

# Skapa, standardisera och underhålla terminologi

Fodina TerminologyBuilder Suite

TERM **C**ATCH  
TERM **K**LUST  
TERM **O**GRAPH

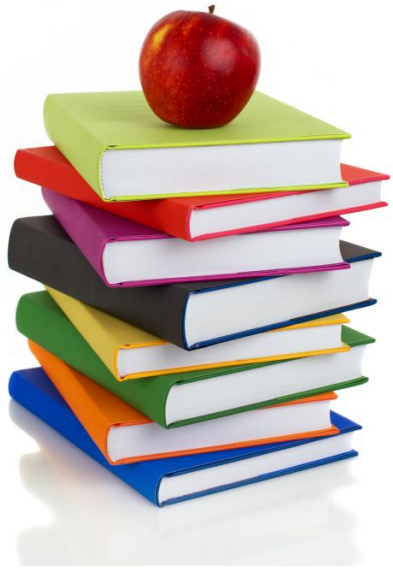


# Fodina Language Technology

- Linköping
- Avknoppning från Linköpings universitet
- Grundat 2004
- Terminologi med språkteknologisk grund
- Kunder främst stora svenska exportföretag



# Fördelar med standardiserad terminologi



- Färre missförstånd bland läsare och översättare
- Minskade översättningskostnader
- Högre kvalitet i original och i översättningar
- Starkare företags-/organisationsprofil
- Bättre sökbarhet och "hittbarhet" (searchability och findability)

# Terminologi på företag

- Tidigare endast Scania som hade terminologer, nu finns det på fler företag: Volvo, Epiroc, Tetra Pak...
- ROI (return-of-investment) måste bevisas.
- Många är medvetna om att de har spretig terminologi och vill göra något åt detta.
- Måste finnas en tillämpning (som skrivstöd, översättningsstöd och integration till andra system, till exempel databaser med artikelbenämningar).
- Enbart publicering av termdatabaser är inte nog.
- Nationella och internationella standarder är inte tillräckliga i dagsläget.
- Intresserade av terminologier med många begrepp och termer (rekommenderade, förbjudna ...)
- Nästan alla större företag skriver originaldokumentation på engelska.
- **Men** ökande intresse:
  - Inte bara för teknisk information, utan även för konstruktion, marknadsföring, m.m.
  - Finns mycket pengar att spara.
  - Kvalitetsbrister kostar.
  - Språk (och terminologi) är en del av företagets "branding".

## Terminologi inom offentlig sektor? Situationen 2018...

- Ett fåtal myndigheter har terminologer, de flesta inte.
- Finns betydligt fler som arbetar med Klarspråk.
- Språklagen från 2009 har haft liten effekt på intresset för terminologi.
- Terminologiska projekt är ofta avgränsade och rör ett begränsat antal begrepp.
- Flera myndigheter publicerar termer, men väldigt få har integrerat termerna i andra applikationer.
- Börjar intressera sig för ROI, men ekonomiska argument har historiskt sett inte varit lika viktiga som i industrin.
- Svårt att få myndigheter att samarbeta med varandra?

## Skapa en standardiserad terminologi och använda den

- Termerna finns redan i dokument, på webbplatser och i "gamla" termlistor.
- Behöver "hittas" och struktureras
  - Vad är viktigast?
  - Vad är inkonsekvent?
  - Nyckelord?
- Beslut måste tas
- Se till att terminologin används

# Fodina TerminologyBuilder Suite

## Fodina TERMATCH

### Extrahera TERMER

Identifiera fraser och uttryck som verkar vara termer från ert eget innehåll:

- Dokument
- Presentationer
- Websidor
- ...

## Fodina TERMKLUST

### Klustra TERMER

Gruppera termer som är synonymmer eller stavningsvarianter i synonymgrupper.

## Fodina TERMGRAF

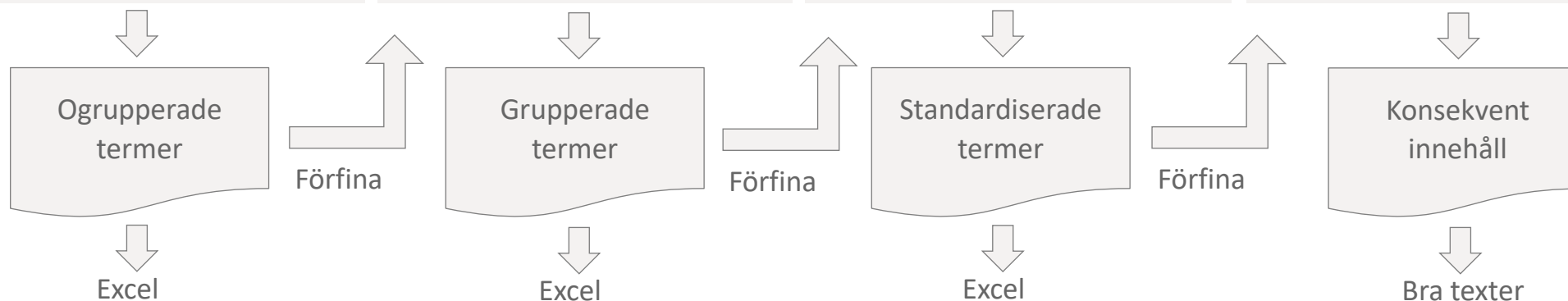
### Ta beslut om TERMER

Granska synonymgrupper i ett grafiskt gränssnit och bestäm vilka termer som ska rekommenderas och vilka som bör undvikas.

## Acrolinx

### Använda TERMER

Assistera skribenter så de använder rekommenderade termer med ett skrivstöd.



Går att använda verktygen var för sig men också hela kedjan.

# Termextraktion med TermCatch (Bolagsverket)

**Numeric Characters**  include  
**Special Characters**  include  
Included Files 47/47  
Hiding Discarded

**Min Frequency** : 1  
**Num Tokens** : 1  
=< = = >=

**Regex Filter:**  
PostgreSQL Regex Expression

**Labels:**  
Search Labels

**POS:**  
Search POS Tags

**Actions:**

Gazetteer Name	Type	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
sv_NounNoSingleTermLemma.gzt	exact	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
sv_NoNounStopWordLemma.gzt	endswith	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
sv_NounStopWordLemma.gzt	endswith	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Copy (page) Excel (page) Excel TSV PDF (page) Show 10 rows Search:

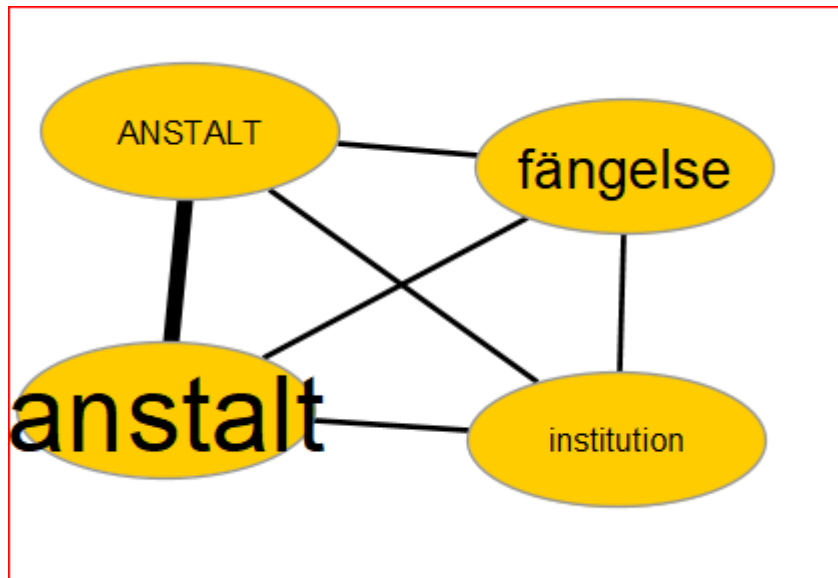
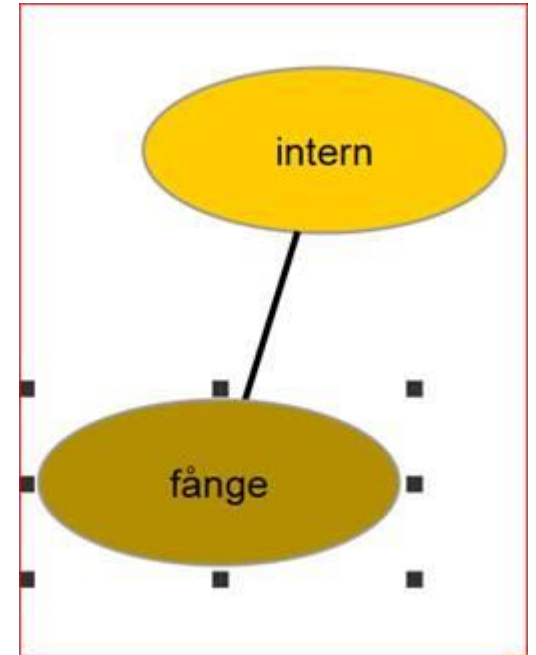
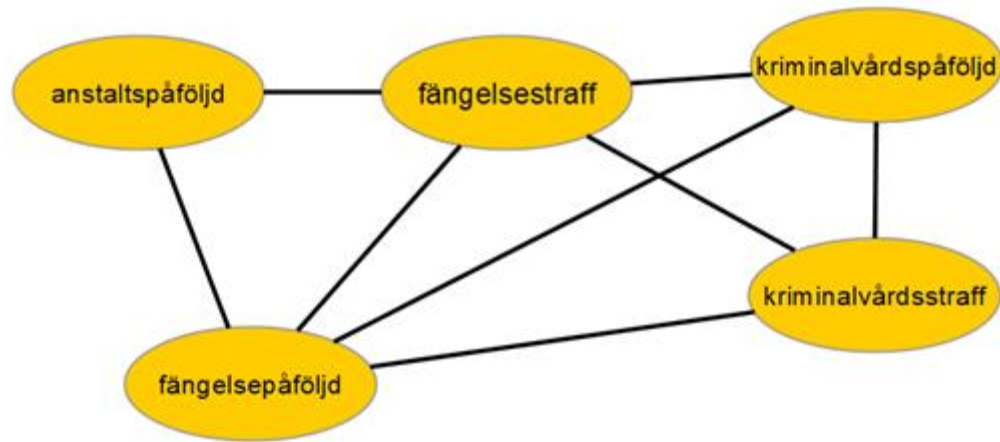
ID	Term	Root	POS	Frequency	
46012	årsredovisning	årsredovisning	NOUN	342	<input type="checkbox"/>
28737	företag	företag	NOUN	328	<input type="checkbox"/>
28802	tjänst	tjänst	NOUN	309	<input type="checkbox"/>
48427	föreningsstämma	föreningsstämma	NOUN	192	<input type="checkbox"/>
45664	bolag	bolag	NOUN	182	<input type="checkbox"/>
46833	förening	förening	NOUN	179	<input type="checkbox"/>
45707	aktiebolag	aktiebolag	NOUN	153	<input type="checkbox"/>
48656	coronavirus	coronavirus	NOUN	144	<input type="checkbox"/>
28852	webbplats	webbplats	NOUN	143	<input type="checkbox"/>
45915	lag	lag	NOUN	135	<input type="checkbox"/>

ID Search Term Search Root Search POS Frequency

Showing 1 to 10 of 2,341 entries (filtered from 2,342 total entries)



# Kriminalvården...



# Standardisering i TermoGraph

☰ TERMOGRAPH by Fodina
ACROLINX

Start Clusters Concepts Review Tools

<4 4-8 >8 All

Filter by working set  
 Wikipedia

Columns  
 Select, ID, Size, Repr...

SeID	SizeRepr	Freqsum
<input type="checkbox"/> 7	15 CO2	370
<input type="checkbox"/> 11	15 manual transmission	201
<input type="checkbox"/> 89	14 automaker	91
<input type="checkbox"/> 133	14 driver's license	63
<input type="checkbox"/> 134	14 drive shaft	66
<input type="checkbox"/> 233	14 sump	61
<input type="checkbox"/> 20	13 Motor Company	145
<input type="checkbox"/> 741	13 permanent resident	25
<input type="checkbox"/> 51	12 semi-automatic transmission	73
<input type="checkbox"/> 111	12 subcategory	66
<input type="checkbox"/> 177	12 railcar	56
<input type="checkbox"/> 268	12 slang	55
<input type="checkbox"/> 294	12 propulsion	52
<input type="checkbox"/> 1022	12 cabriolet	25
<input type="checkbox"/> 79	11 bogie	82
<input type="checkbox"/> 87	11 nitrogen oxide	90
<input type="checkbox"/> 136	11 emission standard	80
<input type="checkbox"/> 260	11 fuel system	65
<input type="checkbox"/> 299	11 alternator	33
<input type="checkbox"/> 450	11 ECU	28
<input type="checkbox"/> 477	11 rail transport	30
<input type="checkbox"/> 643	11 taillight	22
<input type="checkbox"/> 685	11 cycle track	23
<input type="checkbox"/> 23	10 greenhouse gas	195
<input type="checkbox"/> 58	10 fuel efficiency	94
<input type="checkbox"/> 142	10 garbage truck	31
<input type="checkbox"/> 195	10 semi	43
<input type="checkbox"/> 213	10 tow truck	44

Search... VIN    Source All    Field Term

Active cluster    Search Results

ID	Term	Source	Include	Num
10370	<u>provisionary driving licence</u>		+	1
18404	<u>provisional driving licence</u>		+	1
18620	<u>road driving</u>		+	1
19669	<u>VINCH model</u>		+	1
19825	<u>Self-driving car</u>		+	1
19848	<u>vin</u>			1
19849	<u>VIN</u>			1
19850	<u>VIN number</u>			1

🔗 Add Edge

Selected clusters

CLOSE ALL

vehicle identification number

Workspace actions

CREATE TERM

Term operations

DISCARD TERM

Cluster operations

HIDE TERMS

CREATE CONCEPT

Synonym selection

Level 3 of 5

LESS

MORE

Misc

Graph + Table  
 Graph  
 Table

# Standardisering i TermoGraph

**Concepts**  
Filter by Concept Group  
All

1 - 1 of 1

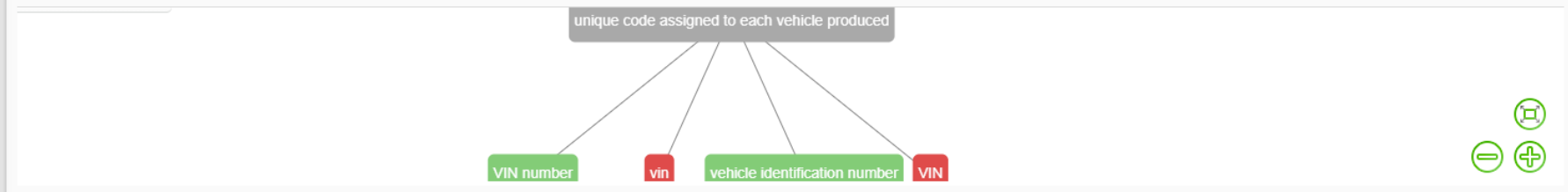
Id	Status	Representative
80727	in_progress	vehicle identification number

Terms in concept #80727  Graph

Representative: vehicle identificatio...

Columns: Term, Status, Doma...

Term	Status	Note	Domain	Frequency	Edit	Remove
<u>vehicle identification number</u>	✓ preferred	...				
<u>VIN</u>	✗ deprecated	...				
<u>VIN number</u>	✓ preferred	...				
<u>vin</u>	✗ deprecated	...				



- Set status for all terms
- Include at least one preferred term
- Include deprecated terms
- Write a note for all admitted terms
- Write a definition (optional)

All tasks completed ✓

Definition: unique code assigned to each vehicle produced

Validators comment:

Comment:

Save as

IN PROGRESS

READY FOR VALIDATION

READY FOR PUBLICATION

Delete

DELETE CONCEPT

# Skapa, standardisera och underhålla termdatabaser

## **Fodina TerminologyBuilder Suite**

- Identifierar termer och synonymer i ditt innehåll
- Hittar och grupperar synonymer och stavningsvarianter
- Underlag för beslut om vilka termer som ska användas och vilka som inte ska användas
- Publicerar standardiserade termer i en termdatabas
- Körs i en webbläsare. Ingen installation krävs.

## Slutord

- Terminologisk kvalitet är viktig
- Kvaliteten på källspråket avgörande för hur bra översättningarna blir
- Behov av en standardiserad svensk myndighetsterminologi
- Mer samarbete mellan myndigheter
- Förutsättningar för att dela terminologi på riktigt (inte bara uppslag i terminologiportaler)
- Flerspråkiga termdatabaser av hög kvalitet behövs om vi vill höja kvaliteten på maskinöversättning.

