



# Kraften i ett digitaliserat kulturarv

Karin Grönvall, riksbibliotekarie, juni 2022



Bildning  
och Kritiskt  
tänkande

Integritet

Tradition  
o förnyelse

Öppenhet

Inkludering

Pålitlighet

Transparens

**2 § Biblioteken i det allmänna biblioteksväsendet ska verka för det demokratiska samhällets utveckling genom att bidra till kunskapsförmedling och fri åsiktsbildning.**

INSTRUKTION

FÖR

KUNGL. BIBLIOTEKET

STOCKHOLM 1922  
KUNGL. BOKTRYCKERIET, P. A. NÖRSTEDT & SÖNER

220592

”Samla in,  
beskriva,  
bevara och  
tillhandahålla  
fysiskt och  
digitalt  
material.”



National Library  
of Sweden



figs in till tigh,  
fin högsta färdig  
p. hållna mig.

Respice finem

5678  
3205

4	2	8	5
6	3	2	4
9	8	7	6
	5	2	2
<hr/>			
2	10	0	7

Carolus.



# Atlantis

## Det Försvunna Riket

CD-ROM  
Windows - Mac  
PC & Mac

"Det är helt enkelt 1997 års bästa spel - alla kategorier."  
★★★★★  
Nora Västerbotten


"Atlantis visar vad vi kan förvänta oss av framtiden." 5/5  
Svenska Speleket

"Slår allt vi hittills sett. Rekommenderas"  
Datavärlden

"Kan det bli bättre?"  
★★★★★  
PC Guider

"...erhört vackert och med en spelbarhet som imponerar"  
★★★★★  
CD-ROM & MultiMedia

"...bland det bästa jag sett"  
★★★★★  
PC Hamna



# pinball

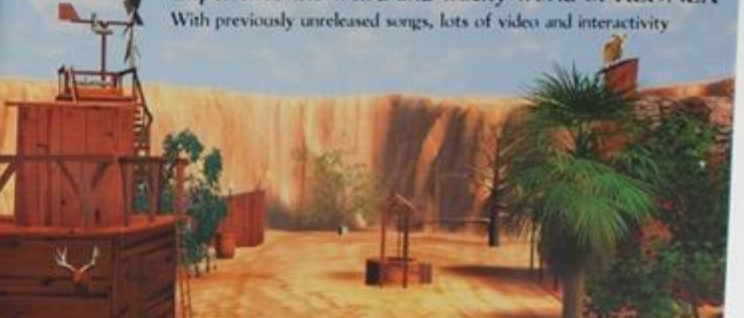
SIMULATOR



# IN BED WITH REDNEX

CD-ROM  
Windows - Mac  
PC & Mac

Experience the weird and wacky world of REDNEX  
With previously unreleased songs, lots of video and interactivity



PC CD-ROM

# MILJONÄR

CELEBRATION  
VE M VILL BLI  
VE M VILL BLI

celador

EIDOS




Wii U

3

www.pap.fr

Directores

# SUPER MARIO MAKER



SMART CLASSIC

# FRIDA I STALLET

FÖR HÄSTFÖRARE

FRIDA  
STALLET  
FÖR HÄSTFÖRARE

SMARTLINE

SPLET ÄR EN SVENSKA

MM00  
A/B 0429







0946-1536

Beställare  
*Joykildens AB*  
*Björngården 16, Stockholm K*

Korrektur skickas till dig  
1. om du inte har svarat på korrekturen på 1 vecka  
2. om du inte har svarat på korrekturen på 2 vecka

Arbete  
*Sättning och bearbetning*  
*100000 ord*

HS  
För beställare (1 till 10) till  
För tryckeriet (1 till 10) till  
För distribution (1 till 10) till

nr	namn	adress	post	stad	land
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					

Tryckare

NATIONAL FONOTEKST  
S 52 314

# DORIS

World some love today Baby

the latin kings

VELDAMMA

NATIONAL FONOTEKST  
S 52 314

SOUNDS FROM THE HEART OF GOTHENBURG

NENEH CHERRY

LIM 4 DISC HARDBOOK

Incl. 2x CD, 2x DVD, 2x CD, 2x DVD artwork, 2x DVD booklets

IN FLAMES PRODUCTION RECORDING IN SCANDINAVIUM, GOTHENBURG

MUSIC BY ANDERS FRIDEN GUITAR BY BÄRN GELTJE DRUMS BY DANIEL SVENSSON BASS BY PETER INYERS GUEST MUSIC BY NICLAS ENGELIN

R

NATIONAL FONOTEKST  
S 52 314

# Agnetha

FÄLTSKOG

Entombed

Wolverine Blues

NATIONAL FONOTEKST  
S 52 314

SUNDSVALLS KAMMARRÖR

Kjell Lönnå

live

NATIONAL FONOTEKST  
S 52 314

PHILIPS  
S 52 314-4  
PACS 316 • STEREO

# Tagels

NATIONAL FONOTEKST  
S 52 314-4

A 01822/23 L

PHILIPS

# NIARA

Epic of Space Flight in 2038 A. D.

Opera in Two Acts by **Karl-Birger Blomdahl**

Soloists of THE ROYAL OPERA, STOCKHOLM

Chorus and Orchestra of THE VIENNA VOLKSOPER

Conducted by **WERNER JANSEN**

Based on the poem "ANIARA" by HARRY MARTENSON Adapted for the stage by ERIK LINDEGREN

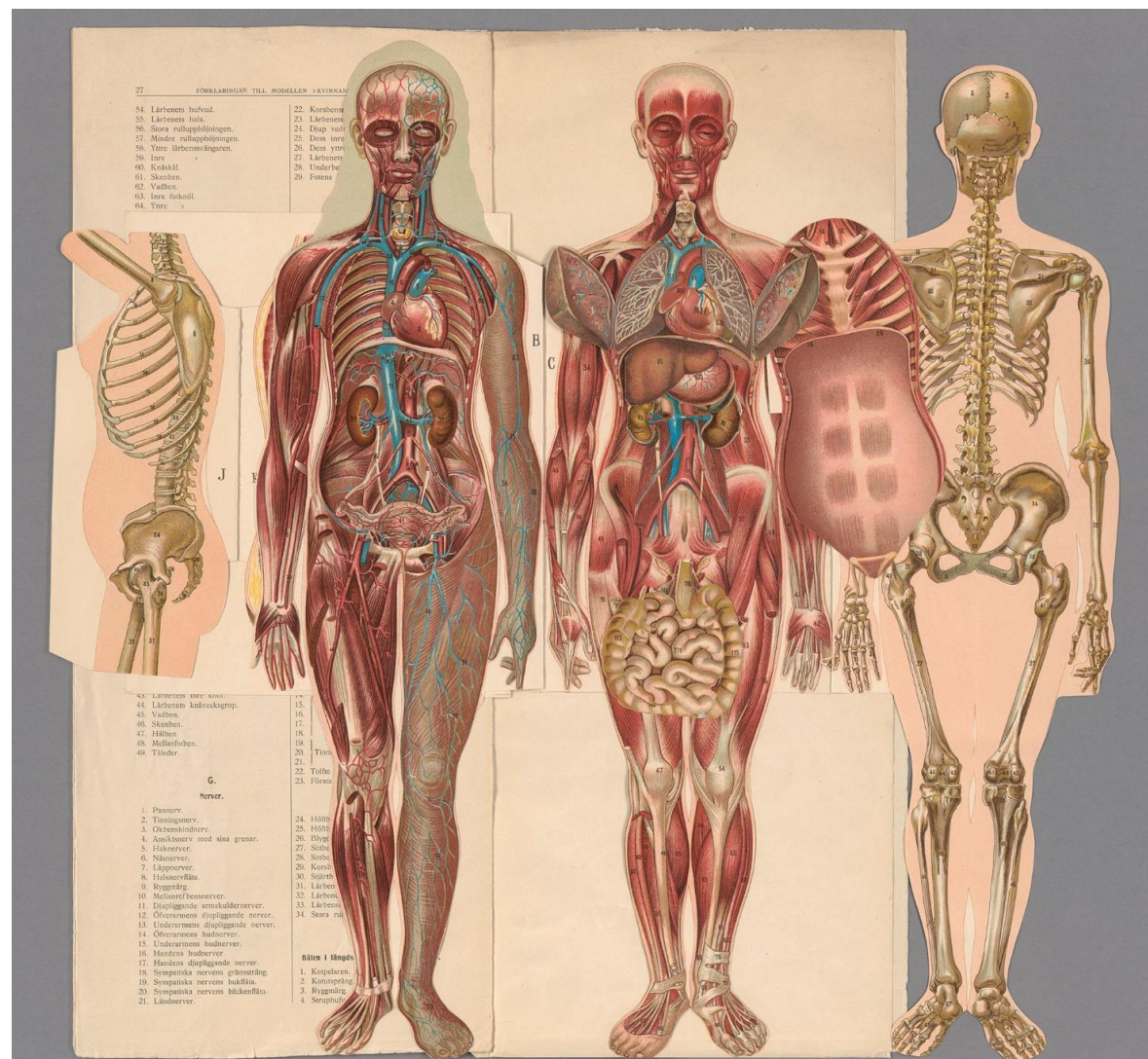
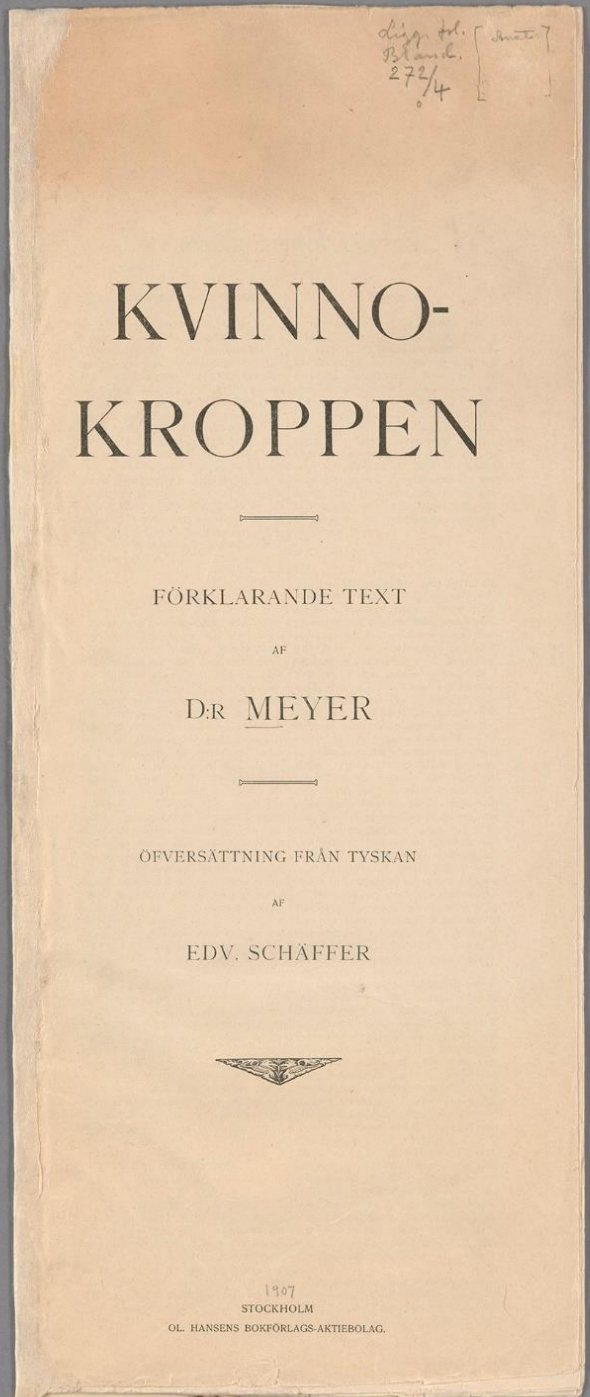
Special Electronic Effects by Swedish Radio

ROLAND CEDERMARK

Mirandas dröm



# Meyer, *Kvinnokroppen*: Stockholm, 1907



# Meyer, *Kvinnokroppen*: Stockholm, 1907

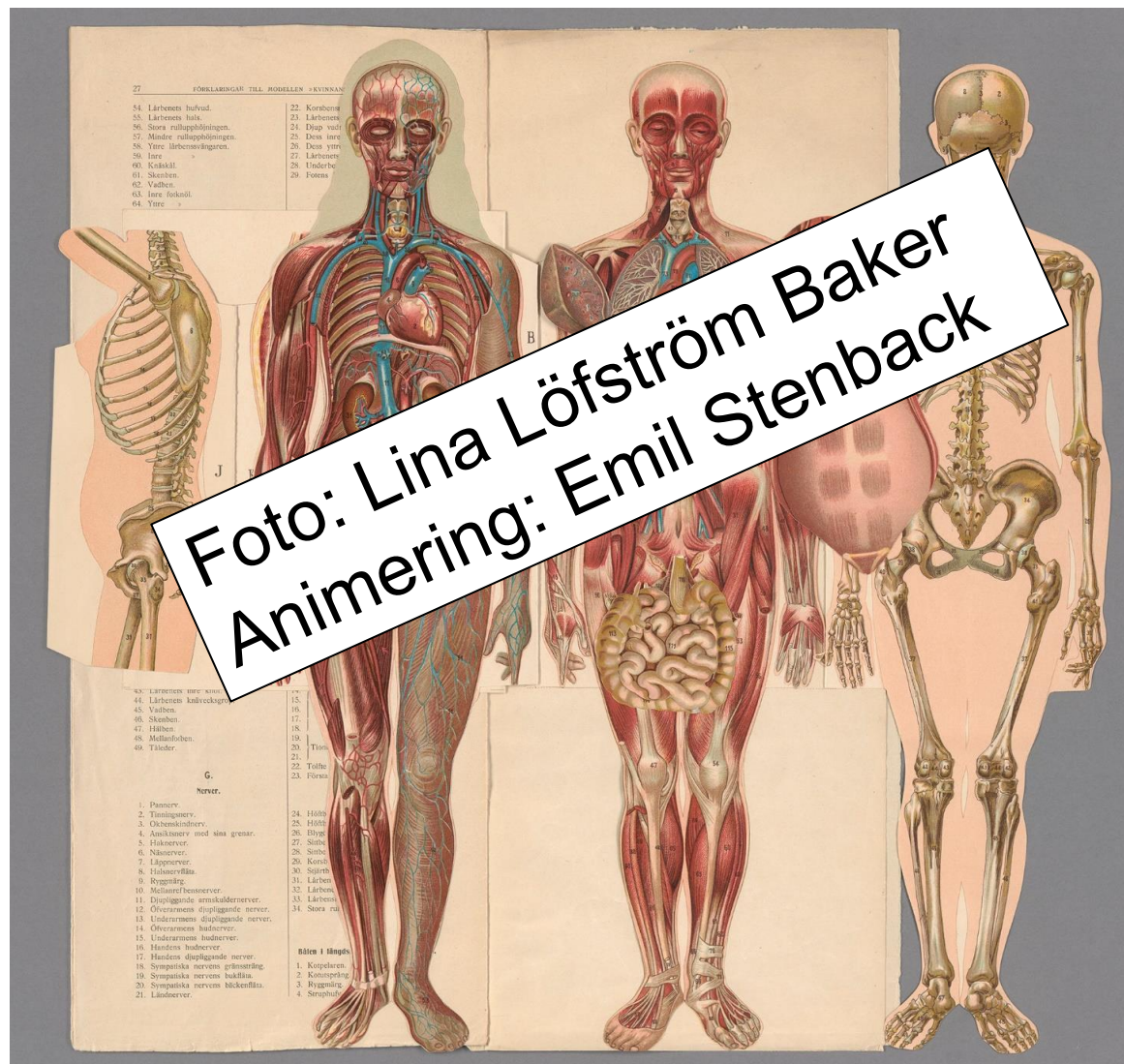
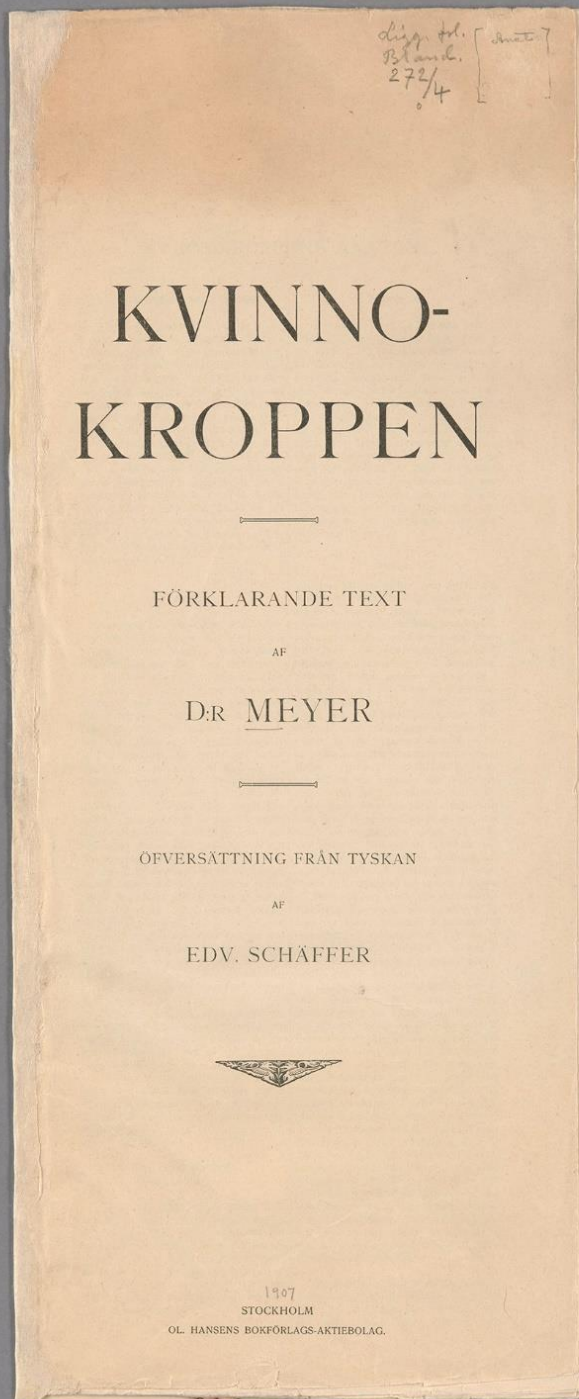
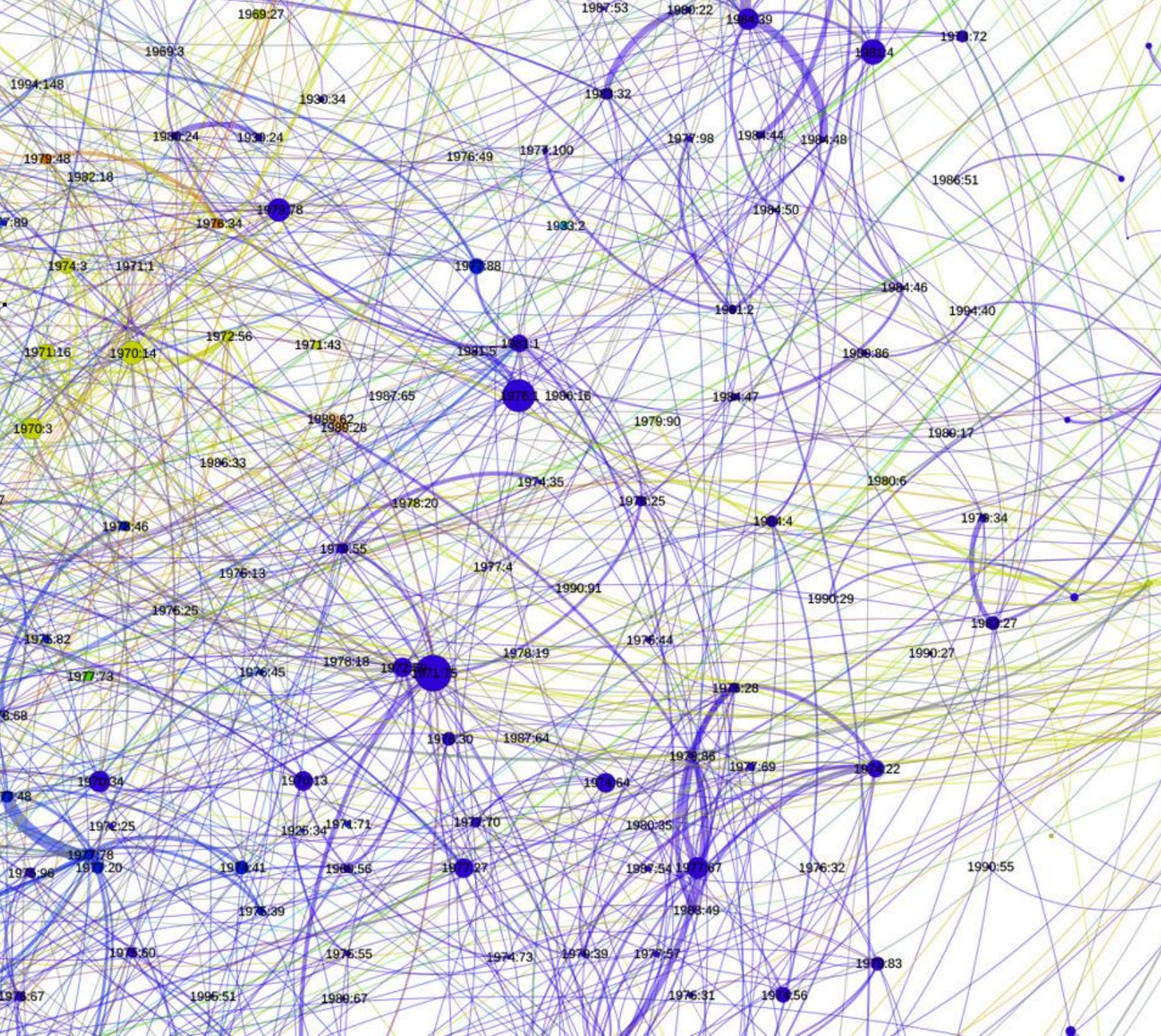


Foto: Lina Lofström Baker  
Animering: Emil Stenback





# Kulturarv som data

# KB-BERT ser dagens ljus på KB-labb 2020

## KB tillgängliggör kraftfulla modeller för språkförståelse

4 februari 2020

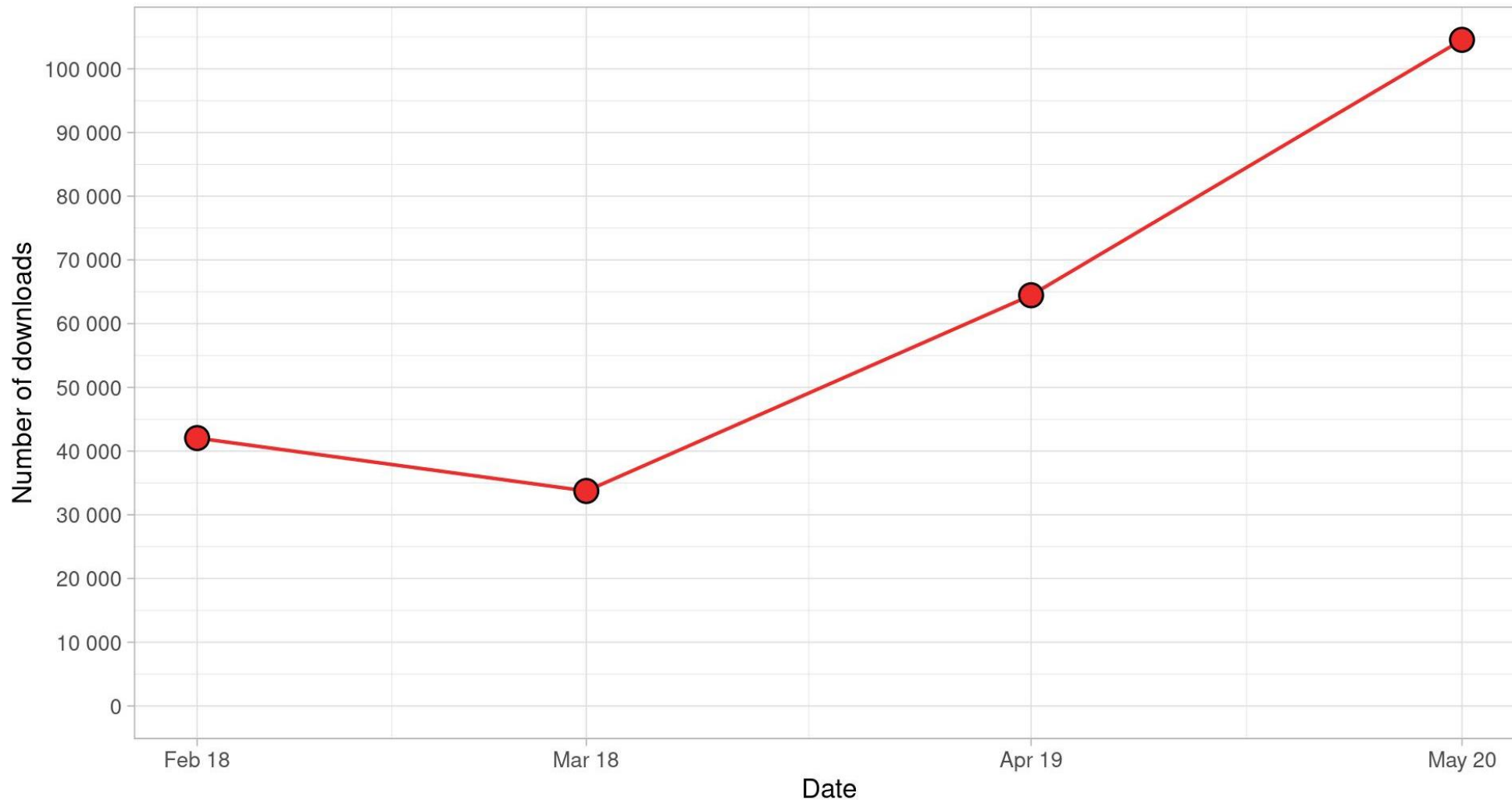
 [Forskning](#)

I dag publicerar KB tre svenska språkmodeller baserade på Googles "BERT" (Bidirectional Encoder Representations from Transformers). De första testerna visar att KB:s modeller överträffar Googles flerspråkiga modell.



# Antalet nedladdningar av KB-labs modeller

från beräkningsramverket Huggingface för rullande 30-dagars perioder, 2022





# Implant Terms: Focused Terminology Extraction with Swedish BERT Preliminary Results

Oskar Jerdhaf<sup>1</sup>, Marina Santini<sup>2</sup>, Peter Lundberg<sup>3,4</sup>, Anette Karlsson<sup>3,4</sup>, Arne Jönsson<sup>1</sup>

<sup>1</sup> Department of Computer and Information Science, Linköping University, Sweden  
oskje724@student.liu.se | arne.jonsson@liu.se

<sup>2</sup> RISE, Digital Health, Sweden

marina

Center for Medical Image Science and Visual

<sup>4</sup> Department of Medical Radiation Physics  
Linköping Uni

Peter.Lundberg@liu.se | Anet

Läkemedel ▾

Antibiotikaresistens

Covid-19

Vaccin

Om LmV ▾

*Gabriel Westman.*

## Lär AI att sortera biverkningsrapporter

Läkemedelsverket bygger kompetens om artificiell intelligens. Både för att kunna granska sådana produkter och för att lära sig använda AI som arbetsverktyg.

2 maj 2022, kl 09:00

0

TAGS

Artificiell intelligens

Läkemedelsverket



# Varför ska ett nationalbibliotek bygga språkmodeller?

- Utvärderingar visar att KB-BERT ger bättre träffsäkerhet än Googles modell.
- Tränad på brett språkligt material och olika genrer.
- Användning långt utöver forskningen, också många myndigheter.
- Demokrati – kontroll, transparens, öppenhet i modellbyggandet.





# VoxRex

- Kan översätta tal till text
- Akustiska modeller specifikt för svenska
- Bygger till stor del på digitiserad lokalradio
- Intern crowdsourcing för att göra modellerna mer träffsäkra
- Börjat transkribera befintligt AV-material



# Digitali- sering av lokalradion

“... vi kan visa upp det vi utvecklar så att invånarna också kan få vara med och tycka och tänka i den här **processengöran** larsson forsknings och utbildningschef på region jämtland **hergedalen** reporter viktor **hanssonså** till den stora dopningsskandalen där det visat sig att ryska staten låg bakom dopning av ryska idrottare enligt en rapport från **vada wold entidoping egiensiv** det här förvånar inte ...”

**P4 lokalradio - transkribering utan språkmodell**

“... vi kan visa upp det vi utvecklar så att invånarna också kan få vara med och tycka och tänka i den här processen. **Göran Larsson** forsknings och utbildningschef på **Region Jämtland Härjedalen**, reporter **Viktor Hansson**. Så till den stora dopningsskandalen där det visat sig att ryska staten låg bakom dopning av ryska idrottare enligt en rapport från **WADA World Wide Antidoping Agency**. Det här förvånar inte ...”

**P4 lokalradio - transkribering med språkmodell**

“... vi kan visa upp det vi utvecklar så att invånarna också kan få vara med och och tycka och tänka i den här processen. **Göran Larsson** forsknings och utbildningschef på **Region Jämtland Härjedalen**, reporter **Viktor Hansson**. Så till den stora dopningsskandalen där det visat sig att **ryska staten** låg bakom dopning av ryska idrottare enligt en rapport från **WADA World Wide Antidoping Agency**. Det här förvånar inte ...”

**P4 Blekinge 2016**

# Tal till text





# AI/språkmodeller som stöd i verksamhetsutveckling



# Pilotprojekt beskrivning av e-plikt genom AI

- 1500 artiklar från Expressen som kommit in som e-plikt "tvättas" från allt annat än det som är ren text..
- KB-labbs NER-modell (Named Entity Recognition) tar ut entiteter som personer, platser och tid.
- Betydelsebärande ord tas ut ur artiklarna för att hitta lättare i stora mängder metadata.
- Pliktlevererad e-pliktsmetadata används (utgivare, årtal m.m.)



## Projektgruppen

**Mattias Åkesson**  
**Martin Malmsten**

**Andreas Mace**  
**Sahar Abbasi**  
**Fredrik Klingwall**  
**Johan Furusvärd**  
**Kalle Wåhlin**

# Resekaoset i Europa det land för land

ARLANDA Publicerad 13 jun 2022 kl 14.05



Oden 2 PROD-v1.2.4

FILTRERA SÖK LEVERANTÖRER 1 KANALER 6

Visar 50 av 2 535 044 paket Sortering: Levererat (ny)

Leverantör	Kanal
✓ Expressen	Expressen,
✓ Expressen	Expressen,
✓ Expressen	Expressen,
✓ Expressen	Expressen,
✓ Expressen	Expressen,
✓ Expressen	Expressen,
✓ Expressen	Expressen,
✓ Expressen	Expressen,
✓ Expressen	Expressen,
✓ Expressen	Expressen, nätartiklar, löpande 2014-0...
✓ Expressen	Expressen, nätartiklar, löpande 2014-0...
✓ Expressen	Expressen, nätartiklar, löpande 2014-0...
✓ Expressen	Expressen, nätartiklar, löpande 2014-0...
✓ Expressen	Expressen, nätartiklar, löpande 2014-0...
✓ Expressen	Expressen, nätartiklar, löpande 2014-0...

<title>Resekaoset i Europa - så är det land för land </title>

<dcterms:accessRights>restricted</dcterms:accessRights>

<dcterms:format>text/html</dcterms:format>

<description>&lt;p&gt;Det väntar ingen större lättnad för den som klarat sig genom hel Arlanda.&lt;/p&gt;&lt;p&gt;Det råder resekaos över hela Europa.&lt;/p&gt;&lt;p&gt;Här några av de stora destinationerna.&lt;/p&gt;</description>

<content:encoded>&lt;p&gt;Kökaoset på Arlanda skapar stora rubriker och även med europ är det illa ställt. Det är ovanligt att vägen till flygplatsen måste stängas.&lt;/p&gt; köer som ringlar sig utanför flygplatserna har blivit en vanlig syn i land efter land i grund och botten desamma som på Arlanda. Pandemin ledde till att över 190 000 flygpl förlorade sina jobb.&nbsp;&lt;/p&gt;&lt;p&gt;Sommaren 2022 blir för många miljoner flygresan utomlands. Det uppdämda behovet är enormt.&lt;/p&gt;&lt;p&gt;Men flygplatser flygbolagen har inte hunnit anställa tillräckligt många. Lägg därtill lokala problem s strejker.&lt;/p&gt;&lt;p&gt;Så den som ska ut och resa i sommar bör packa med sig en s också.&nbsp;&lt;/p&gt;&lt;p&gt;Så här ser läget ut i några stora destinationer:&lt;/p&gt;&lt;p&gt;&lt;strong&gt;Belgien:&lt;/strong&gt;&lt;/p&gt;&lt;p&gt;väntas i slutet av juni och början av juli, men VD:n för Bryssels flygplats tror att d ska kunna undvikas. Målet är 1 200 nyanställningar i år och 600 av tjänsterna har fyll är bagagehanteringen. En vild strejk bland bagagehanterarna ledde till kaos i början a juni.&lt;/p&gt;&lt;p&gt;&lt;strong&gt;Storbritannien:&lt;/strong&gt;&lt;/p&gt;&lt;p&gt; som ringlar sig utanför terminalerna har drabbat många av de större flygplatserna, frå Manchester. Det har varit en följetong i brittiska medier. Personalbrist både hos flyg säkerhetspersonalen. Heathrows VD tror att det kommer att ta 12 till 18 månader innan som normalt.&lt;/p&gt;&lt;p&gt;&lt;strong&gt;Tyskland:&lt;/strong&gt;&lt;/p&gt;&lt;p&gt; flygplats sade upp 4 000 anställda under pandemin och behöver fylla 1 000 luckor i år. planerar hyra in upp till 2 000 säkerhetspersonal från Turkiet för att minska på köern minska på trycket är också att ställa in flygningar och det är precis vad tyska Luftha

Material utklart, löpande 2014-0... Meddelning meddelning kan vara d ordning...

Bearbetning påbörjades: 2022-06-13 14:15:14

Filer: 9

Filtyper i paket: text/html, text/xml, image/jpeg

Storlek: 392.86 kB (392 863 bytes)

Versioner: 1

# So far...

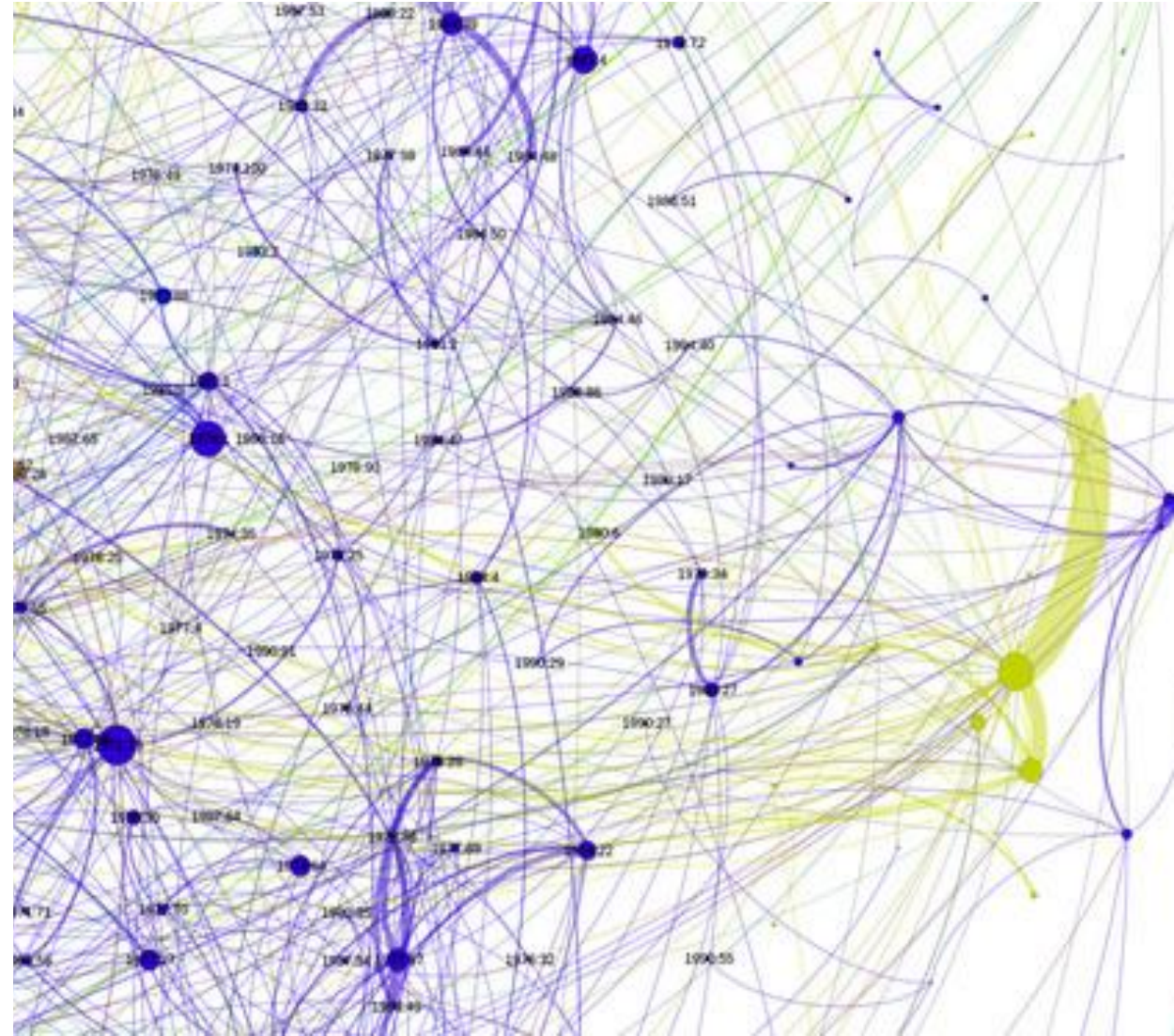
- Förståelse på KB för att manuell katalogisering inte kommer kunna tillämpas på stora mängder metadata
- Viss rörelse från att beskriva OBJEKT till att kunna använda sig av INNEHÅLLET i det pliktlevererade
- AI inte ska användas på denna typ av metadata för att efterlikna manuell katalogisering, utan göra det som människor inte kan på stora mängder metadata



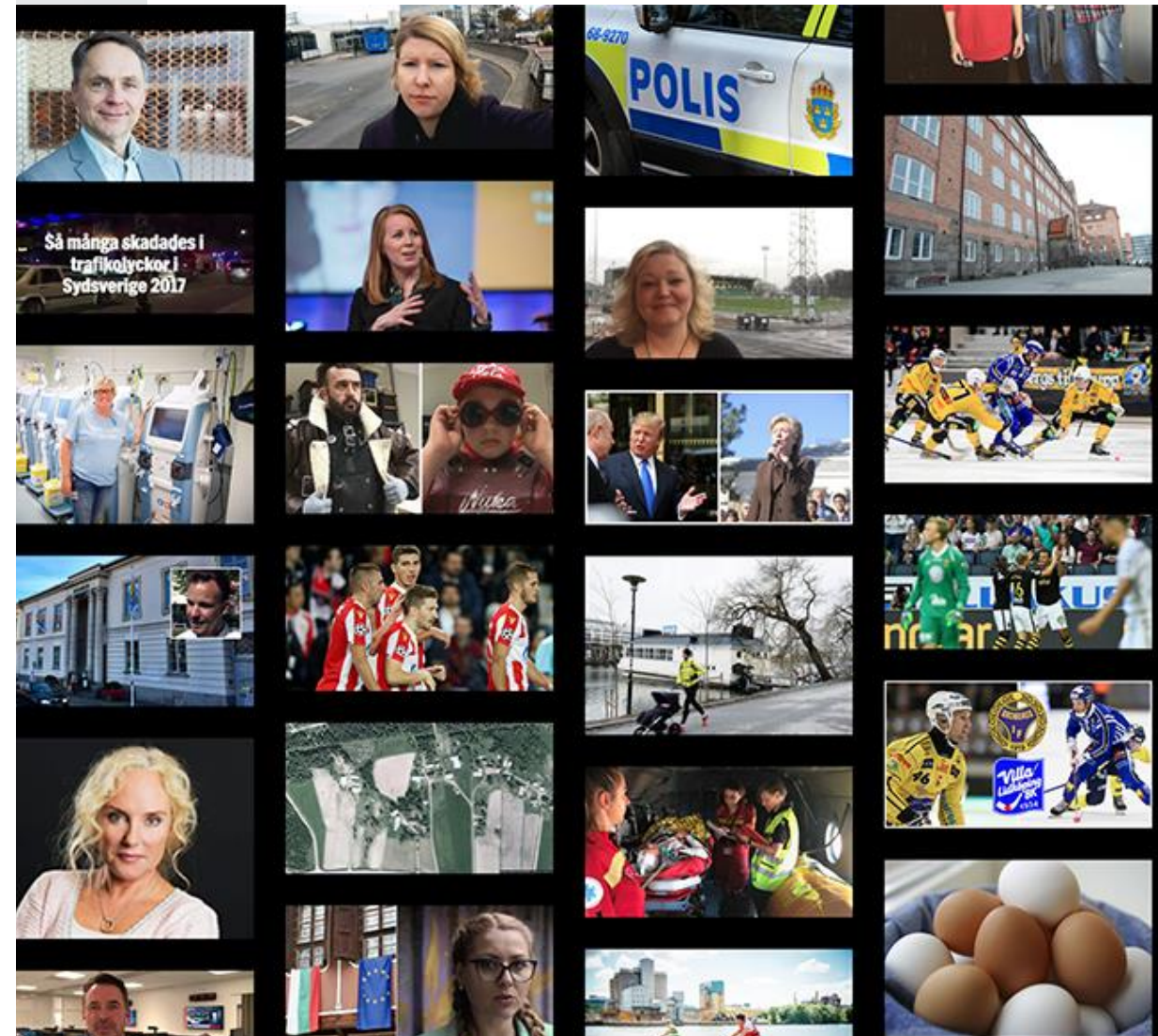
# Digitization vs Digital Transformation



# Researchers vs Research



# Collections vs Content/Data



# Experts vs Cross-competence/Networking





# Möjligheter och utmaningar

- Förväntningar från forskningen – hitta allt material och använda materialet som data.
- Utmaningar i tillhandahållandet: upphovsrätt, dataskydd. Kompetensbehov hos forskare att kurera materialet och utveckla metoder.
- Internt: Från katalogiseringsregler till användarbehov och möjligheter. Från att katalogisera till att kvalitetskontrollera och träna modeller.
- Kompetensfrågan: "Bryggor", data scientists i samverkan med bibliotekarier, systembibliotekarier.
- Vi kommer att ta emot allt mer digitalt material – måste tänka och göra nytt.
- Vi kommer inte ha full kontroll...



**Fråga inte  
enbart vad AI  
kan göra för  
biblioteket –  
fråga också vad  
biblioteket kan  
göra för AI.**



# TACK!

- Helene Wallskär "[Lär AI att sortera biverkningsrapporter](#)" *LäkemedelsVärlden*, 220502.
- Haffenen, Fanto, Malmsten, Börjeson "[Making and Using AI in the Library: Creating a BERT Model at the National Library of Sweden](#)" *College & Research Libraries pre-print* (January 2023)
- Malmsten, Haffenden, Börjeson "[Hearing voices at the National Library -- a speech corpus and acoustic model for the Swedish language](#)" ArXiv
- Jerdhaf, Santini, Lundberg, Anette Karlsson, Arne Jonsson "[Implant Terms: Focused Terminology Extraction with Swedish BERT Preliminary Results](#)"
- [Om KB-labb med länkar till nedladdningsställen mm.](#)



[Karin.gronvall@kb.se](mailto:Karin.gronvall@kb.se)  
[@riksbibliotekarien](#) (Insta)