

### Semantic Data Platform for Healthcare

ICT-611388

D4.2 Terminology Management Platform

WP4- Terminological resource acquisition

V1.1 Final

Lead beneficiary: **EMC & Averbis** Date: **09/01/2015** Nature: **Prototype** Dissemination level: **PU** 



#### TABLE OF CONTENTS

D	OCUME		3
D	OCUME	ENT HISTORY	3
D	EFINITIO	ONS	4
E	XECUTI	VE SUMMARY	5
K	EY WOF	RDS (WORDLE STYLE)	6
1.	INTE	RODUCTION	7
2.	TER	MINOLOGY MANAGEMENT PLATFORM	7
	2.1.	PROTOTYPE DATA	7
	2.2.	AVAILABLE TERMINOLOGIES	7
3.	USE	R MANUAL	9
	3.1.	USER INTERFACE	9
	3.2.	SEARCH	10
	3.3.	HIERARCHICAL NAVIGATION	11
	3.4.	CONCEPT DETAILS	12
	3.5.	EDITING EXISTING TERMINOLOGIES	14
	3.6.	EXPORT OF TERMINOLOGIES	18





# 2

#### TABLE OF FIGURES

FIGURE 1: USER INTERFACE	9
FIGURE 2: EXAMPLE OF THE AUTOSUGGEST FUNCTION	
FIGURE 3: SEARCH RESULT FOR "ST SEGMENT"	
FIGURE 4: SEARCH RESULT USING THE WILDCARD CHARACTER IN "ISCHEM*"	11
FIGURE 5: PART OF THE HIERARCHY OF THE SEMCARE TERMINOLOGY	12
FIGURE 6: CONCEPT DETAILS FOR "ST SEGMENT"	13
FIGURE 7: CONCEPT DETAILS: RELATIVES	13
FIGURE 8: CONCEPT DETAILS: XREFS	14
FIGURE 9: IMPORT TERMINOLOGY	14
FIGURE 10: DIALOGUE WINDOW: IMPORT	15
FIGURE 11: CONCEPT HIERARCHY: MANUAL CHANGES	16
FIGURE 12: EDITING CONCEPT HIERARCHY	16
FIGURE 13: ADD CONCEPT	17
FIGURE 14: PROPOSING SYNONYMS	17
FIGURE 15: EDITING TERMS	
FIGURE 16: QUICK EXPORT	19
FIGURE 17: CHOOSE PROFILE	19
FIGURE 18: EXPORT FILTERS	20
FIGURE 19: MANUAL SETTING OF EXPORT MODES	20
FIGURE 20: AVERBIS CONCEPT MAPPER EXPORT	21



# semcare

D4.2 – Terminology management platform WP4: Terminological resource acquisition Dissemination level: Public Author: C. Riede (Averbis), J. Kors (EMC) Version: 1.1 Final

# DOCUMENT INFORMATION

Grant Agreement Number	ICT-611388		Acronym		SEMCARE				
Full title	Semantic Dat	Semantic Data Platform for Healthcare							
Project URL	www.semcare	www.semcare.eu							
EU Project officer	Saila Rinne (	Saila Rinne ( <u>Saila.Rinne@ec.europa.eu</u> )							
Deliverable	Number 4.2 Title		Terminology	Terminology management platform					
Work package	Number	4	Title	Terminologic	urce acquisition				
Delivery date	Contractual	31/12/1	4	Actual	09/0	1/2015			
Status	Version V1.1	Final		Draft 🗆 🛛 F	Draft 🗆 Final 🗹				
Nature	Report D Pr	ototype	☑ Other <b>□</b>						
Dissemination Level	Public 🗹 Co	nfidentia							
L									

Authors (Partner)	C. Riede (	Averbis), J. Ko	rs (EMC)	
Deepensible Author	Claudia Riede		Email	Claudia.riede@averbis.com
Responsible Author	Partner	AVERBIS	Phone	+49 761 203 97690

# DOCUMENT HISTORY

NAME	DATE	VERSION	DESCRIPTION
C.Riede (AVERBIS)	08.11.2014	0.1	First draft
X. Wang (SGUL)	14.11.2014	0.2	Review and comments
J. Kors (EMC)	07.01.2015	1.0	Review
C. Riede (AVERBIS)	08.01.2015	1.1	Final version





# DEFINITIONS

 Partners (also named as beneficiaries) of the SEMCARE Consortium are referred to herein according to the following codes:

AVERBIS - Averbis GmbH (Germany) Coordinator

EMC - Erasmus Universitair Medisch Centrum Rotterdam (Netherlands) - Beneficiary

MUG - Medical University of Graz (Austria) - Beneficiary

SGUL - Saint George's University of London (UK) - Beneficiary

SYNAPSE - Synapse Research Management Partners S.L. (Spain) - Beneficiary

- Project: The sum of all activities carried out in the framework of the Grant Agreement.
- **Consortium**: The SEMCARE Consortium, conformed by the above-mentioned legal entities.
- AEP: Averbis Extraction Platform; text analysis tool to extract information units such as facts and relations from unstructured text
- EHR: Electronic health record; clinical data record of a patient.
- **Terminology:** General term for information artefacts that provide controlled terms for a domain, identifiers of meaning and semantic relations. e.g. SNOMED, ICD-10, MeSH
- **Term Browser:** Tool provided by Averbis to load, view, modify and export terminologies. It can also be used to create new terminologies.
- UMLS: Unified Medical Language System, metathesaurus.
- **OBO:** OBO format is the text file format used by OBO-Edit, the open source, platform-independent application for viewing and editing ontologies.
- **SKOS:** Simple Knowledge Organization System. SKOS is part of the Semantic Web family of standards
- LOINC: Logical Observation Identifiers Names and Codes. LOINC was initiated in 1994 by the Regenstrief Institute and provide a universal code system for tests, measurements, and observations.
- SNOMED CT: systematically organized computer processable collection of medical terms
- **UIMA:** Unstructured Information Management Architecture; framework by Apache enabling the generation of analysis pipelines for arbitrary content such as text, image or video data





# EXECUTIVE SUMMARY

The aim of the Terminology Management Platform is to provide a user-friendly software which can support the management of terminologies needed within the SEMCARE project. This document presents a user manual explaining how to import terminologies into the platform, how to browse the terminology, how to add or modify terms and how to export the terminology for usage in the text analytics pipeline and the integrated search platform. Besides the SEMCARE terminology, which has been developed especially for the use case, further terminologies which are also used for the text analytics pipeline or support the refinement of the SEMCARE terminology are additionally available in the Terminology Management Platform. They are described shortly in this document.





6

#### KEY WORDS (Wordle style)<sup>1</sup> based use Manual w referenced German symbol Search appropriate Click results case corresponding SKOS \_ platform text listed downloader setting optional lower cking IS e name 🖥 entered existing version created mapper direct details see <sup>2</sup> information children included User editing to ots Choose Additional side added nica a **O**O m edited structure ID Management Butch import support parents CSV different marked empty following formats ata 0 $\mathbf{\omega}$ arci modus mport led function dictionary llera ar a Quic e displayed P preferred DrO Ð multilingual medication patient-leve type Add ٠Ĕ attributes mod containing dialogue OBO def mining

<sup>1</sup> <u>http://www.wordle.net/</u>.



© Copyright 2014-2015 SEMCARE Consortium. This project has received funding from the European Union's Seventh Programme for research, technological development and demonstration under grant agreement No 611388.



# 1. Introduction

The research project SEMCARE aims to build a semantic data platform to support case identification for clinical studies and to facilitate clinical decision making based on patient-level criteria (e.g., age, gender, diagnosis, indication, symptoms, lab results).

Most patient-level data can be found in unstructured text documents. Text mining tools will be used to extract and convert relevant data. These text-mining tools require multilingual terminologies in English, German and Dutch to standardize and address specific idiosyncrasies of medical language.

The SEMCARE Terminology Management Platform is using the Averbis Terminology Platform, called Term Browser. It provides software services to support the management of the terminologies used for the SEMCARE text analytics pipeline. It is possible to import terminologies into the platform, to browse easily the terminologies, to add and modify concepts and terms. The modified terminology can be exported for utilisation in the text-mining platform.

# 2. Terminology Management Platform

#### 2.1. Prototype Data

The prototype is available via the internet with the URL <u>http://www.termbrowser.de.</u>

Username and password will be sent on request, please send a message via the SEMCARE homepage (<u>http://semcare.eu/contact/</u>).

#### 2.2. Available Terminologies

The partners of the SEMCARE project have agreed on an initial use case 'Risk Stratification and Differential Diagnosis of Patients suffering from transient loss of consciousness'. Based on the topic of interest as defined in deliverable D3.2 Requirement specification<sup>2</sup> a distinct SEMCARE terminology has been created. Further terminologies are used to enrich the SEMCARE terminology but are also needed to cover laboratory data, medication and diagnosis.

The following terminologies are currently available in the Terminology Management Platform:

<sup>&</sup>lt;sup>2</sup> D3.2: Requirements Specification, chapter 2.1.3.





8

UMLS SEMCARE:

multilingual terminology provided by EMC<sup>3</sup>

• SEMCARE:

terminology created for the defined use case. The UMLS SEMCARE terminology has been used to enrich the vocabulary. Additional concepts and terms have been added based on the free-text phrases found in the test documents. The terminology is multilingual, for all concepts English, German and Dutch terms are included.

• semCare Modifier:

list of descriptive and evaluative qualifiers, which are linked to the corresponding concepts in the SEMCARE terminology by cross references. The terminology is multilingual, for all concepts English, German and Dutch terms are included.

• Labor 2.0:

terminology for laboratory parameters developed by Averbis. Its hierarchical structure is based on the hierarchy used in LOINC. The terminology is multilingual, for all concepts English, German and Dutch terms are included.

• SNOMED CT:

English version from 2014

• ICD AHA 2014:

German version of ICD-10-GM 2014, enriched by Averbis with further synonyms in order to use it for text mining purposes.

Further terminologies can be imported into the Terminology Management Platform, especially terminologies covering the medication in the different countries.

<sup>&</sup>lt;sup>3</sup> D4.1: Initial multilingual biomedical terminology.





# 9

# 3. User Manual

The terminology management platform offers both reading and editing functions. Terminologies can be exported in different formats, e.g., OBO<sup>4</sup> or SKOS<sup>5</sup>.

#### 3.1. User Interface

atp Export Admin ShowLogs Abo	out logged in as: semcare (click to logout)										
	ay	verbis	3. Coronaries		Q Sea	rch + Limit 25	•	10			
Terminologies 1	Hierarchy - SEMCARE 1.0 (xyz)		۰	earch re	sults (1)						
CD10 AHA 2014 1	Modern Classical	2		Concept	D	4	Label		Ter	minology	(
Octabor 2.0 II	SEMCARE 1.0			2LemLN	1		normal	coronary arteries	SE	MCARE 1.0	
ISSUMED 2001 <sup>- G</sup> LI GS SYOMED 2001 R2 <sup>- G</sup> LI S YOMED 2001 R2 <sup>- G</sup> LI S UMLS SEMCARE 10 <sup>- G</sup> LI	Panly hatos of disorder      Connay angiogram obsenable     Dectocardogram obsenable     With Signs     Pracedure     Substance			1 - 1 von 1 Ergebnissen © Concept details (SEMCARE 1.0 (vpz)) 21.emL/NI			5	normal coronary arteries			د : 2
				Term	(Synonyms (9) Pres	twine (17.0) 887	Attributes (0)	🖉 Anda (1) 🔿 Linese Data 🗄 ima	yes 🔝 Settings	ber.	4-14
				= 0	coronaire vaten		~	(ns		imported from	
				= 0	coronairvaten					anomed_an Imported from	
				-	coronation					shomed_en	
					Coronanes					data di	
					Coronary anenes					union ( )	
				as ()	coronary arteries					shomed_en	-
				0	Coronary artery					default	
				• •	Herzkranzgefaße					default	
				• •	Koronararterien					cetaut	
Echocardogam obsensée     Decinardogam obsensée     Decinardo		imported from umits_sempare									
	English # Setting time 25 *	W Hide IOs								inology CARE 1.0	

Figure 1: User Interface

Description of the areas:

- List of the available terminologies. The terminologies to be used in the search can be selected by ticking the box before the name of the terminology. By clicking on the name of the terminology this terminology can be browsed in the area rightwards.
- 2. The hierarchical structure of a selected terminology. The result of a free-text search will be shown here in its hierarchical position.
- 3. The search box which allows a free-text search.
- 4. List of the search results.

<sup>5</sup> Simple Knowledge Organization System: <u>http://www.w3.org/2004/02/skos/</u>.



<sup>&</sup>lt;sup>4</sup> OBO Flat File Format: <u>http://oboformat.googlecode.com/svn/trunk/doc/GO.format.obo-1\_4.html</u>.



5. Details of the concept.

An area can be faded out in order to enlarge the disposable space for the remaining areas by clicking on <sup>(()</sup>) top right.

#### 3.2. Search

A basic functionality of the Terminology Platform is the possibility to search terms via a free-text search in different terminologies. All terminologies which are marked in the left area *Terminologies* are included into the search. The search covers all synonyms of a concept.

#### Search box

The search box allows a free text search. During the entry of the word the AutoSuggest function helps to find the desired term.

ischem Q Search 1	<ul> <li>Limit 25</li> </ul>
Ischemia [SNOMED 2014 02: 52674009]	
Ischemia [UMLS SEMCARE 1.0: C0022116]	
Ischemia (disorder) [SNOMED 2014 02: 52674009]	
Ischemia (disorder) [UMLS SEMCARE 1.0: C0022116]	
Ischemia and infarction of kidney [UMLS SEMCARE 1.0: C0495064]	
Ischemia Anoxia, Brain [UMLS SEMCARE 1.0: C0752306]	
Ischemia Anoxia, Cerebral [UMLS SEMCARE 1.0: C0752307]	
ischemia bowel [UMLS SEMCARE 1.0: C0156149]	
Ischemia bowel [UMLS SEMCARE 1.0: C2004435]	
ischemia brain [UMLS SEMCARE 1.0: C0007786]	
Ischemia cerebral [UMLS SEMCARE 1.0: C0917798]	
Ischemia cerebrovascular [UMLS SEMCARE 1.0: C0917798]	

Figure 2: Example of the AutoSuggest function

#### Search results

The search results are listed on the right hand side.

ST segment	Q Se	earch 💌 Limit 10 🔹			
٩	Search results (2)				(
	ConceptID		Label	Terminology	C
	C0429029		segment st	UMLS SEMCARE 1.0	-
	L2AqQkS		ST segment	SEMCARE 1.0	
					-
	1 - 2 von 2 Ergebnissen				<1.)



Displayed are the ID and preferred term of the concepts and the name of the corresponding terminology. By clicking on one of the search results more details of the selected concept are shown in the area





below and additionally the hierarchical tree structure of the terminology is expanded to show the position of the concept.

The asterisk ("\*") can be used as a wildcard character to expand the search.

ischem*	Q Search	Limit 25		
۹	Search results (52)			۲
	ConceptID	Label	Terminology	0
	14	ischemic	semCare Modifier 1.0	-
	IBq1eOR	Ischemic Heart Disease	SEMCARE 1.0	
	10076004	Ischemic muscular atrophy (disorder)	SNOMED 2014 02	E
	129611009	Ischemic peripheral neuropathy (disorder)	SNOMED 2014 02	
	13026000	Ischemic stricture of intestine (disorder)	SNOMED 2014 02	
	13954005	Ischemic ulcer (disorder)	SNOMED 2014 02	
	140911000119109 1 - 10 von 52 Ergebnissen	Ischemic stroke with coma (disorder)	SNOMED 2014 02	

Figure 4: Search result using the wildcard character in "ischem\*"

#### 3.3. Hierarchical Navigation

Besides searching for a specific term it is possible to navigate through the hierarchic structure of a terminology. By clicking on the name of a terminology in the left area *Terminologies* this terminology is displayed in the area *Hierarchy* opening the highest level of the hierarchical tree. More detailed sub-trees can be opened by clicking on the "+" symbol before the appropriate superordinate concept.





Hierarchy - SEMCARE 1.0 (xyz)	٩
Modern Classical	
E- CSEMCARE 1.0	
🔄 🗔 Clinical finding	
🔄 🗔 Cardiovascular finding	
- 📰 Atrial fibrillation	
— 📰 Bradycardia	
— 📰 Brugada syndrome	
- 🖅 Bundle Branch Block	
🖅 🛅 Electrocardiogram finding	
⊕⊢ 🗂 Heart Failure	
🖅 💼 Ischemic Heart Disease	
🕀 💼 Myocardial infarction	
<ul> <li>— E Pulmonary Hypertension</li> </ul>	
- E Stenosis	
— 📰 Sudden cardiac death	
— 📰 Syncope	
🖅 🚞 Neurological finding	
- Cobservable entity	
i ⊡- 🛅 Substance	
English 💌 Sibling limit 25 💌 🗹 Hide IDs 🔂 Add (root) concept	

Figure 5: Part of the hierarchy of the SEMCARE terminology

If the selected terminology is multilingual, the language can be selected below. By default the concept IDs are not visible; this can be changed by removing the tick besides *Hide IDs*. The classical view is only available for monolingual terminologies.

#### 3.4. Concept Details

For each concept there is additional information available, called *Concept details*, which are displayed in the lower right area.





2AqC	QkS	ST segme	ent			5
o Te	erms/Sy	nonyms (12) 🗗 Relatives (1 / 0) 🔡 Attributes (0) 🔗 Xrefs (2)	C Linked Data	III Settings		
	Pref	Label -	Xrefs	Attributes	Туре	(
	0	electrocardiogram: ST segment			imported from umls_semcare	
	۲	ST-segment			imported from umls_semcare	
	$\circ$	ST-Strecke			imported from umls_semcare	
20 20 20 20	۲	ST segment			imported from umls_semcare	1
	0	ST				
	۲	Elektrokardiogramm ST-Strecke			imported from umIs_semcare	
	0	Electrocardiogram ST segment			imported from umls_semcare	
	0	ECG ST segment			imported from umIs_semcare	



In the first tab, **Terms/Synonyms**, all terms or synonyms of the concept are listed. For each term the language, the type and possible attributes are specified. In each language a preferred term can be defined, this is indicated in the column *Pref*. Term attributes provide additional information in a typified format, where types can be defined by the user. E.g., common typing errors can be added to a concept as synonyms and can be marked by the attribute "typing error".

In the second tab, **Relatives**, the direct relatives of the concept within the terminology are listed.

/R7fyMb			Electroo	card	liogram observa	ble		<u> </u>	5	
<b>o:</b> Terms/	/Synonyms (4) 🗗 Relat	ives (1 / 8)	(0) 🔗 Xrefs (0)	C	Cinked Data	🖥 Images 🛛 🔝 Settings				
Туре	Label	Concept ID			Туре	Label	Concept ID			
s-a	Cardiovascular	tZ6GLKZ	Ŷ	*	is-a	ST segment	L2AqQkS	Ŷ		
	observable				is-a	QT interval duration	f2DnPA8	Ŷ		
					is-a	PR Interval Duration	VGzKgzj	Ŷ		
					is-a	Q wave	1MXCUkY	Ŷ		
						is-a	T wave	8ovlPKf	Ŷ	
					is-a	ST segment duration	3J3UgVk	Ŷ		
					is-a	Q wave height	EtfqNCt	Ŷ		
					is-a	QRS duration	PF9itOW	Ŷ		
				-						







The direct parents are listed in the left table and the direct children are listed in the right table. With a click on the pin symbol next to a concept ID it is possible to switch to the corresponding concept.

**Concept Attributes** can be used for specifying additional information on concept level, e.g. describing the scope of the concept ("Ejection fraction (MRI)": only to be used when the ejection fraction is measured by a MRI).

In the tab **Xrefs** the relations with concepts in other terminologies are listed. The relationship between the concept and the referenced terminology can be described by the type. If no type is specified the terminology is containing the same concept.

⊙ Conce	t details (SEMCARE 1.0 (xyz))				
L2AqQkS	ST segment			^ _ <b>ይ</b>	
.; <b>o</b> : Term	(Synonyms (12) 🗗 Relatives (1 / 0) 🔠 Attributes (0) 🔗 Xrefs (2) 🔿 Linked Data 🗗 Images	III Settings		Ŧ	
Туре	Label	ID in target	Namespace		0
relatedTo	depressed - elevated	27	semCareModifier	Ŷ	*
	segment st	C0429029	umls_semcare	Ŷ	

Figure 8: Concept details: Xrefs

If the referenced terminology is part of the Terminology Management Platform it is possible to go directly to the referenced concept by clicking on the pin symbol.

The tabs *Linked Data*, *Images* and *Settings* can be used in prospective versions to display information form external sources or images.

#### 3.5. Editing existing terminologies

#### Import

Users with the appropriate rights can import additional content into an existing terminology under the menu item *Admin*.



Figure 9: Import Terminology





In the opening dialogue window the target terminology and the source file have to be chosen. Additionally the existing terminology can be emptied before import. Allowed import formats are csv and OBO.

Averbis Terminology Platform Import content into a ter Please choose a OBO or CSV file. Th	minology he values will get appended to the chosen terminology.
Choose target	SEMCARE 1.0 (xyz)
Empty terminology before import	
Choose source file (*.obo, *.csv)	Browse
0 %	L Import ① Stop 🛞 Close

Figure 10: Dialogue window: Import

The fields of a source file in csv format must be separated by ";". The following fields are expected:

Table 1: Import fields

ConceptID	TermID	TermLabel	TermType	Language	isPreferred	parentConceptIDs	MappingForms <sup>6</sup>
ID of the concept	ID of the term (can be empty)	label of the term	type of the term (if empty then type = default)	identification code of the language, e.g. de, en	1 = preferred term 0 = no preferred term	concept ID of the parent concept [optional]	mapping forms of the term [optional}

#### Manual editing

Terminologies can be edited manually provided that the terminology is editable and that the user has write permission.

#### Concepts:

With a right-mouse click on the concept in the hierarchical view the concept can be deleted, its hierarchy

<sup>&</sup>lt;sup>6</sup> if a homonym term has an additional specification which will not be found in a text: (e.g. note <finan.> and note<mus.> ) the text analytics pipeline can use the mapping form (e.g. note)





can be edited or a new child can be added. Below the hierarchical tree there is also a  $\oplus$  – Button to add new concepts.

Electrocardiogra Add child PR Interval D Edit hierarchy Q wave heigh Q waves Remove	
QRS duration	
QT interval duration	
_	
ST segment duration	
— 📰 T wave	
⊞- 🧰 MRI observable	
i - <mark>"</mark> Vital Signs	
🔄 🗔 Cardiovascular procedure	
🗄 💼 Angioplasty	-
English Sibling limit 25 V Hide IDs Add (root) concept	

Figure 11: Concept hierarchy: Manual changes

• Editing concept hierarchy

Averbis Terminology Platform			8
Edit concept hierar	rchy		
Concept ID:	tfY6RG4		
Preferred term label:	Heart Rate		
Current parents:	bCahZp2: Vital Signs 🟁		
Add parent:	Bra Braducardia (6d24x1c)	🕀 Add	
Current children:	Bundle Branch Block (g223G1B)		
Add child:		🕀 Add	
	(	OK & Clos	е

Figure 12: Editing concept hierarchy

Both parents and children can be added. In the fields after *Add parents* and *Add child* any term of the desired concept can be entered. The system will suggest applicable concepts. Existing parent or children relationships can be deleted by clicking on the red X.





#### New concepts:

Averbis Terminology Pla	tform 😵
Add concept	(SEMCARE 1.0)
Child of:	Vital Signs
Concept ID:	Autogenerated if empty
Preferred term	label: Weight
Language:	English
Suggest synony	/ms: 🗹
	OK Scancel

Figure 13: Add concept

The concept ID can be entered manually, otherwise the ID is generated automatically.

By ticking the box *Suggest synonyms* all synonyms from terminologies that were marked as selected in the left area of the Terminology Platform, are proposed. The terms can be selected individually. The checkmark on concept level inserts an XREF to the referencing terminology.



Figure 14: Proposing synonyms





#### Terms

The terms of a concept can be edited in the Concept details area.

2Aq	QkS		ST segment			⊗		ጜ
<b>o</b> 1	erms/S	Synonyms (12)	) & Xrefs (2) 🔿 Linked Data 🗗	Images 🖽 Settings			•	
	Pref	Label	Xrefs	Attributes	Туре			
	0	ECG ST segment			imported from umls_semcare	Ø	۳,	⊗
	0	Electrocardiogram ST segment			imported from umls_semcare	Ø	×	⊗
	0	electrocardiogram: ST segment			imported from umls_semcare	Ø	۳,	⊗
	0	elektrocardiogram ST-segment			imported from umls_semcare	Ø	۳,	⊗
-	۲	Elektrokardiogramm ST-Strecke			imported from umls_semcare	Ø	۳,	⊗
	0	segment st			imported from umls_semcare	Ø	۳ĸ	⊗
	0	segments st			imported from umls_semcare	Ø	*	⊗
	0	ST				Ø	×ĸ	⊗
2112	0	ST segment				120	2	0

Figure 15: Editing terms

to search for synonyms in the terminologies available in the Terminology Platform



to delete the term



to edit the term

Additional synonyms can be entered by clicking on the  $\oplus$  – Button at the bottom right.

#### 3.6. Export of terminologies

Users with the appropriate rights can export a terminology under the menu item Export.



#### **Quick export**

With the *Quick export* function terminologies can be exported in the data formats OBO, SKOS, and csv. The export in csv format includes only the fields that were accepted when the terminology was imported (see Table 1: Import fields).





Ave	erbis Terminology Platform			8
	Export terminology			
	Choose terminology:	SEMCARE 1.0 (xy	z) 💌	
	Full or subtree:	Full	•	
	Restrict to subtrees:			
	Output format:	CSV	•	
	Successfully exported terminolog Download	gy! You can download it	using this link:	×
	100 %	<u></u> 1 E	xport ① Stop	Close 8

Figure 16: Quick export

The exported terminology can be downloaded directly in the dialogue window.

#### **Export profile settings**

For the export of terminologies into a ConceptMapper dictionary which is used in the text mining pipeline the *Export profile settings* modus is required.

First, an export profile has to be defined or an existing profile has to be selected. The name of the selected profile is displayed at the top of the window

	SENCARE		@ Show all mo	idfied terms 🛛 🔮 Erase all modifications 🖉 😒	Close profi
Choose Profile O Ch	anna Tarmintogan & Fitara 🚹 Engott.				
D	Profile label	Separate files	Created		
160	SEMCARE	8	2014-01-14	5	
244	UMLS Semcare	8	2014-09-08	\$	
54	SEMCARE_Modifier	2	2014-10-07	\$	
) Use selected profile	3 Add new profile				

Figure 17: Choose Profile



© Copyright 2014-2015 SEMCARE Consortium. This project has received funding from the European Union's Seventh Programme for research, technological development and demonstration under grant agreement No 611388.



For each export profile at least one terminology has to be allocated. There are several predefined filters which can be applied, e.g. the *Default Filter English* identifies all terms containing acronyms and stop words and sets them to the EXACT modus for the export. Using the Averbis Concept Mapper format for the export the term or concept will then be written into a separate dictionary (see below). The results of a filter can be downloaded as csv file for evaluation instead of applying the filter directly.

	Drop Filter	Remove selected	AbstractConceptsFilter	
ICD 10 AHA 2014		⊖ Remove	♀ ActivateLeafNodes	0.0
O Labor 2.0		⊖ Remove	Custom (Advanced)     DefaultFilter English	
SEMCARE 1.0	♥ DetaulFilter English	(C) Remave	♥ DefaultFilter German	
SemCare Nodifier 1.0		⊖ Remave	DefaultFilter No Numbers     ONB Ignore Short	
SNOMED 2004		⊖ Remove	Synonyms P DNR DND Kurzwort	
B SNOMED 2014 02		(O Remove	♀ GND-Motor-Filter	
O UMLS SEMCARE 1.0		⊖ Remove	♀ GND-Underspecified Terms Filter (No ambiguity check)	
			<ul> <li>♀ Ignore all</li> <li>♥ Ignore Ambiguous Terms</li> <li>♥ FHD specific</li> <li>♥ SWO GEO specific</li> <li>♥ SWO GEO specific</li> <li>♥ Underspecifice/TermsFilter</li> </ul>	

Figure 18: Export filters

After closing the export dialogue window (clicking on *Close dialog* at the bottom of the window), the export mode of a term or a concept can be set.

⑦ C	oncept	details (SEMCARE 1.0 (xyz))				1		
PF9i	tOW	QRS	duration	DEFAULT		-	8U	û <b>₽</b>
;o:	Ferms/S	Synonyms (4) 🖓 Relatives (1 / 0)	Attributes (0) 🔗 Xrefs (2) 🖒 Lin	iked Data 🐻 Images 🖽 Setti	ngs			
	Pref	Label	Xrefs	Attributes	Туре	Amt	Dict mode	Override O
	0	QRS complex duration			imported from umIs_semcare		DEFAULT 💌	*
	۲	QRS duration					DEFAULT 🔻	
_	۲	QRS Komplex Dauer			default		DEFAULT 🔻	
=	۲	QRS-duur					DEFAULT	
							EXACT	
							IGNORE	







By setting the export mode to IGNORE the term or concept will not be included into an export. The IGNORE mode can be extended to subordinate concepts if the tick box on the right side of the dropdown menu on concept level is marked.

By setting the export mode to EXACT the term or concept is exported into a separate dictionary. The text mining pipeline will use the terms from this exact dictionary only in this exact form and allows no variations like differing lower or upper case or grammatical inflections.

For the export the following formats are available: Averbis Concept mapper Format, Averbis MSI Format, Averbis Lemmatizer Format, OBO, SKOS and csv. The text analytics pipeline expects the Averbis Concept mapper Format, which considers all three export modes.

ath	Export Admin ShowLogs About	logged in an: servcare (	ck to logout) - Current export Prolife: SEMCARE (160)		
	Currently Used Profile SEM	@ Show all modile	ed terms 🔮 Erase all modifications 🔞 Close profile		
	Charge Darthe O'Charge Tarana	per KT Bert			
Teer	Choose export format				
	Averbis Concept Mapper 3 Format	n der AEP besutzen, ist das ihr bevorzugtes Format.	averbis		
	Averbis MSI Format	Das Averbis MSI Form	wrbis Terminology Platform		•erbis
	Averbis Lemmatizer Format	Das Averbis Lemmal	Please choose the coscepts' annotation type the concepts shall be tagged with. Termisology Concept Type ancestors CIDs Language Term langua	nges	erbis
	O 080	Ores CEID Formatist	SEBICARE (ge averbia setration types Concept	English @ Netherland O Deselect all d Select all	Gene Ontology
	© SKOS	Whe OBO tat auch SK			🔊 S W3C 💒
	© CSV	Der CSV Export liefen	By clicking on "OK" a new dictionary will be created. It will replace the old one in the "current" folder.		
	₫ Export			© ok	Cancel

Figure 20: Averbis Concept Mapper export

In the dialogue window the concept type and the language(s) can be defined. The resulting dictionary is stored automatically in a defined data folder in the Averbis file system, but its current version can be downloaded in the register *Choose Profile*.

Currently Used I	Profile: SEMCARE		Show all modified terms	Brase all modification
Choose Profile 🕲 Choose Terminalogies & Filters: 👌 Export				
ID	Profile label	Separate files	Created	1
160	SEMCARE	Ø	2014 01 14	5 <
-				

