimize

🗑️

Concept Set Expression

Included Concepts 11

Included Source Codes

Recommend

Explore Evidence

Export

Import

Compare

Versions

Messages

Enter the concept set description here

Show 50 entries

Filter: Search...

Showing 1 to 1 of 1 entries

Previous 1 Next

<input type="checkbox"/>	Concept Id	Concept Code	Concept Name	Domain	Standard Concept Caption	<input type="checkbox"/> Exclude	<input checked="" type="checkbox"/> Descendants	<input checked="" type="checkbox"/> Mapped
<input checked="" type="checkbox"/>	24609	302866003	Hypoglycemia	Condition	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Remove Selected

Add Concepts

Classification

Non-Standard

Standard

We can see some relevant concepts!

Concept Set #1871394
created by anonymous on 2022-10-12 8:18, modified by anonymous on 2022-10-12 8:19

EXER Hypoglycemia

Concept Set Expression Included Concepts (11) Included Source Codes Recommend Explore Evidence Export Import Compare Versions Messages

Show columns Copy CSV Show 50 entries Filter: Search...

Showing 1 to 10 of 10 entries Previous 1 Next

	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vocabulary										
SNOMED (77)										
Invalid Reason										
Valid (77)										
Class										
Clinical Finding (72)										
Observable Entity (3)										
Procedure (2)										
Domain										
Condition (71)										
Observation (4)										
Procedure (2)										
Relationship										
Lexical via standard (48)										
Ontology-parent (20)										
Lexical via source (13)										
Patient context (10)										
	<input checked="" type="checkbox"/>	45768449	706882009	Hypertensive crisis	Clinical Finding	417,953	32,794,551	Condition	SNOMED	1
	<input checked="" type="checkbox"/>	4098604	190855004	Hypomagnesemia	Clinical Finding	10,179,029	10,199,536	Condition	SNOMED	1
	<input checked="" type="checkbox"/>	4029423	237633009	Hypoglycemia due to diabetes mellitus	Clinical Finding	152,801	5,193,289	Condition	SNOMED	4
	<input checked="" type="checkbox"/>	45757363	120731000119103	Hypoglycemia due to type 2 diabetes mellitus	Clinical Finding	3,775,853	3,878,312	Condition	SNOMED	3
	<input checked="" type="checkbox"/>	439233	212545006	Poisoning by antidiabetic agent	Clinical Finding	479,506	480,821	Condition	SNOMED	1
	<input checked="" type="checkbox"/>	4096804	190448007	Drug-induced hypoglycemia without coma	Clinical Finding	177,278	177,278	Condition	SNOMED	3
	<input checked="" type="checkbox"/>	380688	267384006	Hypoglycemic coma	Clinical Finding	151,841	174,301	Condition	SNOMED	2
	<input checked="" type="checkbox"/>	4226798	421725003	Hypoglycemic coma due to diabetes mellitus	Clinical Finding	23,893	153,078	Condition	SNOMED	2
	<input checked="" type="checkbox"/>	36714116	719216001	Hypoglycemic coma due to type 2 diabetes mellitus	Clinical Finding	101,728	101,728	Condition	SNOMED	2
	<input checked="" type="checkbox"/>	4048805	230796005	Non-diabetic hypoglycemic coma	Clinical Finding	22,160	22,160	Condition	SNOMED	1

Showing 1 to 10 of 10 entries Previous 1 Next

You can add them to the concept set and iterate as described in the option 1.