

VINGE

# Maskininlärning och upphovsrätt

## En introduktion

Tobias Kempas

2023-06-12

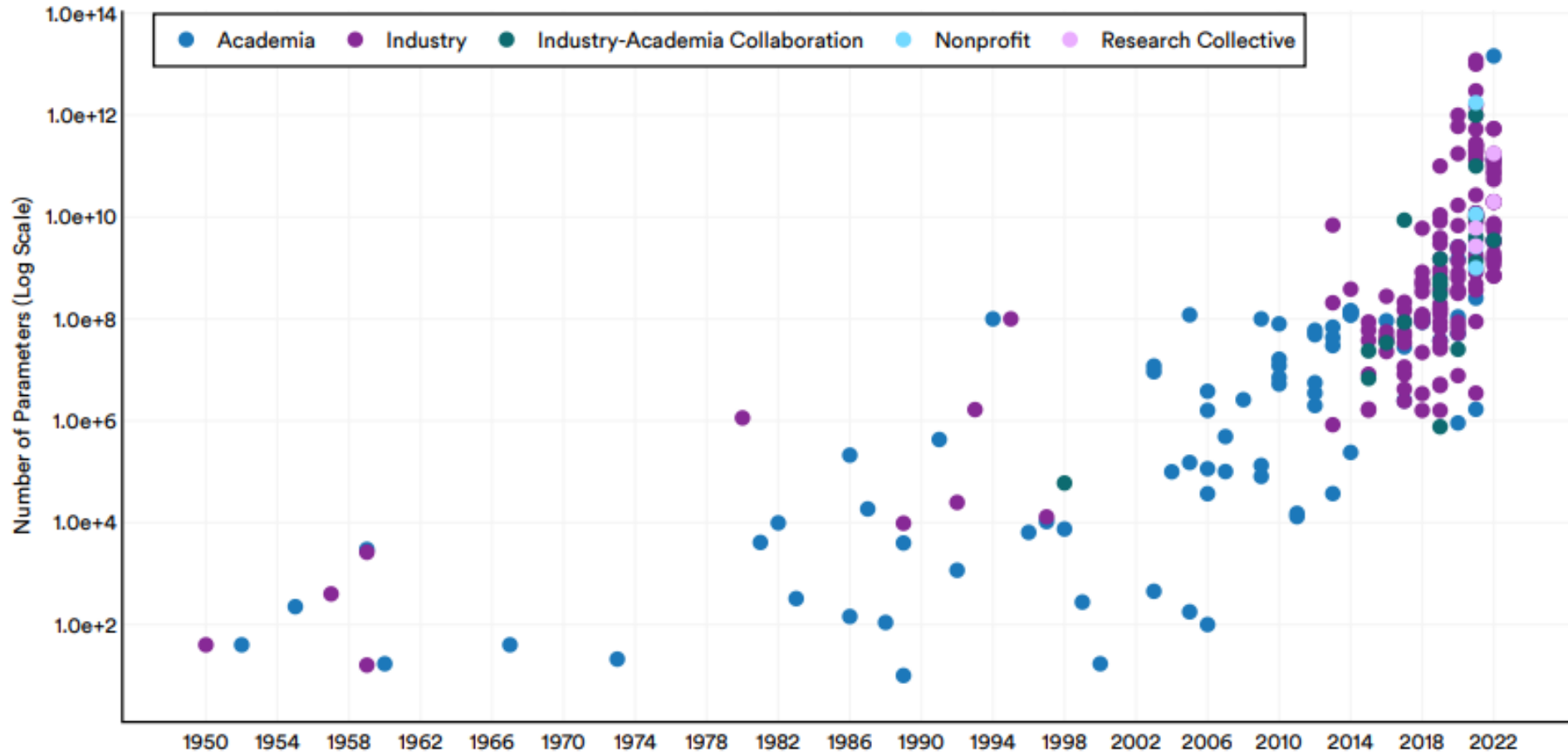


**Kort bakgrund**

# Exponentiell utveckling

### Number of Parameters of Significant Machine Learning Systems by Sector, 1950–2022

Source: Epoch, 2022 | Chart: 2023 AI Index Report



# Samspelet mellan teknik och samhälle



*”Technology is not an exogenous force over which we have no control. We are not constrained by a binary choice between “accept and live with it” and “reject and live without it””*

**Professor Klaus Schwab, 2016**

# Användningsområden

## *Exempel*

Skapande och innovativ verksamhet

Sjukdomsdiagnos

Översättningar

Nätsäkerhet

Hushållsapparater

Bank & finans

Kundservice

Marknadsföring

Utbildning och information

Trafiksäkerhet

Militärt försvar

Jordbruk

Energiproduktion

Avfallshantering

Robotar i industrin

Beslutsstöd vid myndighetsutövning



# Politiska målsättningar

*Exempel*



EU ska bli ”den **världsledande** regionen för utveckling och användning av etisk och säker AI av spetskvalitet” och ”**världsledande** inom högpresterande datorsystem som tillhandahåller datalösningar och ett förbättrat samarbete kring avancerad vetenskaplig forskning”

Meddelande från kommissionen till Europaparlamentet, rådet, Europeiska ekonomiska och sociala kommittén samt Regionkommittén. Samordnad plan om artificiell intelligens, COM(2018) 795 final av den 7 december 2018, Bilaga 1 s. 1.



Sverige ska vara ”**bäst i världen** på att använda digitaliseringens möjligheter” och ”**ledande** i att ta tillvara möjligheterna som användning av AI kan ge, med syftet att stärka både den svenska välfärden och den svenska konkurrenskraften”

Regeringskansliet, Nationell inriktning för artificiell intelligens, 16 maj 2018 s. 4 f.

# Immateralrättens utformning

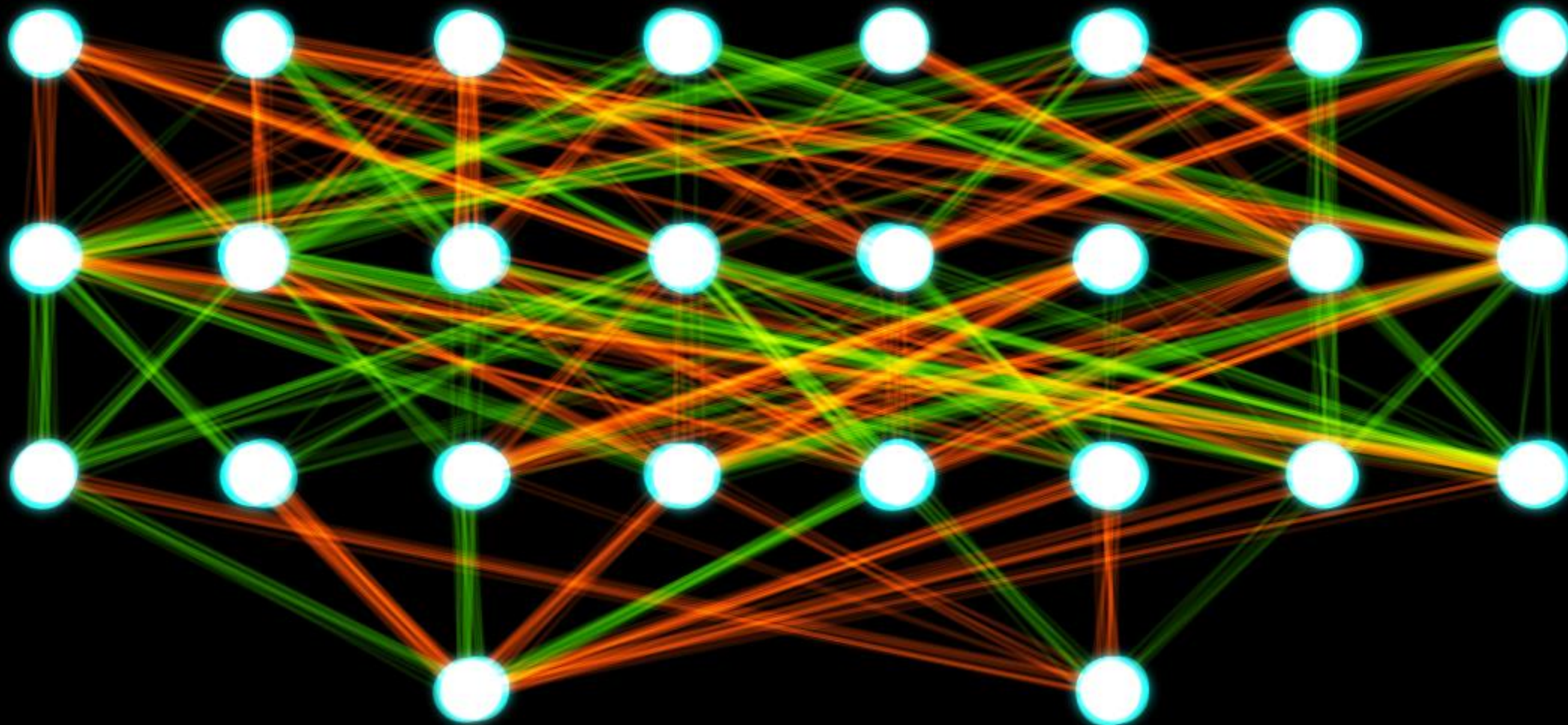


Lagen bör ”tillhandahålla ett balanserat och innovationsdrivet immaterialrättsligt skydd till förmån för de europeiska utvecklarna av AI-teknik, för att stärka de europeiska företagens konkurrenskraft”

**Europaparlamentets resolution om immateriella rättigheter för utveckling av artificiell intelligens (2020/2015(INI)) av den 20 oktober 2020**

Ensamrättigheter är önskvärda, ”men samtidigt måste EU ha strategier som gör att skyddet för de immateriella rättigheterna går att förena med behovet av friktionsfri utbyggnad av ny och innovativ teknik”

**Meddelande från kommissionen till Europaparlamentet, rådet och Europeiska ekonomiska och sociala kommittén. Ett balanserat system för säkerställande av ett skydd för immateriella rättigheter som svarar mot dagens samhällsutmaningar, COM(2017) 707 final av den 29 november 2017**



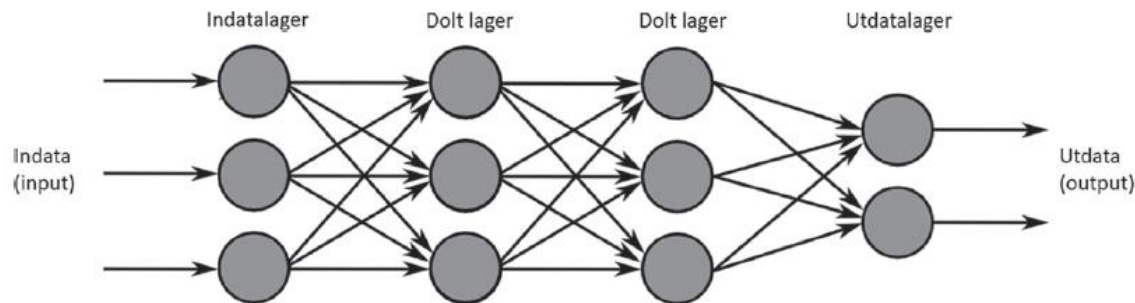
**Vad är maskininlärning?**



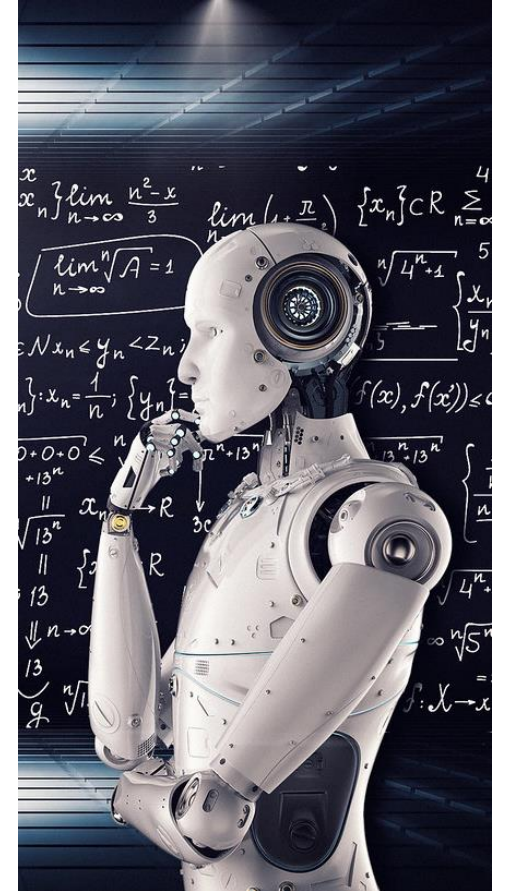
# En samling metoder för utveckling av datorprogram (AI)

## *Syfte, särdrag och centrala begrepp*

- Datorprogrammet ska kunna ändras när det konfronteras med ny information, utan att systemutvecklaren uttryckligen behöver beskriva *hur* det ska förändras
- Maskininlärning syftar till att systemet själv ska "förstå" samband och korrelationer (d.v.s. självständigt "lära sig" regler för att lösa en eller flera specifika uppgifter)
- Huvudtyper: Övervakad, oövervakad och förstärkt
- Maskininlärning associeras ofta med artificiella neurala nätverk ("ANN"). Huvuduppgiften för ett ANN är att känna igen mönster



- Ytterligare begrepp: Djupinlärning (många lager, hierarkisk inlärning av särdrag)

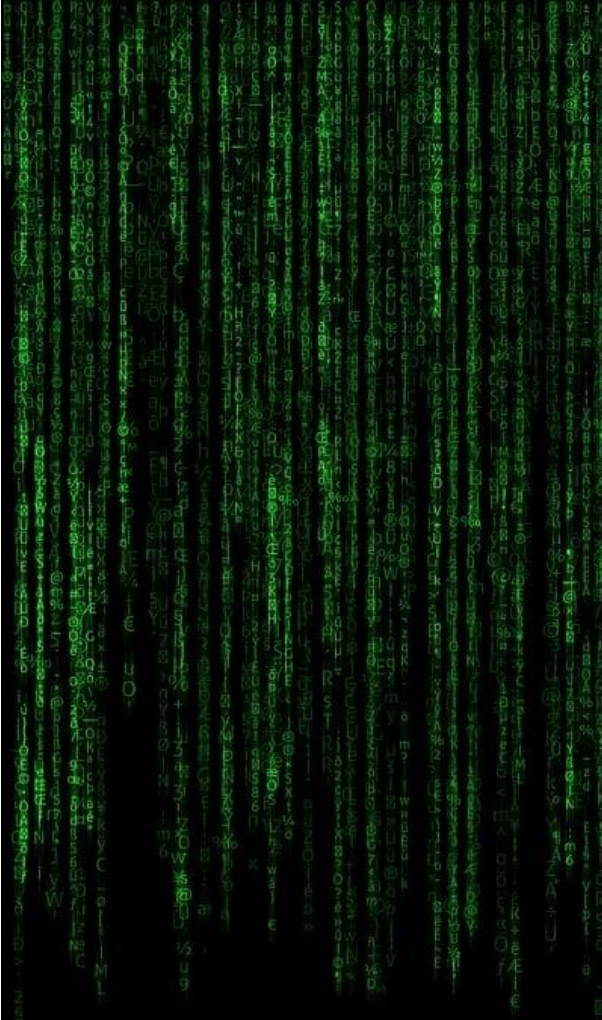


# Träningsdata

## *Behovet av data*

- Vid maskininlärning används digital information för att ”lära” algoritmerna att hitta samband och mönster, att dra slutsatser och/eller att göra förutsägelser
- Stora datavolymer krävs (särskilt vid djupinlärning)
- Ju större datavolymer algoritmerna har tillgång till, desto mer kan AI-systemet lära sig
- Felaktiga och/eller bristfälliga datavolymer innebär en risk att AI-systemen drar felaktiga slutsatser och/eller fattar irrationella eller oetiska beslut
- *”[W]ithout data, there is no AI”*

EU-kommissionen, Vitbok. Om artificiell intelligens – en EU-strategi för spetskompetens och förtroende, COM(2020) 65 final av den 19 februari 2020 s. 19



# Träningsdata

*Volym och typ(er) avgörs av träningens syfte*



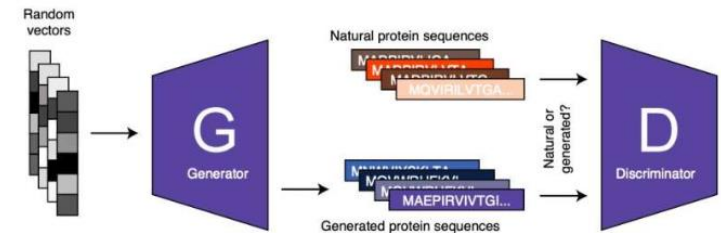
- 346 st. Rembrandt-målningar



- Skrapning av internet (Big Data)
- Artiklar, böcker, webbsidor  
programspråk m.m.

## ProteinGAN: A generative adversarial network that generates functional protein sequences

by Ingrid Fadelli, Phys.org



- Information om (a) välkända proteiner och (b) artificiella proteinstrukturer

# Träningsdata

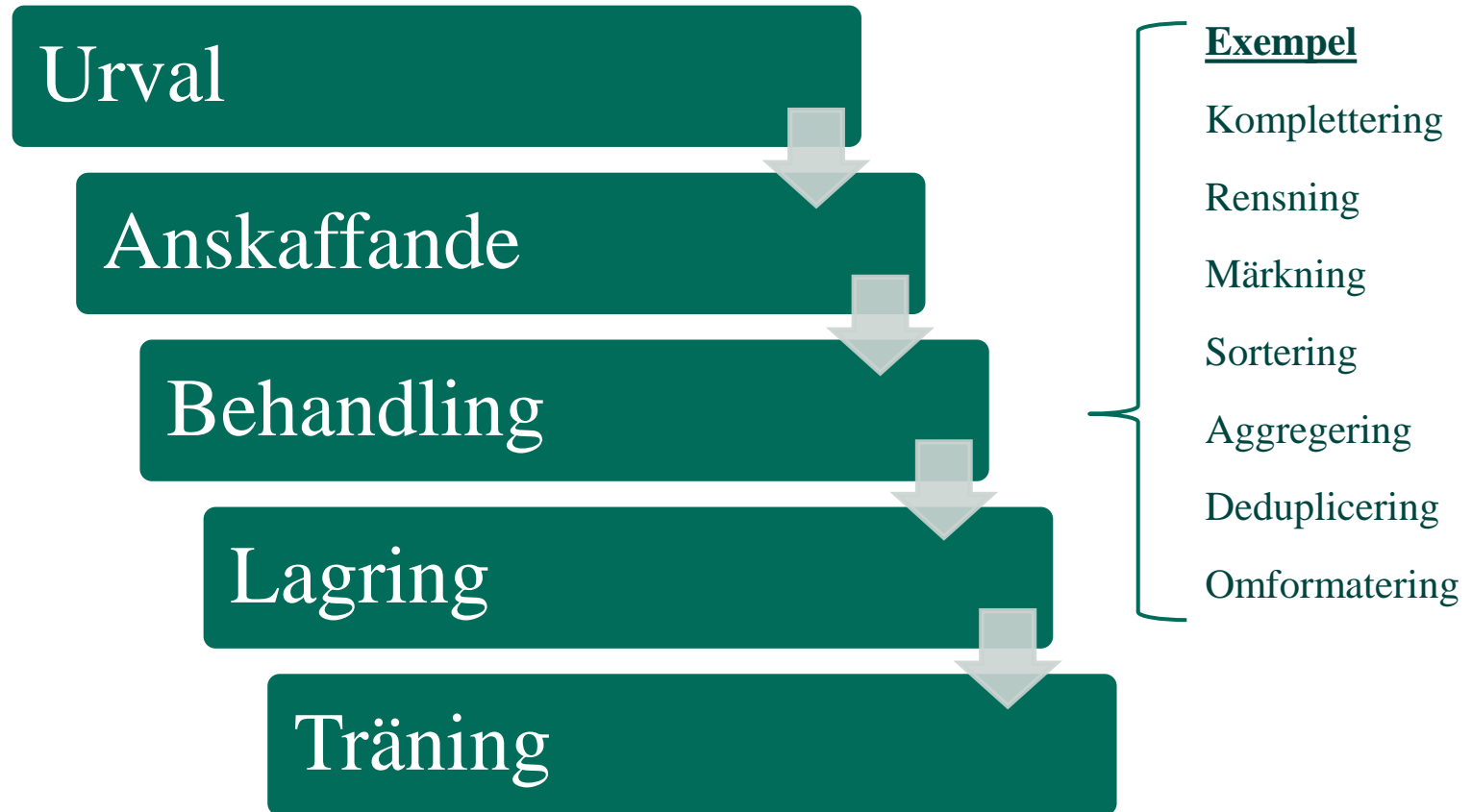
*Exemplarframställning (URL)*

- ”... i ursprungligt eller ändrat skick, i översättning eller bearbetning, i annan litteratur- eller konstform eller i annan teknik”

**2 § 1 st. URL**

- ”Framställning av exemplar innefattar varje direkt eller indirekt samt tillfällig eller permanent framställning av exemplar av verket, oavsett i vilken form eller med vilken metod den sker och oavsett om den sker helt eller delvis”

**2 § 2 st. URL**





**Inskränkningar (URL)**

# Framställning av tillfälliga exemplar

*Se t.ex. C-5/08 (Infopaq I), C-302/10 (Infopaq II) och C-360/13 (Public Relations Consultants Association)*

*”Tillfälliga former av exemplar av verk får framställas, om framställningen utgör en integrerad och väsentlig del i en teknisk process och om exemplaren är flyktiga eller har underordnad betydelse i processen. Exemplaren får inte ha självständig ekonomisk betydelse.”*

*Framställning av exemplar enligt första stycket är tillåten bara om det enda syftet med framställningen är att möjliggöra*

- 1. överföring i ett nät mellan tredje parter via en mellanhand, eller*
- 2. laglig användning, dvs. användning som sker med tillstånd från upphovsmannen eller dennes rättsinnehavare, eller annan användning som inte är otillåten enligt denna lag.*

*Första och andra styckena ger inte rätt att framställa exemplar av litterära verk i form av datorprogram eller sammanställningar.”*



# Framställning av tillfälliga exemplar

## *Några iakttagelser*



- ”Integrerad och väsentlig del i en teknisk process”
  - (a) Kopiering som är (a) direkt nödvändig *eller* (b) behövlig för att för att processen ska fungera korrekt och effektivt
- Alt 1: ”Flyktiga” exemplar
  - Kopiornas livslängder ska vara begränsade till vad som behövs för att processen ska fungera väl. Kopiorna ska automatiskt raderas, utan mänskligt ingripande, så snart de inte längre fyller sin funktion att möjliggöra genomförandet av processen
- Alt 2: ”Underordnad betydelse” i processen
  - Exemplaren får inte ha någon självständig existens eller något självständigt syfte i förhållande till den tekniska process mångfaldigandet ingår i

# Framställning av tillfälliga exemplar

## *Några iakttagelser*

- Ingen ”självständig ekonomisk betydelse”
  - Den ekonomiska nyttan av att genomföra mångfaldigandet får inte vara avgränsad eller avskiljbar från den ekonomiska nyttan av den lagliga användningen
  - Mångfaldigandet får inte medföra någon ytterligare ekonomisk nytta, utöver den nytta som följer av den lagliga användningen av verket (de tillfälliga kopiorna får inte i sig utnyttjas ekonomiskt)
  - Test: Om nyttan av kopieringen förverkligas först vid den lagliga användningen är kravet uppfyllt
- Laglig användning
  - Automatiserad analys av data/information (jfr med situationen att användaren betraktar, läser eller lyssnar till verk – eller skapar sammanfattningar av verk)
  - *”ML statistical analysis is arguably as lawful as the preparation of summaries and is not a right reserved to the right holder by EU copyright law”* (Margoni, 2018)
  - Syftet får inte vara att återge eller återskapa de verk som har tillfälligt kopierats i träningsprocessen (om det saknas tillstånd eller annat undantag/inskränkning).





# Framställning av exemplar inom arkiv och bibliotek

*Begränsad krets. Forskningsändamål*



*”De statliga och kommunala arkivmyndigheterna, de vetenskapliga bibliotek och fackbibliotek som drivs av det allmänna och folkbiblioteken har rätt att framställa exemplar av verk, dock inte datorprogram, även*

- 1. för kompletterings- eller forskningsändamål,*
- 2. för att tillgodose lånesökandes önskemål om enskilda artiklar eller korta avsnitt eller om material som av säkerhetsskäl inte bör lämnas ut i original, eller*
- 3. för användning i läsapparater.*

**16 § 2 st. URL  
(jfr artiklarna 5.3 a) och 5.3 n) infosoc-direktivet)**

# Framställning av exemplar inom arkiv och bibliotek

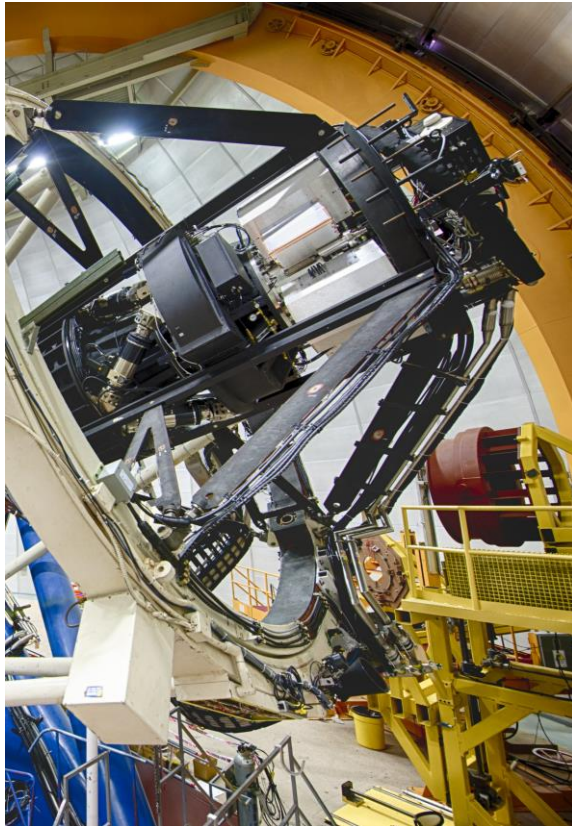
## *Begränsad krets*

- Forskningen (med tillhörande exemplarframställning) måste ske för institutionens räkning eller inom ramen för dess verksamhet
- Undantaget omfattar:  
  
Anställda och uppdragstagare som handlar för de angivna institutionernas räkning eller inom ramarna för institutionernas verksamhet
- Undantaget omfattar *inte*:  
  
Personer som helt eller delvis agerar för egen räkning och som inte har någon stark anknytning till en berättigad institution
- Detta utesluter inte att forskningen skapar nytta för utomstående



# Framställning av exemplar inom arkiv och bibliotek

## *“Forskningsändamål”*



- Jfr artikel 5.3 n) infosoc-direktivet: *”Forskningsssyfte”*
- Jfr artikel 5.3 a) infosoc-direktivet: *”Vetenskaplig forskning”*
- Ordet *”vetenskaplig”* används *inte* i 16 § 2 st. URL
- Enligt 16 § 2 st. URL ska det åtminstone vara fråga om någorlunda kvalificerad forskning och inte om kopiering för vanliga elevuppsatser eller liknande
- Ståndpunkt: Berättigade svenska institutioner får använda stora datavolymer (d.v.s. framställa exemplar av skyddade verk) vid maskininläring, *om* ändamålet är att bedriva någorlunda kvalificerad forskning

- Lästips:

Triaille, m.fl., *Study on the legal framework of text and data mining (TDM)*, 2014

Rosati, *Copyright as an Obstacle or Enabler? A European Perspective on Text and Data Mining and its Role in the Development of AI Creativity*, 2019

# Text- och datautvinning

*En analytisk process som syftar till generering av information*

## Text- och datautvinning

Säkerställande av datatillgång

Inhämtande av originaldata

Vidare behandling och analys (t.ex. med hjälp av maskininlärning)

- *”en automatiserad teknik som används för att analysera text och data i digital form i syfte att generera information”*

**15 c § 1 st. URL**

- *”automatiserad analysteknik som används för att analysera text och data i digital form för att generera information, inklusive, men inte begränsat till, mönster, trender och samband”*

**Artikel 2 DSM-direktivet**

- Definitionen tar sikte främst på det generella och analytiska syftet med utvinningen och inte på de tekniker som används
- *”[T]he process of discovering meaningful correlations, patterns and trends by sifting through large amounts of data stored in repositories. Data mining employs pattern recognition technologies, as well as statistical and mathematical techniques”*

**<https://www.gartner.com/en/information-technology/glossary/data-mining>**

# Text- och datautvinning

*Begränsad krets. Forskningsändamål*

*Forskningsorganisationer, sådana bibliotek och museer som är tillgängliga för allmänheten, arkiv samt institutioner för film- eller ljudarvet får framställa exemplar av verk som de har lovlig tillgång till, dock inte datorprogram, för att utföra text- och datautvinning för forskningsändamål. Exemplaren får inte behållas längre än vad som är nödvändigt för ändamålet och får inte användas för andra ändamål. Exemplaren ska lagras på ett sätt som hindrar obehörig användning.*

*Första stycket hindrar inte att upphovsmannen vidtar proportionerliga åtgärder för att säkerställa integritet och säkerhet i nätverk och databaser som innehåller verk.*

*Avtalsvillkor som inskränker rätten att använda verk enligt denna paragraf är ogiltiga.*

**15 b § URL**  
**(jfr artikel 3 DSM-direktivet)**



# Text- och datautvinning

## *Begränsad krets*

- Undantaget omfattar:

1. Forskningsorganisationer (15 c § 2 st. URL, artikel 2 DSM-direktivet)

- *”ett universitet med tillhörande bibliotek, en forskningsinstitution eller en annan enhet vars främsta mål är att bedriva vetenskaplig forskning eller utbildningsverksamhet som omfattar vetenskaplig forskning”*
- *”verksamheten i organisationen [måste] bedrivs på ett sådant sätt att ett företag som utövar ett avgörande inflytande på organisationen inte kan ha prioriterad tillgång till resultaten av forskningen”*
- utan vinstsyfte / återinvestering i forskningen, alternativt i enlighet med ett uppdrag i det allmännas intresse som har erkänts av en MS

2. Allmänt tillgängliga bibliotek och museer

3. Arkiv

4. Institutioner för film- eller ljudarvet

} *”Kulturarvsinstitutioner”*  
(artikel 2 DSM-direktivet)



# Text- och datautvinning

## *“Forskningsändamål”*



- Artikel 3 DSM-direktivet (eng): *”scientific research”*
- Prop. 2021/22:278 s. 45: Det finns inte någon möjlighet att bortse från kravet på att forskningen ska vara vetenskaplig
- Ingresspunkt (12) DSM-direktivet: *”Termen vetenskaplig forskning bör ... anses omfatta såväl naturvetenskaplig som humanistisk forskning.”*
- Ds 2021:30 s. 62: *”Uppräkningen är exemplifierande och även andra ämnesområden är alltså avsedda att omfattas.”*
- Exkluderade aktiviteter (exempel):
  - Allmänt sökande efter information
  - Prediktiv marknadsanalys (även kallad ”kommersiell forskning”)
  - Verksamhet som inte rimligen kan resultera i mer än vad som redan följer av beprövad erfarenhet
  - Tillämpning av tidigare forskningsresultat

# Text- och datautvinning

*Övriga aktörer (näringslivet). Opt-out-möjlighet för rättighetshavarna*

*”Den som har lovlig tillgång till ett verk får framställa exemplar av verket för text- och datautvinningsändamål. Exemplaren får inte behållas längre än vad som är nödvändigt för ändamålet och får inte användas för andra ändamål.*

*Första stycket gäller inte om upphovsmannen på lämpligt sätt har förbehållit sig den rätt som avses där.”*

**15 a § URL  
(jfr artikel 4 DSM-direktivet)**





# Text- och datautvinning

*Hur kan en rättighetshavare förbehålla sig användningen “på lämpligt sätt”?*

- Måste förbehållet alltid ske med effektiva *tekniska* medel?
- Kontrollmöjligheten är viktig. Jfr mål C-392/19 (VG Bild-Kunst), punkt 46: ”*[F]ör att säkerställa rättssäkerheten och ett väl fungerande internet [kan det inte] vara tillåtet för upphovsrättsinnehavaren att begränsa sitt samtycke på annat sätt än genom effektiva tekniska åtgärder*”
- Exempel:
  - Robots.txt-filer  
(påverkar besök av sidor/innehåll)
  - Metarobot-tags  
(påverkar indexering av sidor/innehåll)
- OBS: Instruktionerna kan ignoreras i praktiken. Bruket av robots.txt-filer och metarobot-tags bygger på ”best practice”

```
User-agent: *  
Allow: /
```

Alla sökrobotar tillåts  
att utforska allt innehåll

```
User-agent: *  
Disallow: /
```

Alla sökrobotar förbjuds  
att utforska allt innehåll

```
User-agent: Googlebot  
Disallow: /blog  
Allow: /blog/example-post
```

Googlebot förbjuds att utforska  
samtliga bloggar, med ett undantag



**Är regelverket tillräckligt balanserat?**

# Är regelverket tillräckligt balanserat?

*Jfr pågående tvister i USA*

Anderson, et al v. Stability AI Ltd., et al,  
(3:23-cv-00201)

Case 3:23-cv-00201-WHO Document 50 Filed 04/18/23 Page 1 of 29

1 LATHAM & WATKINS LLP  
2 Andrew M. Gass (SBN 259694)  
3 *andrew.gass@lw.com*  
4 Michael H. Rubin (SBN 214636)  
5 *michael.rubin@lw.com*  
6 505 Montgomery Street, Suite 2000  
7 San Francisco, California 94111-6538  
8 Telephone: 415.391.0600  
9  
10 *Attorneys for Defendant DeviantArt, Inc.*

11

12 UNITED STATES DISTRICT COURT  
13 NORTHERN DISTRICT OF CALIFORNIA  
14 SAN FRANCISCO DIVISION

15 SARAH ANDERSEN, an individual;  
16 KELLY MCKERNAN, an individual;  
17 KARLA ORTIZ, an individual,  
18 Individual and Representative Plaintiffs,  
19 v.  
20 STABILITY AI LTD., a UK corporation;  
21 STABILITY AI INC., a Delaware corporation;  
22 MIDJOURNEY, INC., a Delaware corporation;  
23 DEVIANTART, INC., a Delaware corporation,  
24 Defendants.

25

26

27

28

DEVANTART'S ANTI-SLAPP  
MOTION TO STRIKE  
CASE NO. 3:23-cv-00201-WHO

Getty Images (US), Inc. v. Stability AI, Inc.  
(1:23-cv-00135)

Case 1:23-cv-00135-UNA Document 1 Filed 02/03/23 Page 1 of 36 PageID #: 1

IN THE UNITED STATES DISTRICT COURT  
FOR THE DISTRICT OF DELAWARE

GETTY IMAGES (US), INC. )  
Plaintiff, )  
v. ) C.A. No.:  
STABILITY AI, INC. ) **DEMAND FOR JURY TRIAL**  
Defendant. )

**COMPLAINT**

Plaintiff Getty Images (US), Inc. ("Getty Images" or "Plaintiff"), by and through its undersigned attorneys, for its Complaint against Defendant Stability AI, Inc. ("Stability AI" or "Defendant"), hereby alleges as follows:

**NATURE OF ACTION**

1. This case arises from Stability AI's brazen infringement of Getty Images' intellectual property on a staggering scale. Upon information and belief, Stability AI has copied *more than 12 million* photographs from Getty Images' collection, along with the associated captions and metadata, without permission from or compensation to Getty Images, as part of its efforts to build a competing business. As part of its unlawful scheme, Stability AI has removed or altered Getty Images' copyright management information, provided false copyright management information, and infringed Getty Images' famous trademarks.

2. Getty Images brings this action to recover damages that it has suffered and is continuing to suffer, and to prevent the irreparable harm caused by Stability AI's intentional and willful acts.

3. Getty Images is one of the world's leading creators and distributors of digital content. At great expense, over the course of nearly three decades, Getty Images has curated a

1

Doe 1, et al v. GitHub Inc., et al,  
(4:22-cv-06823)

Case 3:22-cv-06823 Document 1 Filed 11/03/22 Page 1 of 56

1 Joseph R. Saveri (State Bar No. 130064)  
2 Cadjo Zirpoli (State Bar No. 179108)  
3 Travis Manfredi (State Bar No. 281779)  
4 JOSEPH SAVERI LAW FIRM, LLP  
5 601 California Street, Suite 1000  
6 San Francisco, California 94108  
7 Telephone: (415) 500-6800  
8 Facsimile: (415) 395-9940  
9 Email: jsaveri@saverilawfirm.com  
10 czirpoli@saverilawfirm.com  
11 tmanfredi@saverilawfirm.com

12 Matthew Butterick (State Bar No. 250953)  
13 1920 Hillhurst Avenue, #406  
14 Los Angeles, CA 90027  
15 Telephone: (323) 968-2632  
16 Facsimile: (415) 395-9940  
17 Email: mb@buttericklaw.com

18 *Counsel for Individual and Representative  
19 Plaintiffs and the Proposed Class*

20 UNITED STATES DISTRICT COURT  
21 NORTHERN DISTRICT OF CALIFORNIA  
22 SAN FRANCISCO DIVISION

23 J. DOE 1 and J. DOE 2, individually and on  
24 behalf of all others similarly situated,  
25 Individual and Representative Plaintiffs,  
26 v.  
27 GITHUB, INC., a Delaware corporation;  
28 MICROSOFT CORPORATION, a Washington  
corporation; OPENAI, INC., a Delaware  
nonprofit corporation; OPENAI, L.P., a  
Delaware limited partnership; OPENAI GP,  
L.L.C., a Delaware limited liability company;  
OPENAI STARTUP FUND GP I, L.L.C., a  
Delaware limited liability company; OPENAI  
STARTUP FUND I, L.P., a Delaware limited  
partnership; OPENAI STARTUP FUND  
MANAGEMENT, LLC, a Delaware limited  
liability company,  
Defendants.

Case No.  
**COMPLAINT**  
**CLASS ACTION**  
**DEMAND FOR JURY TRIAL**

CLASS ACTION COMPLAINT

# Är regelverket tillräckligt balanserat?

## Ytterligare perspektiv

- Upphovsrättens särskilda syfte är ”att säkerställa skyddet av de berörda rättighetsinnehavarnas rätt att kommersiellt utnyttja att de skyddade verken förs ut på marknaden eller tillhandahålls i form av licenser som beviljas mot betalning”.

**C-403/08 och C-429/08 (Football Association Premier League m.fl.), punkt 107**

- ”The development of new products and services based on the Data Economy, specifically those relying on text and data mining and the reuse of vast public sector information, will offer a good opportunity to revisit the ’internal’ adjustments in copyright laws and examine their effects in the market.”

**ALAI (2021)**

- ”Researchers expect to access and process the full content of the research literature with their computer programs and should be able to use their machines as they use their eyes. The right to read is the right to mine.”

**Murray-Rust m.fl. (2014)**

- ”[W]hen acts of reproduction are carried out for the purpose of search and TDM, the work is not used as a work... The expressive features of the work are not used, and there is no public to enjoy the work, as the work is only an input in a process for searching a corpus and identifying occurrences and possible trends or patterns.”

**Ducato & Strowel (2018)**

- ”[T]he opt-out mechanism that applies to the broader unqualified exception of Art. 4 strongly limits the effectiveness of the reform and its ability to promote competitive advantage for European companies engaged in AI and other disruptive technology development”

**Geiger m.fl. (2019)**

- ”[N]on-expressive uses of copyrighted works – i.e. acts of copying that do not communicate the author’s original expression to the public – should not generally be regarded as infringing”

**Sag (2009)**

# Tack för uppmärksamheten!



Tobias Kempas  
Counsel, Advokat  
Immaterialrätt, Marknadsrätt, Medierätt

010 614 30 37  
070 714 30 37  
Tobias.Kempas@vinge.se





VINGE

STOCKHOLM GÖTEBORG MALMÖ HELSINGBORG BRYSEL