



# Language models for Swedish authorities

Language models learn language by reading text

Can be used to solve (all) language processing tasks

Pretrain - finetune - predict



Swedish is a small language and most companies are international

But the **public sector** needs to handle Swedish language data

...and they work primarily with text

# Language models for Swedish authorities project

---

Provide tools and prerequisites to use LMs

Code, data, models, applications

Text similarity (semantic search)

NER

Machine translation

Text categorization





Spin-off project #1:

**SuperLim** - a Swedish evaluation framework for LMs

(cf. SuperGLUE for English)





Spin-off project #2:

**Data readiness factory** - improve data readiness

Accessibility, validity, utility

Annotation, anonymization, evaluation



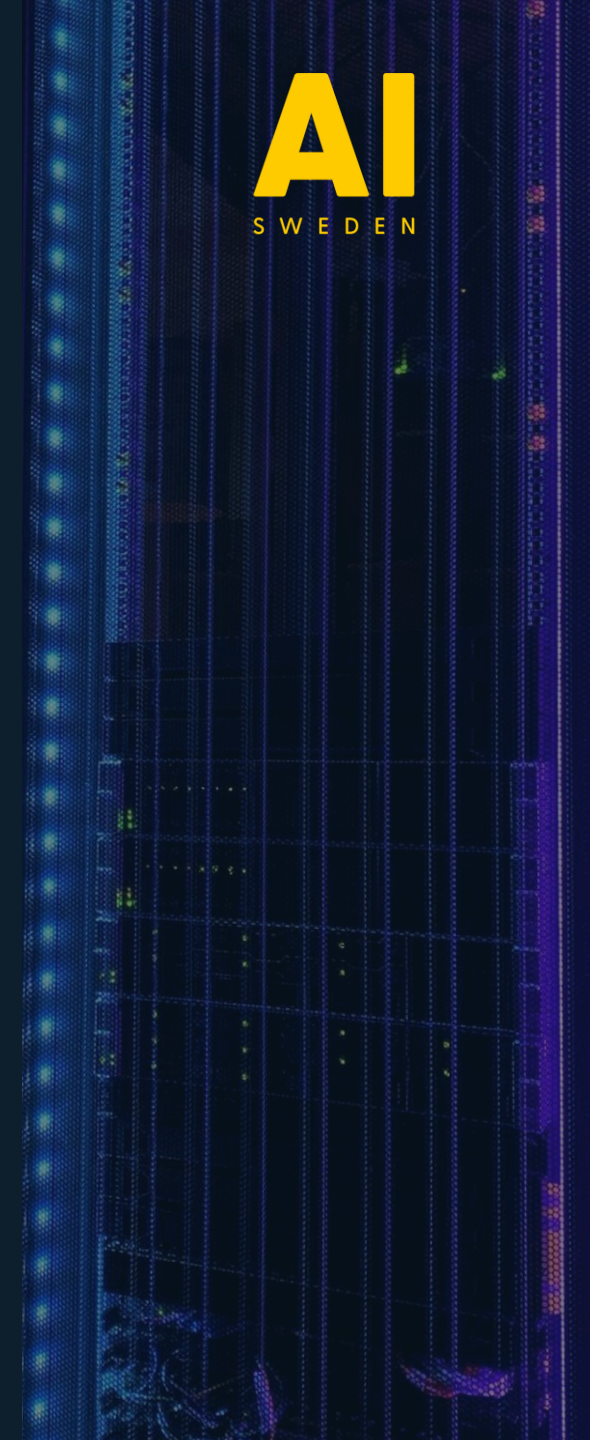
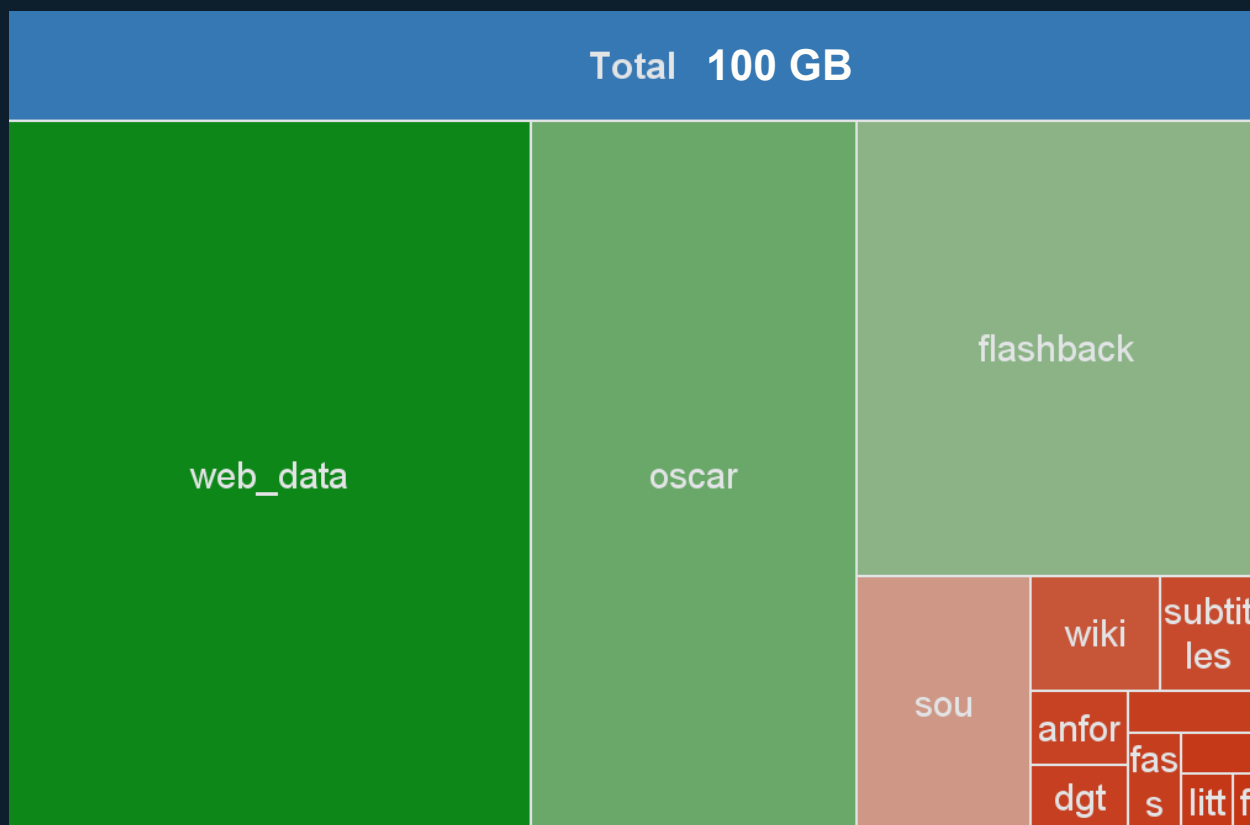
**Large-scale language models for Swedish**

Reduce the need for in-house technical expertise

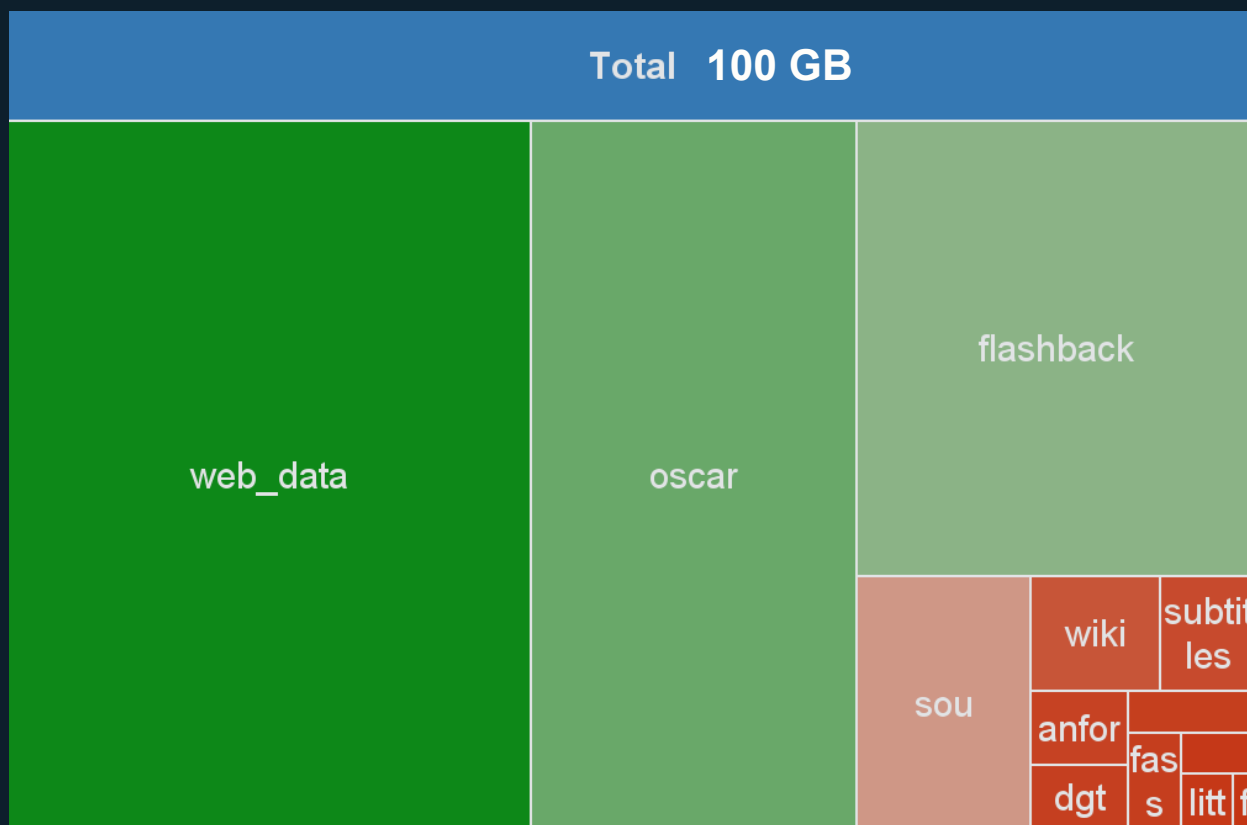
Reduce the need for expensive compute resources

**Pretrain - predict (p-tuning)**



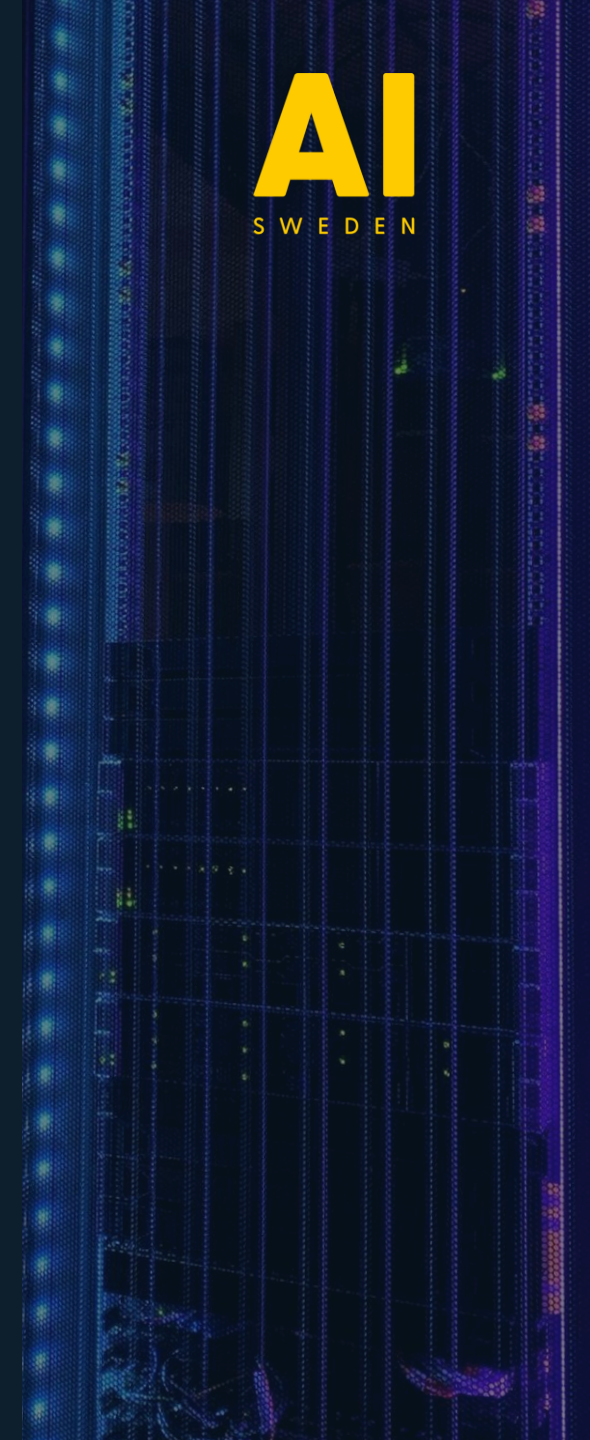




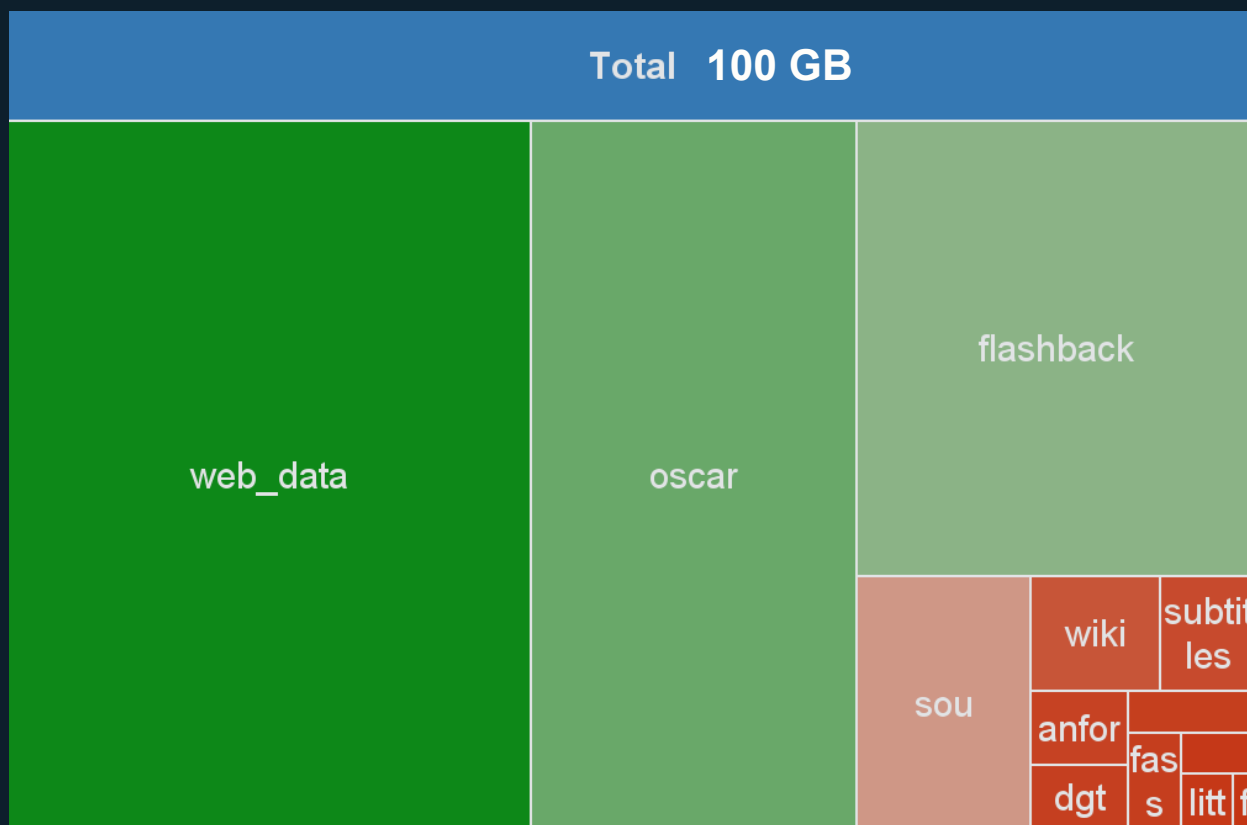


**Layers**            30  
**Heads**             28  
**Seq. len.**        2048  
**Hidden**            3072

**Parameters** 3.5B







**Layers** 30  
**Heads** 28  
**Seq. len.** 2048  
**Hidden** 3072

**Parameters 3.5B**

Model	$PPL_c$ sv	$PPL_c$ en
Flashback-GPT	4.76	6.22
GPT2-xl	4.70	2.30
GPT-NEO	2.87	<b>2.11</b>
GPT-SW3	<b>2.19</b>	2.70





**NVIDIA SuperPOD:  
60 NVIDIA DGX-A100 compute nodes**

**Each DGX-A100:  
8 NVIDIA A100 Tensor Core GPUs**

**Each A100 GPU:  
40 GB on-board HBM2 VRAM**

**Nvidia Mellanox Infiniband networking**



# GPT-SW3 v1

~100 GB data

3.5 B parameters

16 nodes, 2.5 days

# GPT-SW3 v2

~600 GB data

100 B parameters

20 nodes, 90 days



# Pooling Resources

Combining datasets from morphologically similar languages

SENODKIS



## Question & Answering

Answer questions, information extraction, parsing unstructured data and text classification

## Writing aid

Everything from creative writing, copy writing and summarization to grammar correction

## Code related tasks

Translate between programming languages, explain code, write documentation and more

## A general framework

With increased scale come better capabilities and lower barrier to entry for users, pushing the boundaries of applied AI



# GPT-SW3: a foundational resource for Swedish NLP