

Computer Vision 3

Ș.I. dr. ing. Mihai DOGARIU

www.mdogariu.aimultimedialab.ro

Organizare



Ș.I. dr. ing. Mihai DOGARIU

Email: mihai.dogariu@upb.ro

Web: www.mdogariu.aimutimedialab.ro

Materiale curs & laborator:

<https://github.com/MihaiDogariu/CV3>

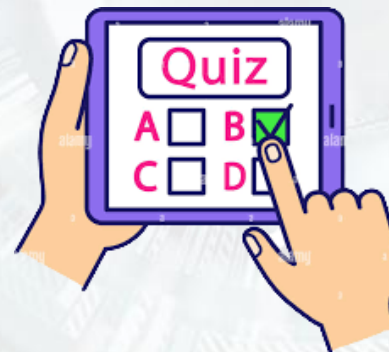
Punctaj:

- 50% colocviu laborator
 - 50% examen final
- } minim 50% din total pentru promovare

Organizare

Bonusuri (facultative):

- quiz săptămânal:
 - în aceeași zi a săptămânii (de stabilit)
 - 2-3 întrebări, aprox. 1 min/întrebare
 - Clasament (Q1: +20 pct, Q2: +10 pct)
- prelegere la curs:
 - 5-10 minute la începutul fiecărui curs
 - stabilim tema împreună în avans
 - maxim 2 prelegeri, +10 pct/prelegere



Organizare

Pre-requisites:

- Python - cunoștințe de bază
- Git/GitHub - cunoștințe minime (<https://github.com/git-guides>)
- CV 1 & CV 2
- IC 1 & IC 2
- PyTorch – reprezintă un avantaj

Structura cursului



M1. Introducere

M2. Fundamentele Învățării Adânci (Deep Learning Fundamentals)

M3. Învățare Adâncă Supervizată (Supervised Deep Learning)

M4. Învățare Adâncă Nesupervizată (Unsupervised Deep Learning)

M1. Introducere

1.1. Motivația

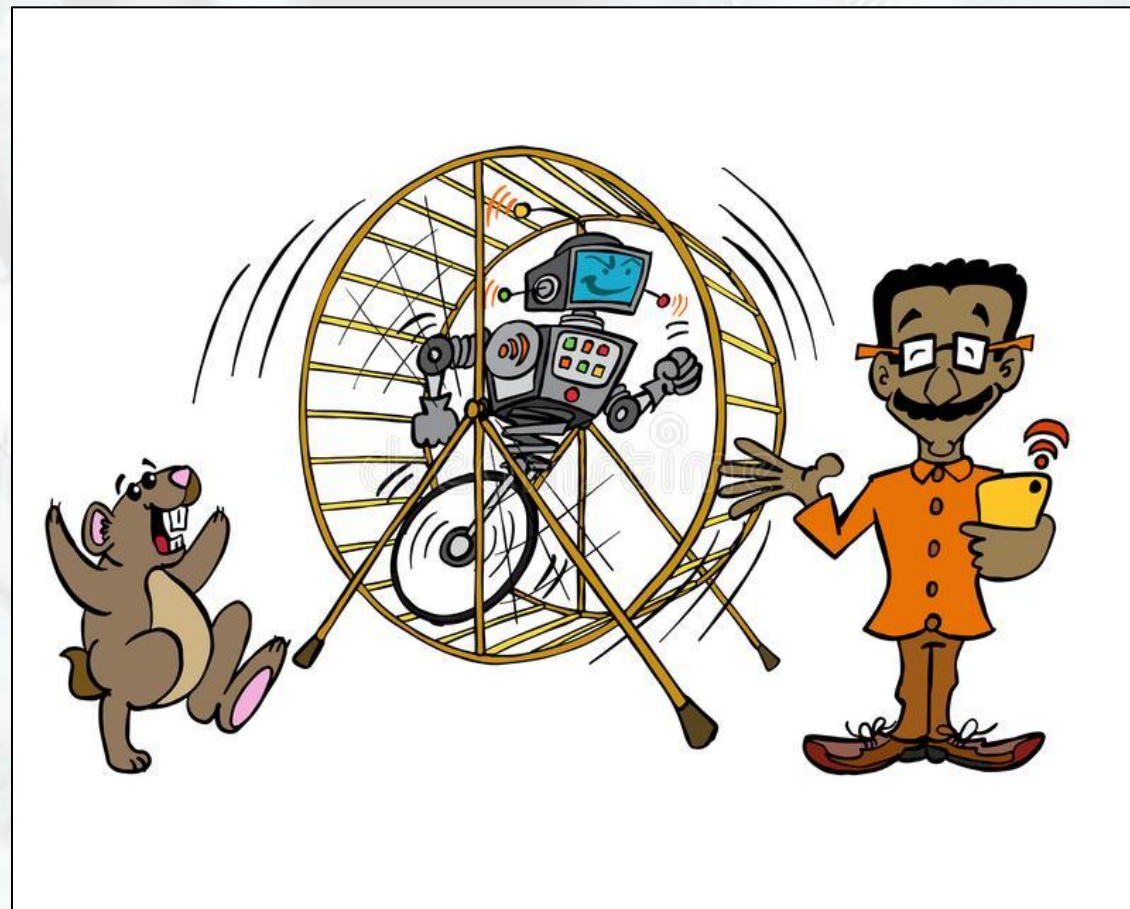
1.2. Scurt istoric

1.3. Aplicații existente

1.4. Utilitare

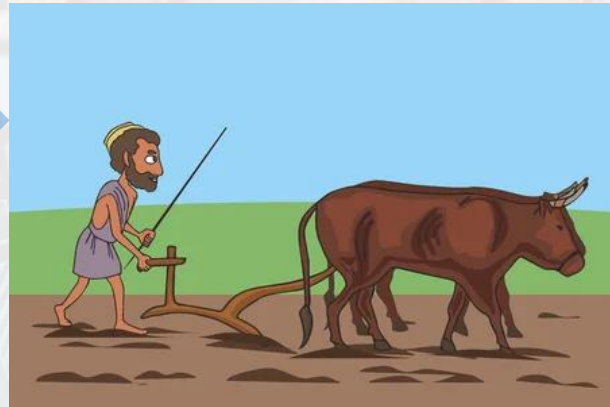
Identificarea nevoii

- Nevoia de automatizare a unor procese (de cele mai multe ori, redundante).
- Capacitatea umană de lucru este limitată și supusă greșelilor, mai ales după un timp îndelungat de lucru (~6 ore).
- Creșterea raportului eficiență/cost.



Rezolvarea problemei

- Utilizarea „mașinilor” pentru a lucra în locul nostru.



Rezolvarea problemei => alte probleme

- Cum înlocuim partea umană?

Inteligență artificială

- Cum evaluăm performanța unui astfel de sistem?

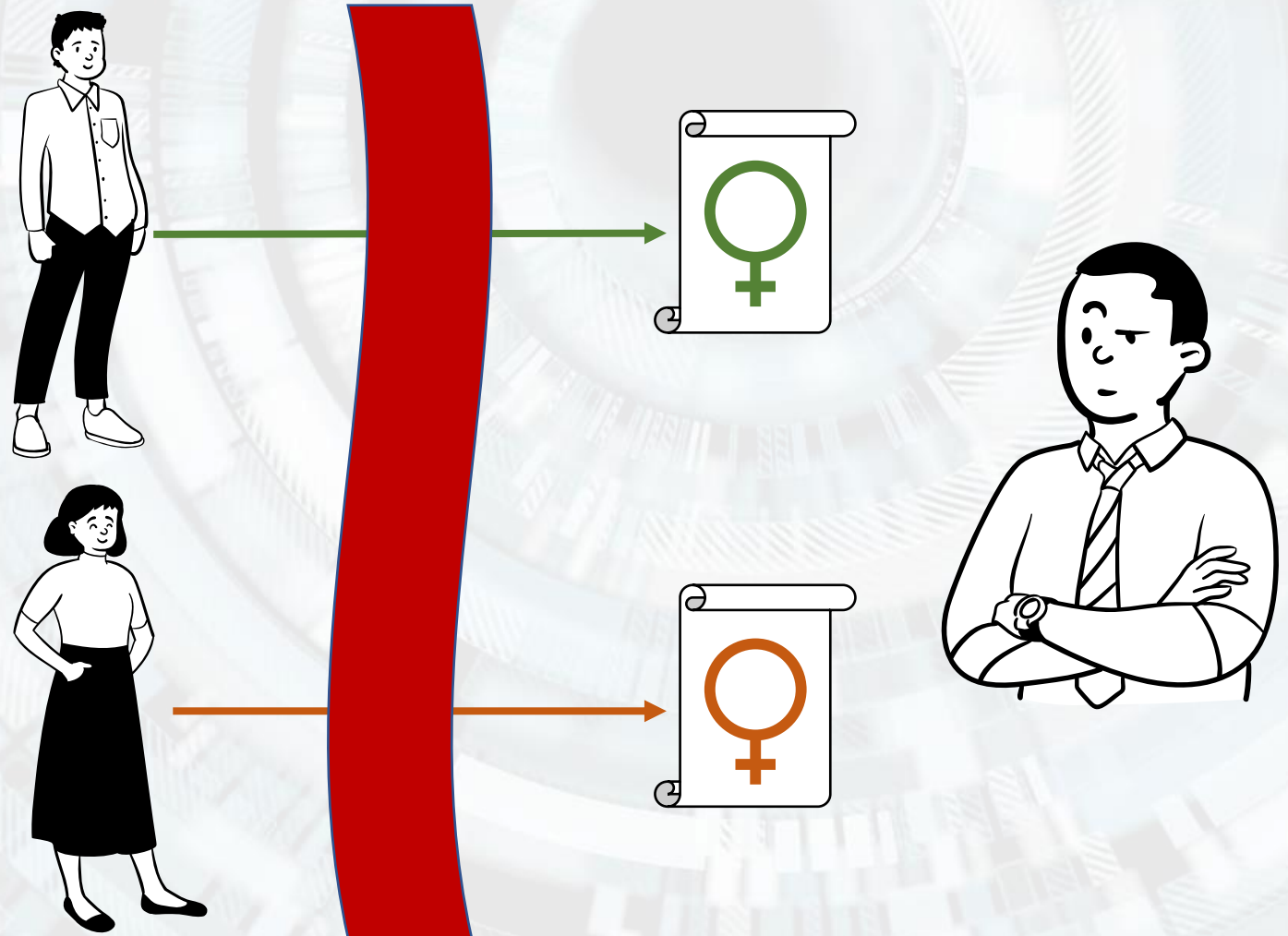
Metrici: mAP, accuracy, SSIM, RMSE etc.

- Când decidem că s-a ajuns la inteligență artificială pură?

Test Turing, Test Feigenbaum, CAPTCHA, etc.

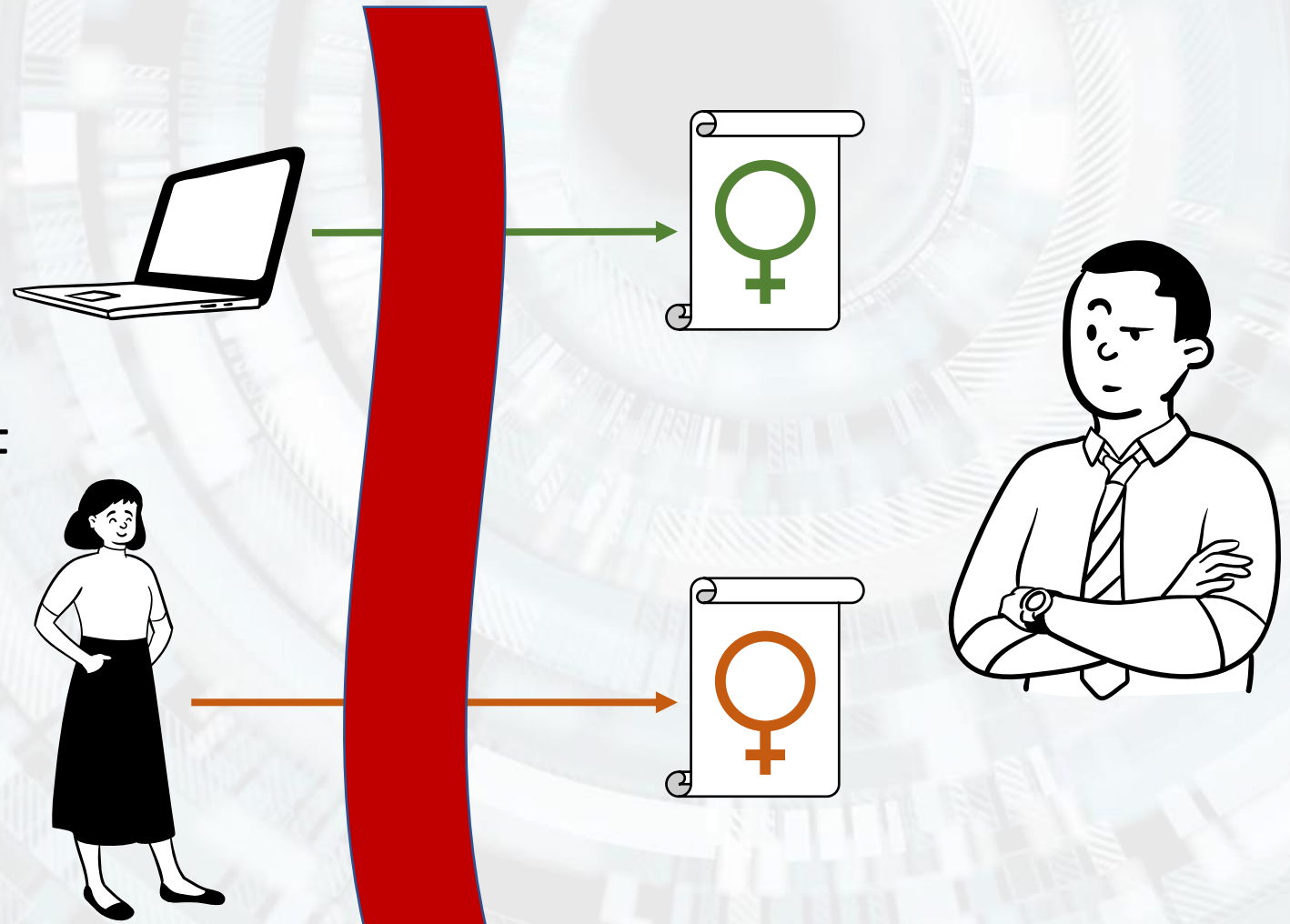
Testul Turing v1 (Imitation game)

- 2 persoane (M+F) scriu un bilet către arbitru
- M încearcă să convingă arbitrul că este F
- F încearcă să convingă arbitrul că este F
- Arbitrul trebuie să își dea seama cine este M și cine este F



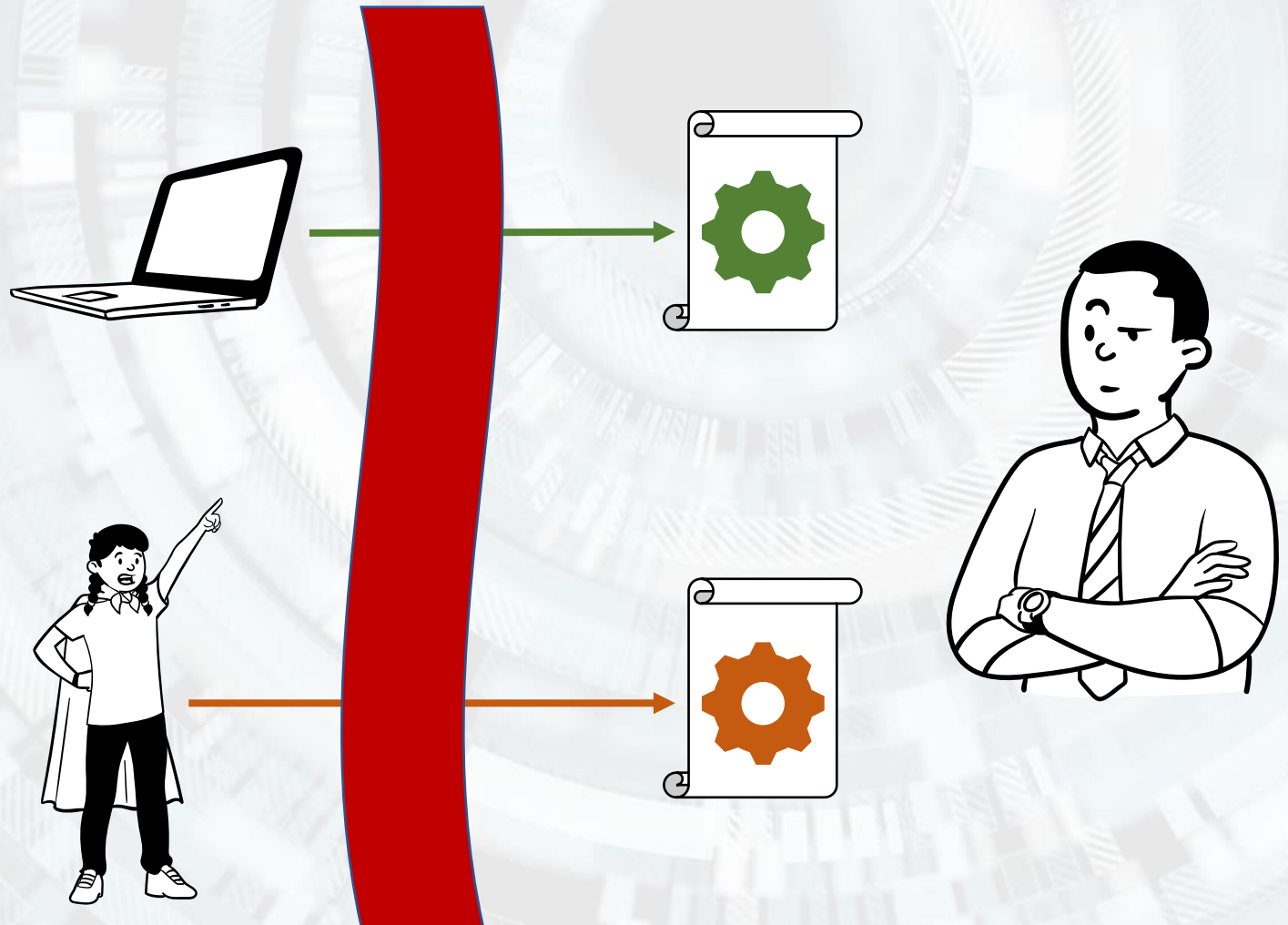
Testul Turing v2

- M este înlocuit de un calculator
- Arbitrul trebuie să își dea seama cine este calculatorul și cine este F



Testul Feigenbaum

- M este înlocuit de un calculator
- F este înlocuită de un expert în domeniu
- Calculatorul și expertul trebuie să rezolve aceeași sarcină de lucru
- Cine se descurcă mai bine?



AI, ML, DL, CV?

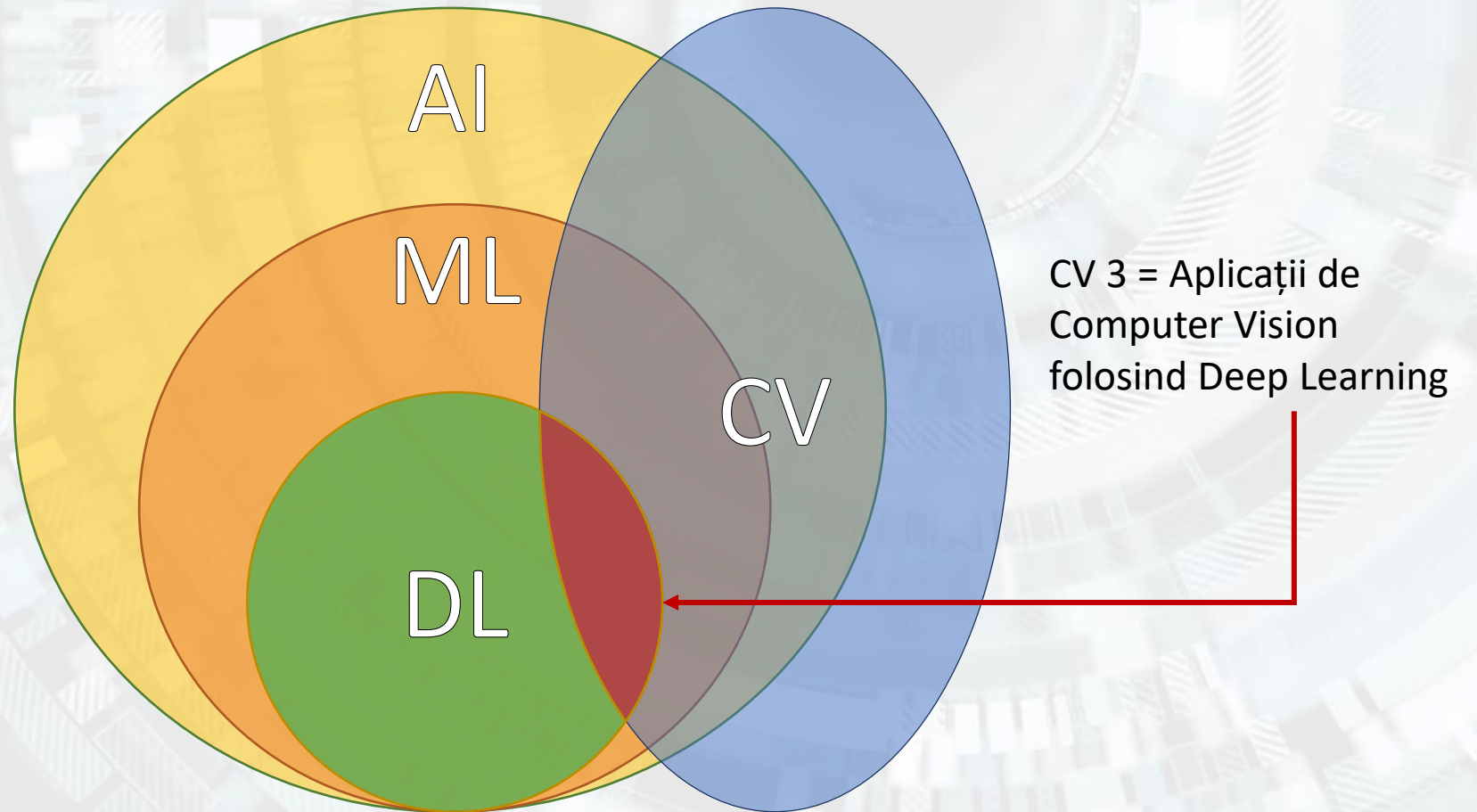
Artificial Intelligence (AI) = abilitatea unui calculator de a îndeplini sarcini pentru care este necesară inteligența (non-/umană).

Machine Learning (ML) = subdomeniu al AI în care sistemele sunt dezvoltate cu capacitatea de a învăța pe baza exemplelor întâlnite.

Deep Learning (DL) = subdomeniu al ML în care sistemele sunt dezvoltate pe modelul rețelei de neuroni a creierului uman.

Computer Vision (CV) = domeniu interdisciplinar în care calculatoarele sunt învățate să interpreteze la nivel înalt informația din imagini digitale.

AI, ML, DL, CV?



Scurt istoric

Scurt istoric

Walter Pitts & Warren McCulloch – model matematic ce imită funcționalitatea unui neuron uman: neuron artificial.

1943

Scurt istoric

Frank Rosenblatt – primul algoritm de clasificare binară supervizat bazat pe neuroni artificiali: perceptronul.



Scurt istoric

Henry J. Kelley – primul model de propagare înapoi continuă: backpropagation.



Scurt istoric

Stuart Dreyfus – model de propagare înapoi folosind regula înlănțuirii derivatelor: backpropagation with chain rule.



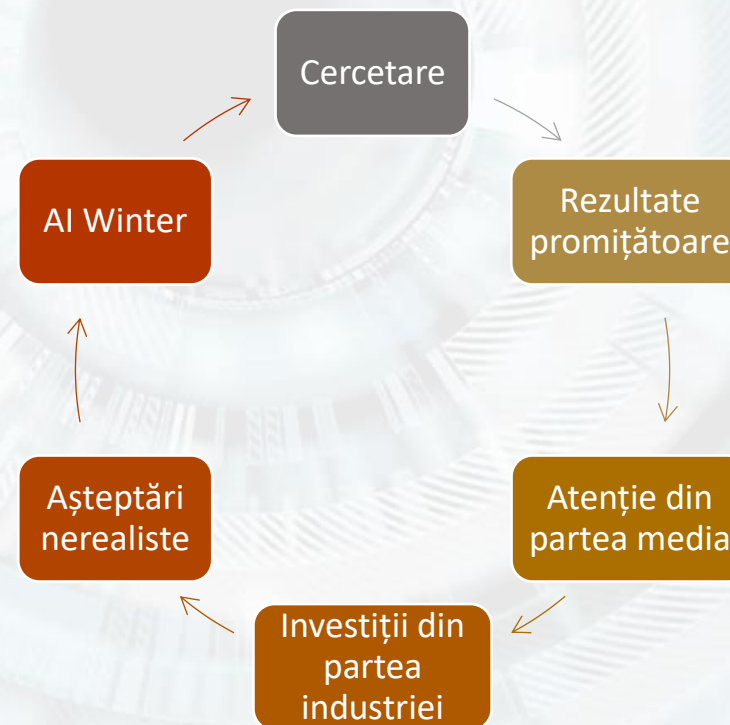
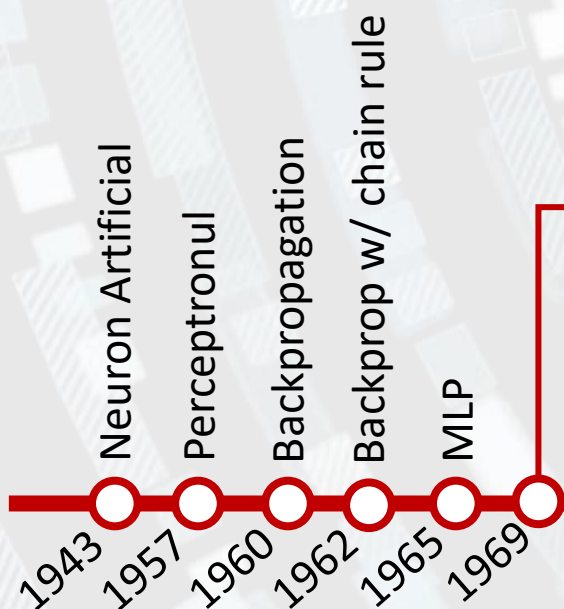
Scurt istoric

Alexey Grigoryevich Ivakhnenko & Valentin Grigorevich Lapa—
reprezentare ierarhică a unei rețele neuronale: primul
perceptron multistrat (MLP).



Scurt istoric

Marvin Minsky & Seymour Papert –
limitările rețelelor neuronale; marchează
începutul „iernii AI”.



Scurt istoric

Seppo Linnainmaa – implementează propagarea înapoi în cod.



Scurt istoric

Kunihiko Fukushima – prima rețea convoluțională: Neocognitron.
Poate recunoaște litere.



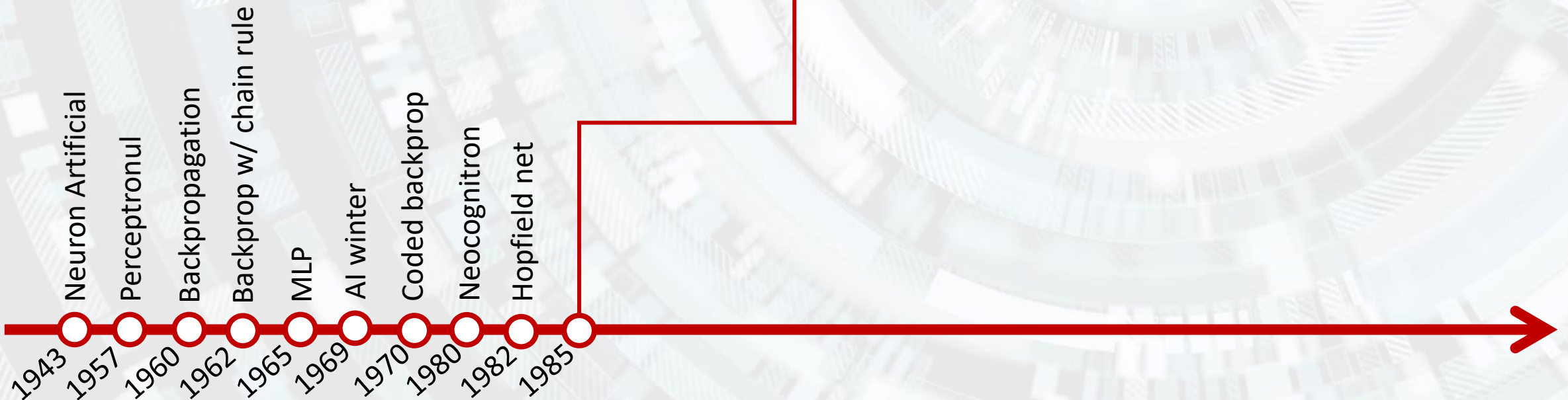
Scurt istoric

John Hopfield – rețelele Hopfield, precursor al rețelelor recurente.



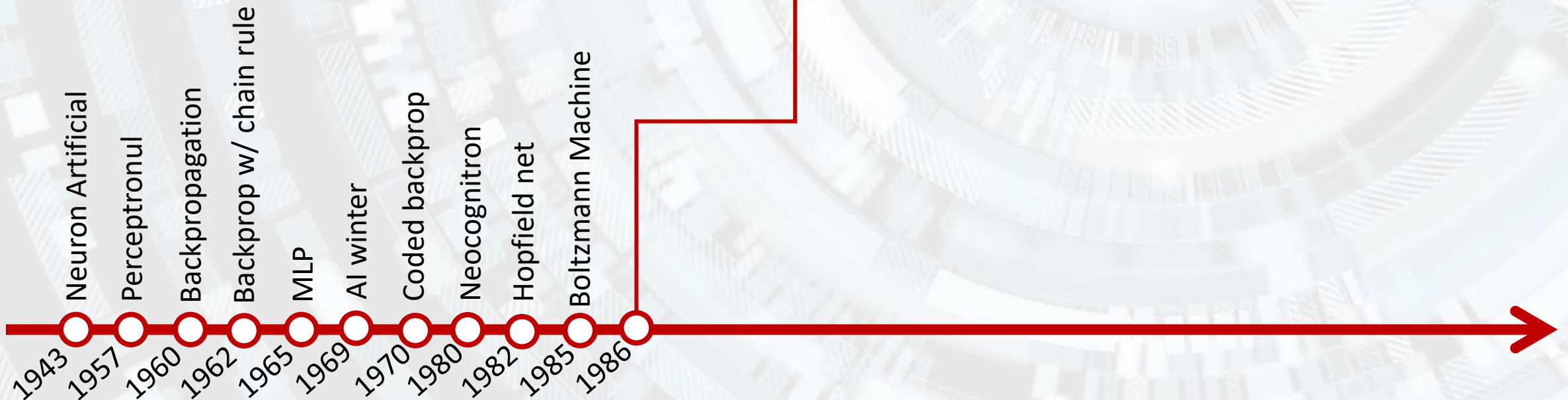
Scurt istoric

David Ackley, Geoffrey Hinton & Terrence Sejnowski – rețea neuronală recurentă stocastică: Boltzmann Machine.



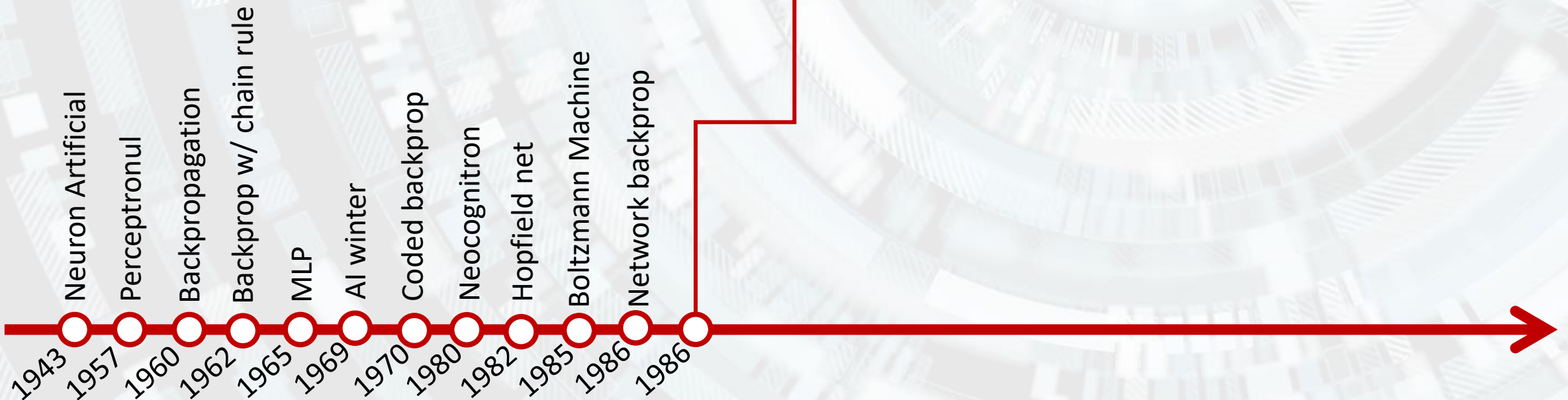
Scurt istoric

David Rumelhart, Geoffrey Hinton and Ronald Williams – implementarea propagării înapoi în rețele neuronale. Milestone.



Scurt istoric

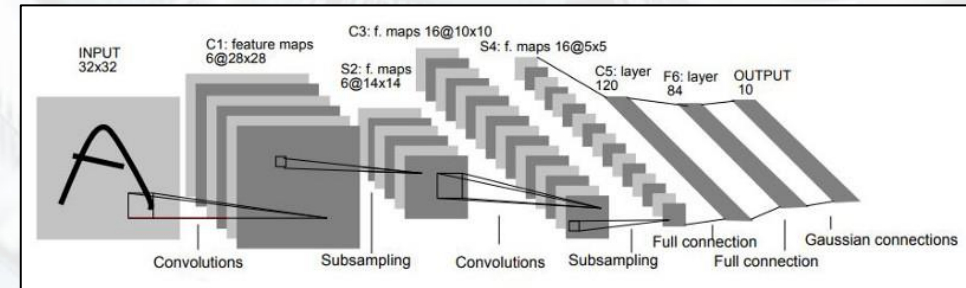
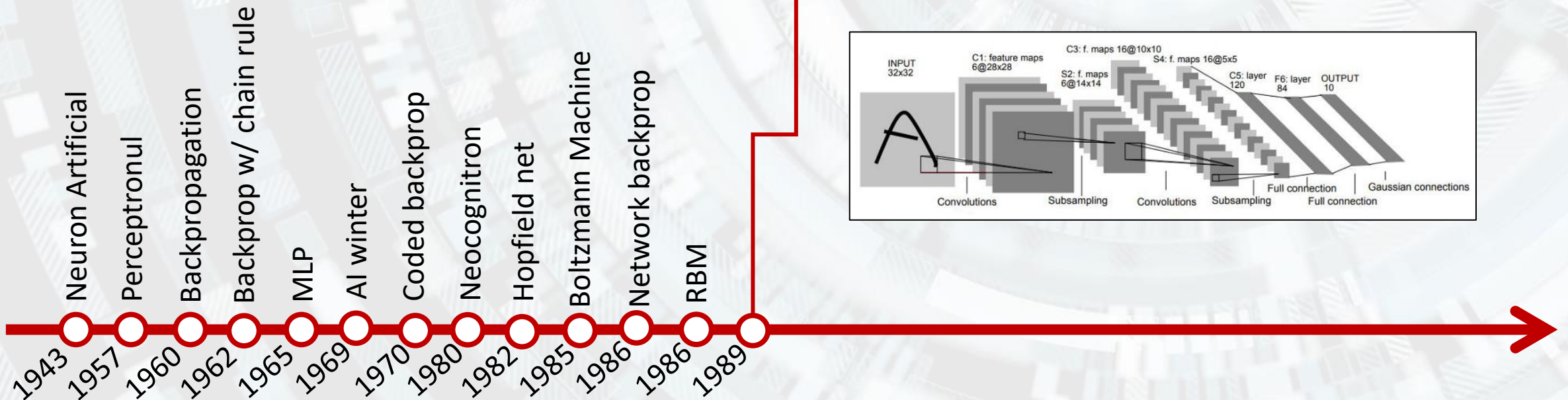
Paul Smolensky – o versiune nouă de Boltzmann Machine, cu restricții: Restricted Boltzmann Machine (RBM).



Scurt istoric

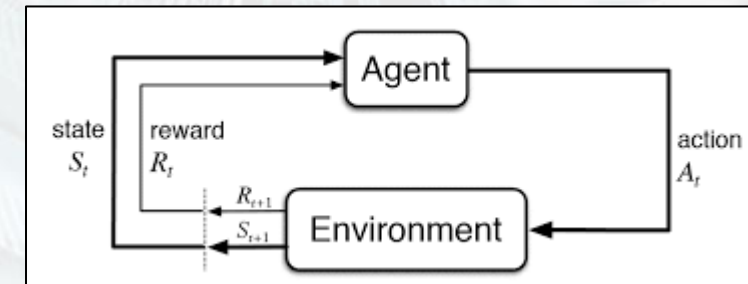
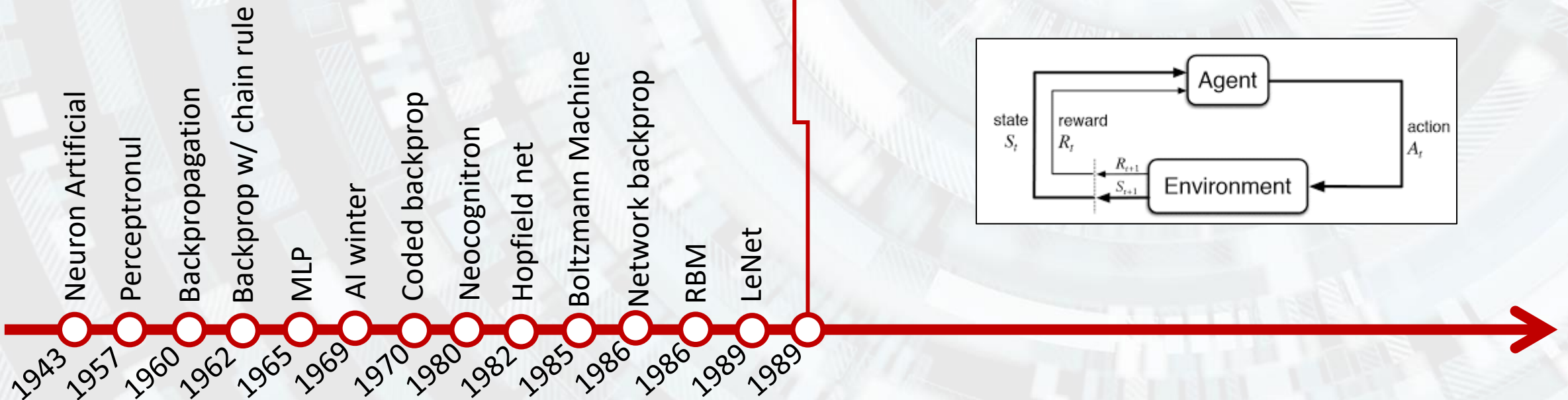
Yann LeCun – prima rețea neuronală convoluțională care poate recunoaște cifre scrise de mână rapid și cu acuratețe: LeNet.

Milestone



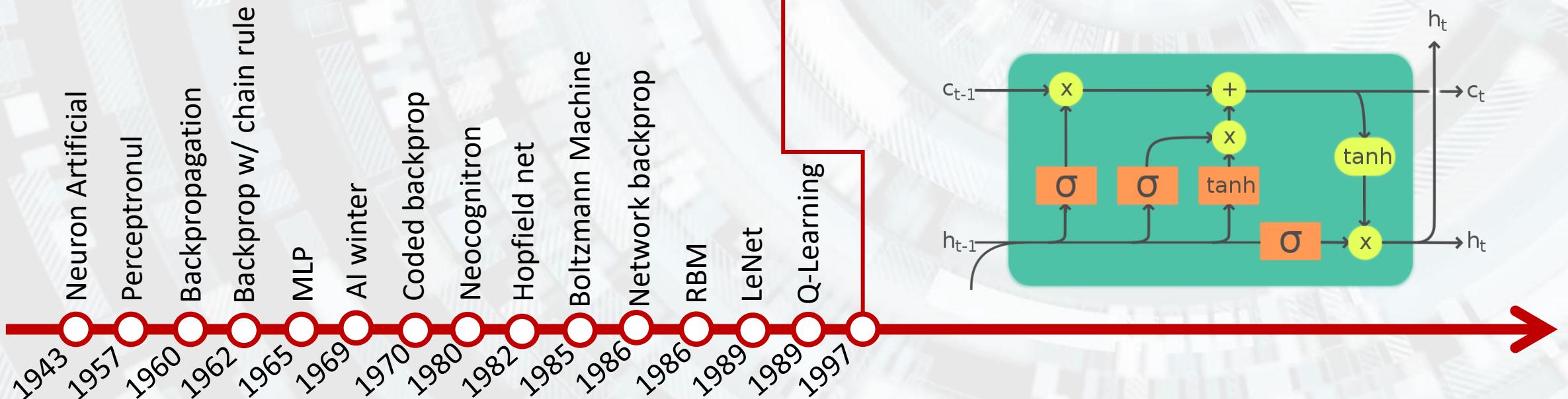
Scurt istoric

Christopher Watkins – progrese în domeniul învățării consolidate:
Q-Learning.



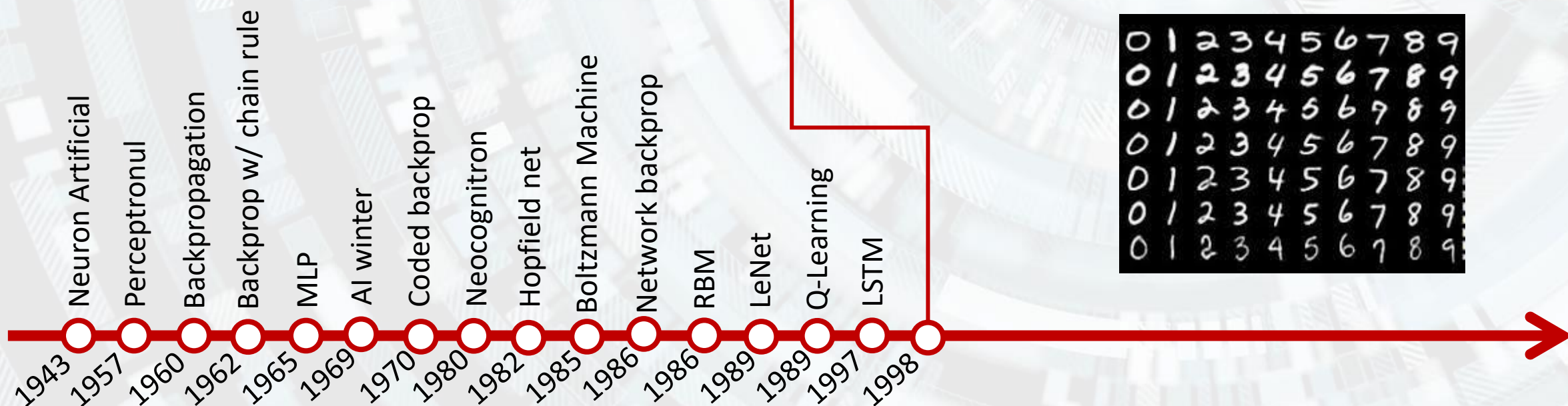
Scurt istoric

Sepp Hochreiter & Jürgen Schmidhuber – model de rețea neuronală recurentă revoluționar: Long Short-Term Memory (LSTM). Milestone.



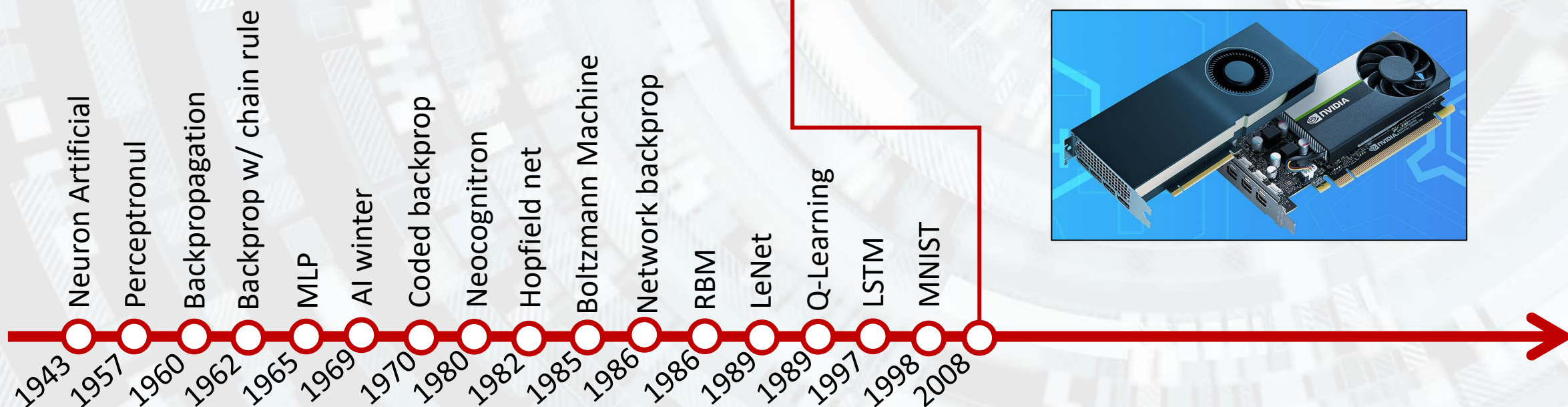
Scurt istoric

Yann LeCun – bază de date pentru recunoașterea cifrelor scrise de mână: MNIST.



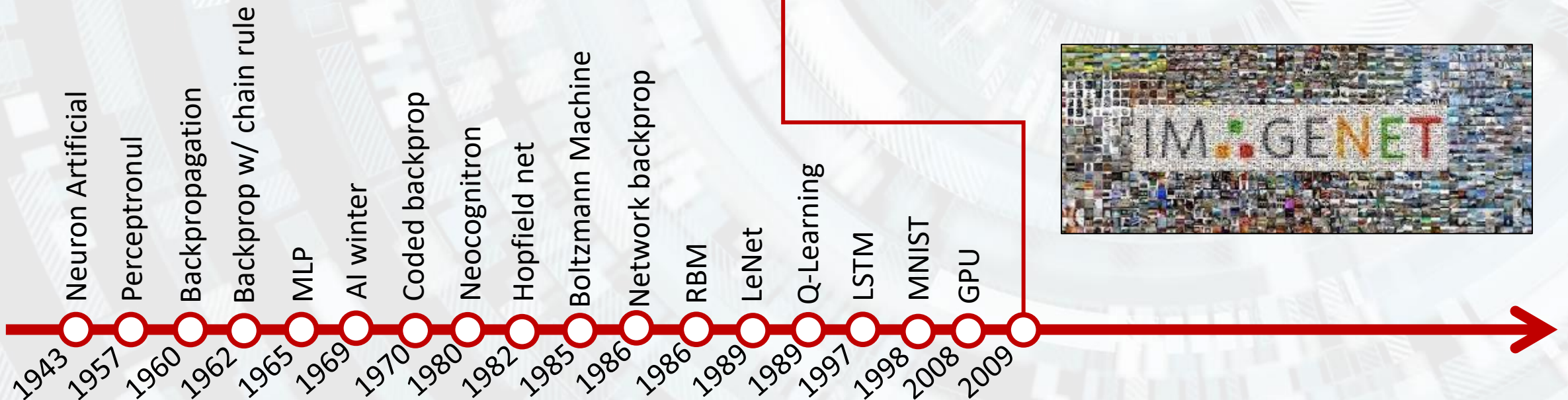
Scurt istoric

Andrew Ng – atrage atenția asupra utilizării plăcilor grafice (GPU) pentru antrenarea rețelelor neuronale.



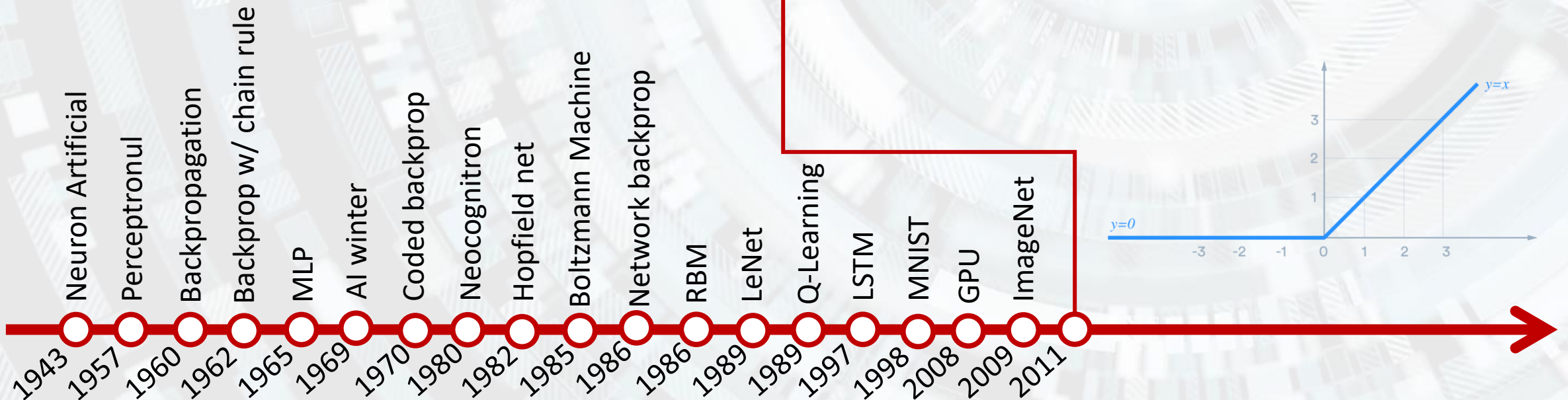
Scurt istoric

Fei-Fei Li – lansează cea mai populară (și exhaustivă la acel moment) bază de date de imagini: ImageNet.



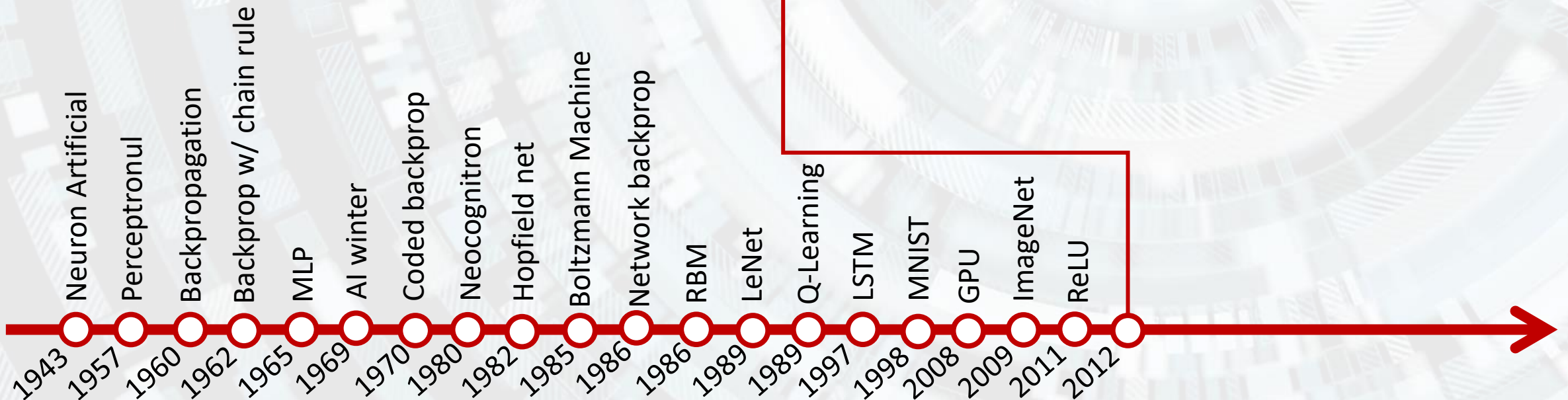
Scurt istoric

Yoshua Bengio, Antoine Bordes & Xavier Glorot – propun o metodă pentru combaterea dispariției gradientului: Rectified Linear Unit (ReLU).



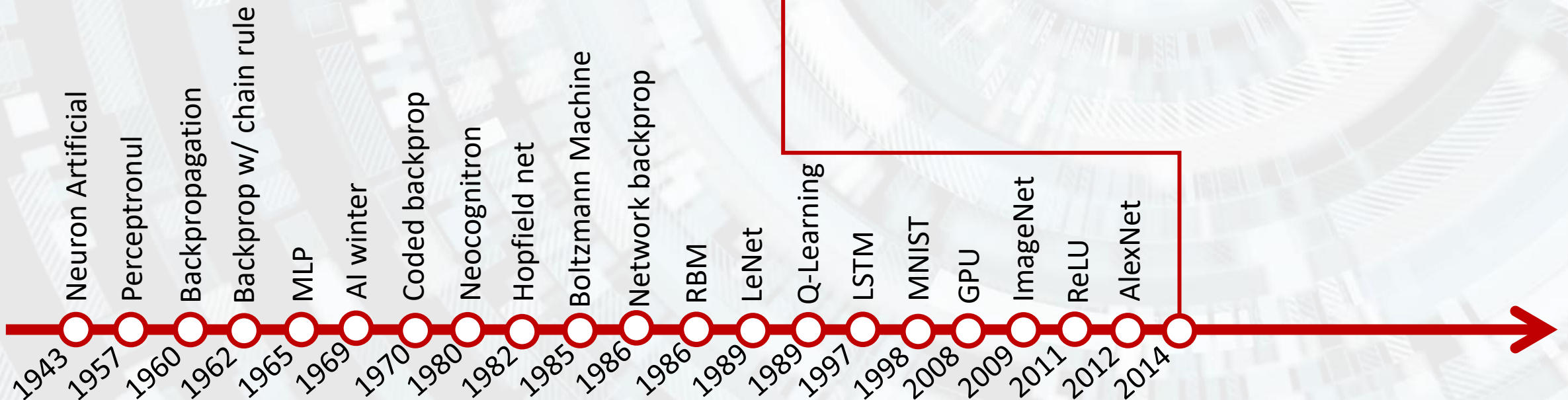
Scurt istoric

Alex Krizhevsky – propune un model de rețea neuronală convoluțională ce depășește categoric cel mai bun rezultat la acel moment: AlexNet. Milestone.



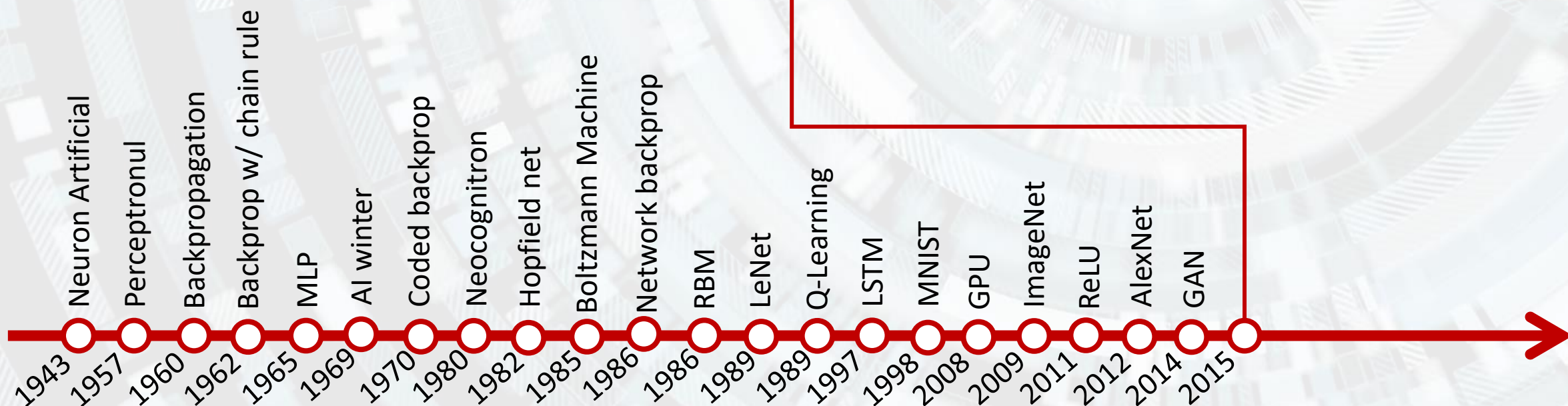
Scurt istoric

Ian Goodfellow – propune un model de rețea care poate sintetiza imagini realiste: Generative Adversarial Network (GAN).
Milestone.



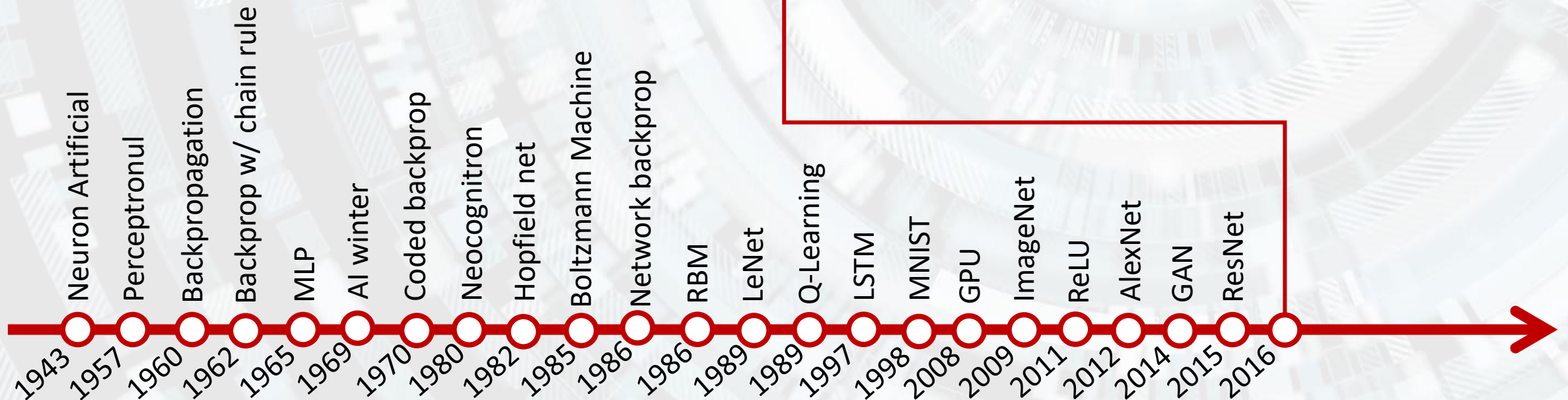
Scurt istoric

Kaiming He – prima rețea neuronală foarte adâncă (sute de straturi): ResNet. Milestone.



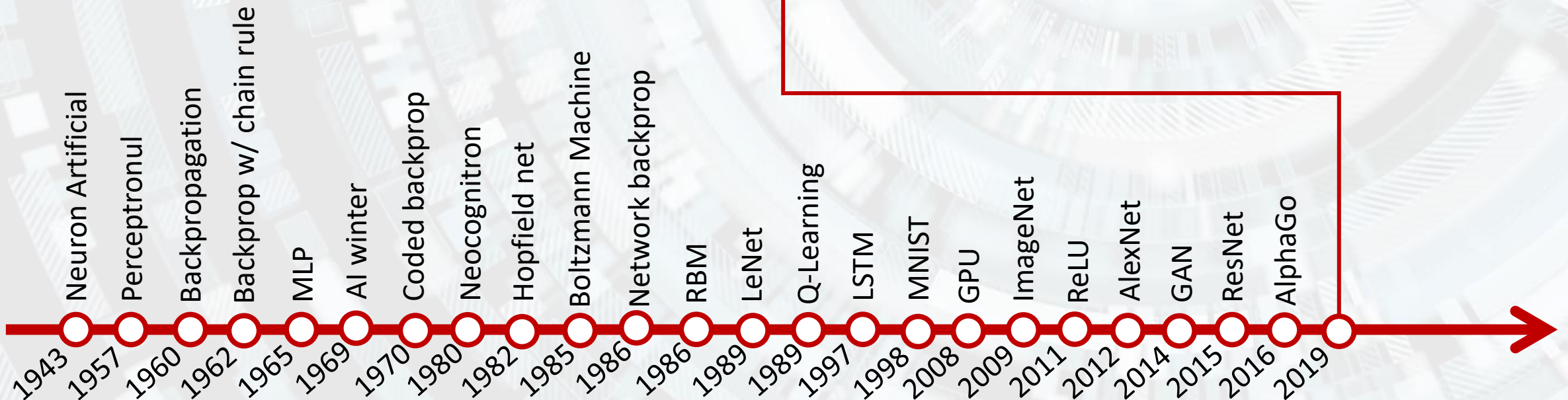
Scurt istoric

DeepMind – departamentul de AI al Google creează o rețea neuronală care bate cel mai bun jucător de Go al lumii: AlphaGo.



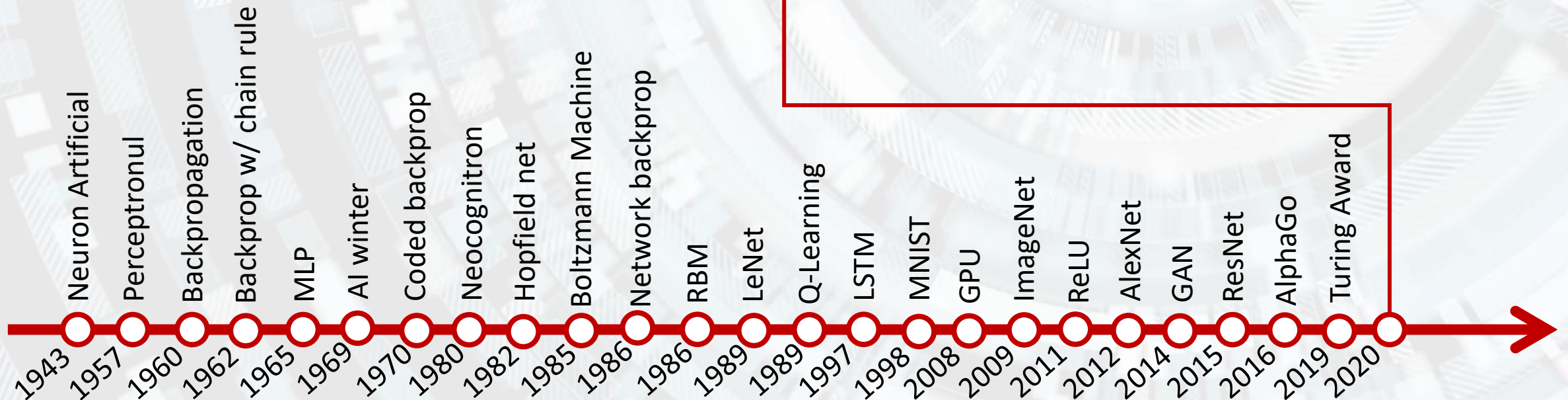
Scurt istoric

Yoshua Bengio, Geoffrey Hinton & Yann LeCun – Turing Award.



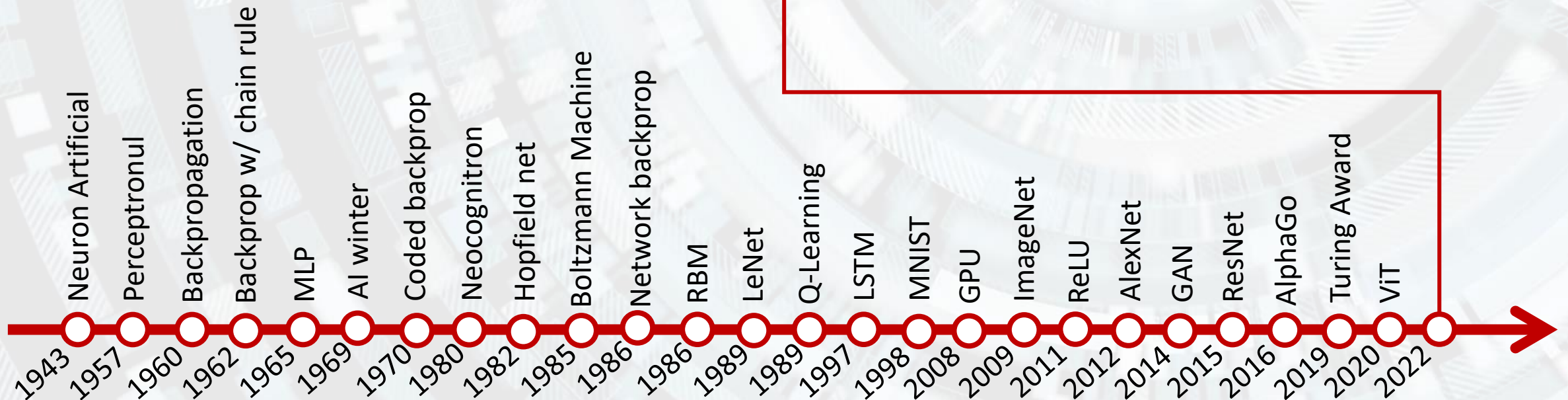
Scurt istoric

Alexey Dosovitskiy et al – o echipă a Google Brain adaptează arhitectura de tip Transformer pentru recunoașterea imaginilor – Visual Transformers (ViT).



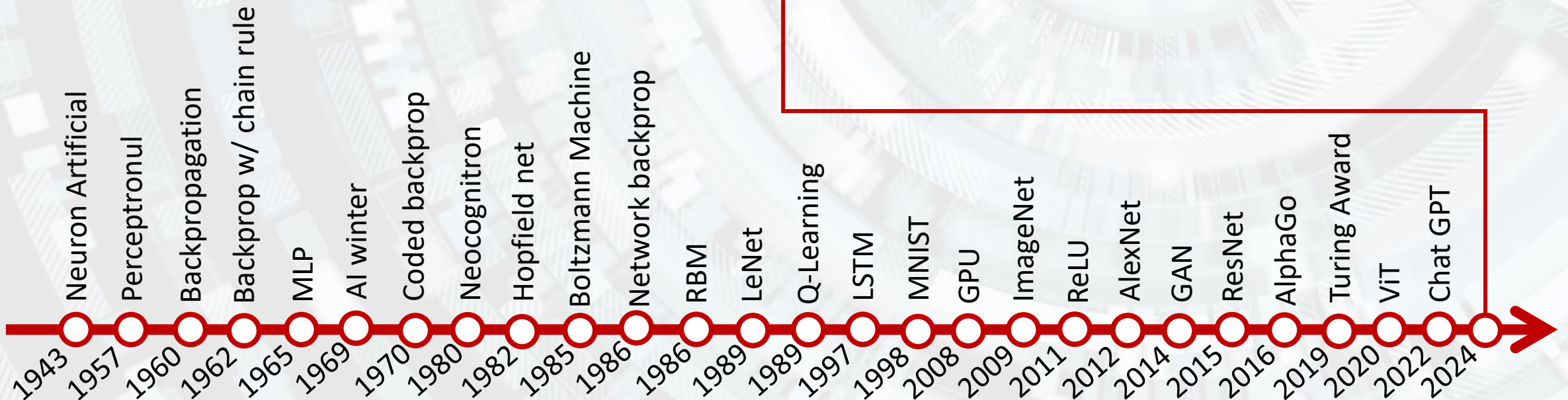
Scurt istoric

Chat GPT – Open AI lansează cel mai performant chatbot.



Scurt istoric

John Hopfield și Geoffrey Hinton primesc premiul Nobel pentru fizică *“for foundational discoveries and inventions that enable machine learning with artificial neural networks”*.



Explozia Deep Learning

Ce a dus la avansul exponențial al Deep Learning în ultima perioadă?

1. putere de calcul superioară (hardware) – GPU
2. mult mai multe date disponibile => rezultate mai bune
3. framework-uri optimizate (software): Tensorflow, PyTorch, Caffe, MXNet etc.
4. efort financiar și atenție din partea industriei: Facebook AI Research (FAIR), Google Deepmind, NVIDIA, Microsoft Research, AWS Deep Learning etc.

Exemple celebre

Șah: Deep Blue vs Gary Kasparov (1997)



Exemple celebre

Jeopardy: IBM Watson vs
Brad Rutter & Ken Jennings
(2011)



Exemple celebre

Go: Google AlphaGo vs
Lee Sedol (2016)



Aplicații concrete

1. Clasificare de imagini

- input: imagine
- output: scor de apartenență la diferite clase

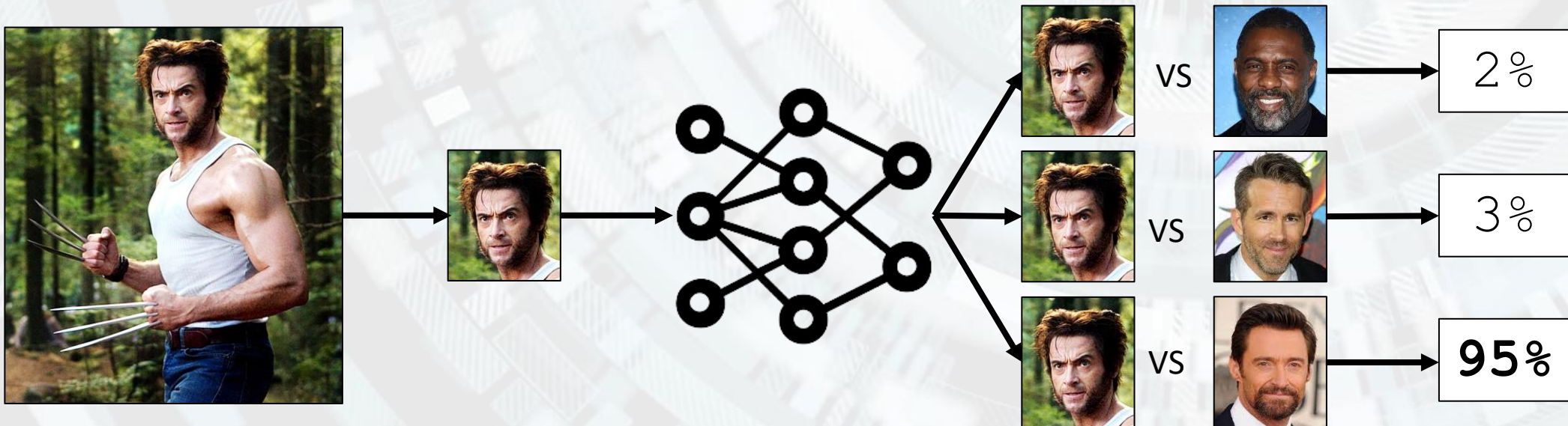


Raton, 90.50%
Chitară, 7.00%
Urs, 1.50%

Aplicații concrete

2. Recunoaștere de fețe (localizare + clasificare, un singur obiect)

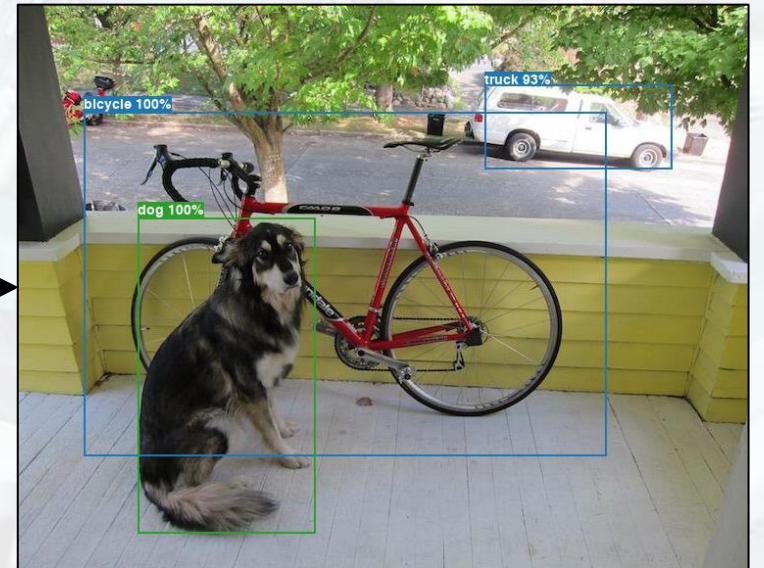
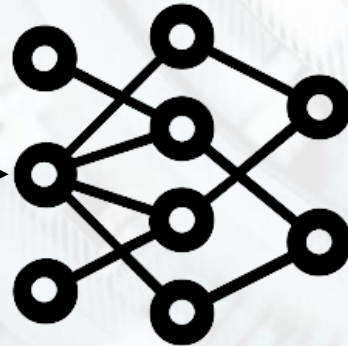
- input: imagine cu fața unei persoane
- output: identitatea persoanei



Aplicații concrete

3. Detecție automată de obiecte

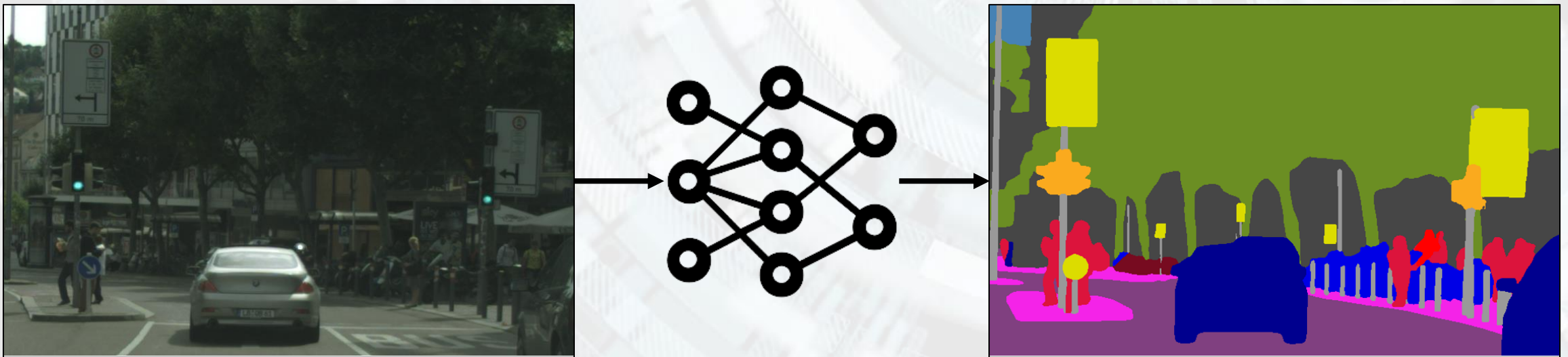
- input: imagine
- output: $[x1, x2, y1, y2, \text{etichetă}, \text{scor}]$ pentru fiecare obiect



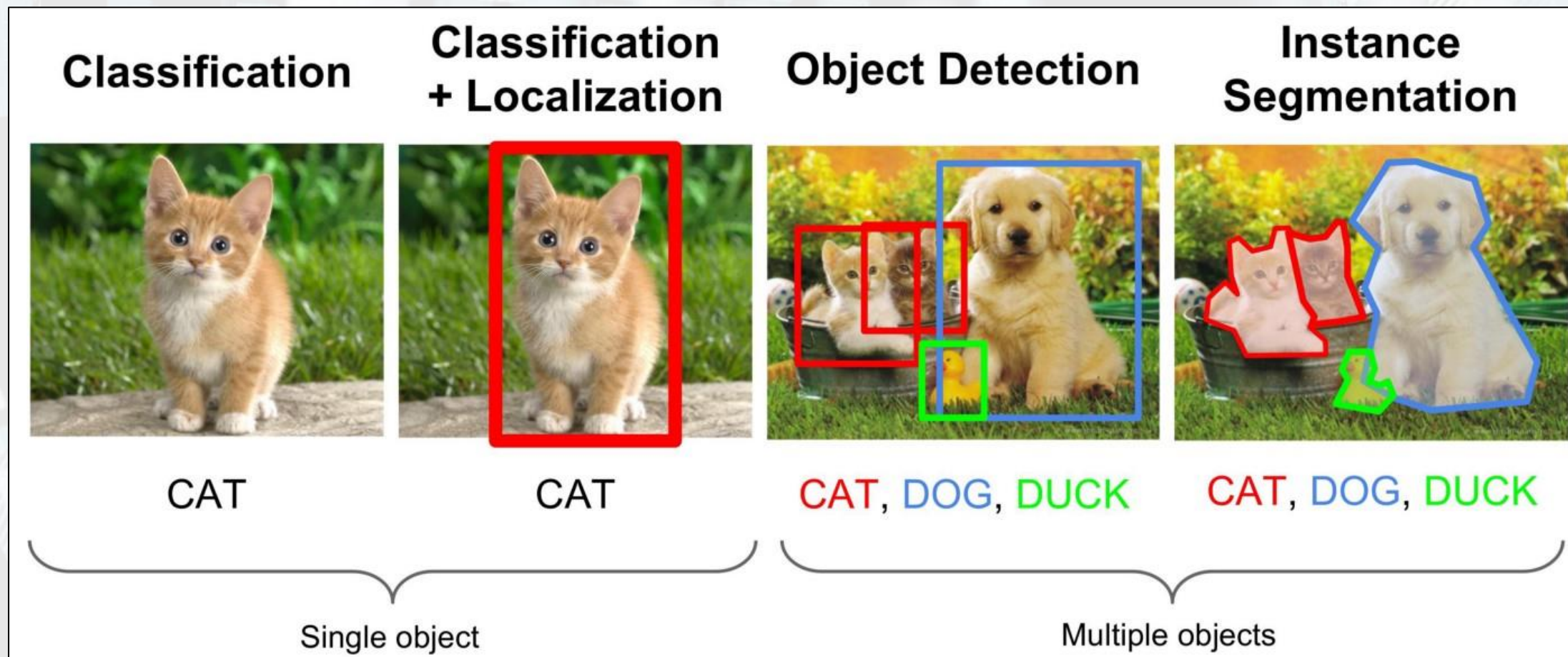
Aplicații concrete

4. Segmentare semantică

- input: imagine
- output: imagine cu o paletă redusă de culori



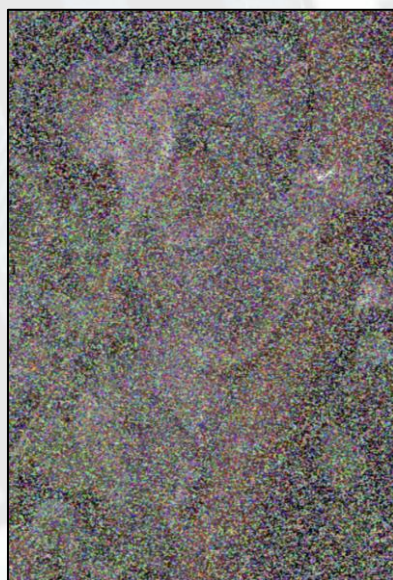
Aplicații concrete



Aplicații concrete

5. Restaurare de imagini – denoising

- input: imagine cu zgomot
- output: imagine fără zgomot



VS



Aplicații concrete

6. Restaurare de imagini – inpainting

- input: imagine cu zone lipsă
- output: imagine completă



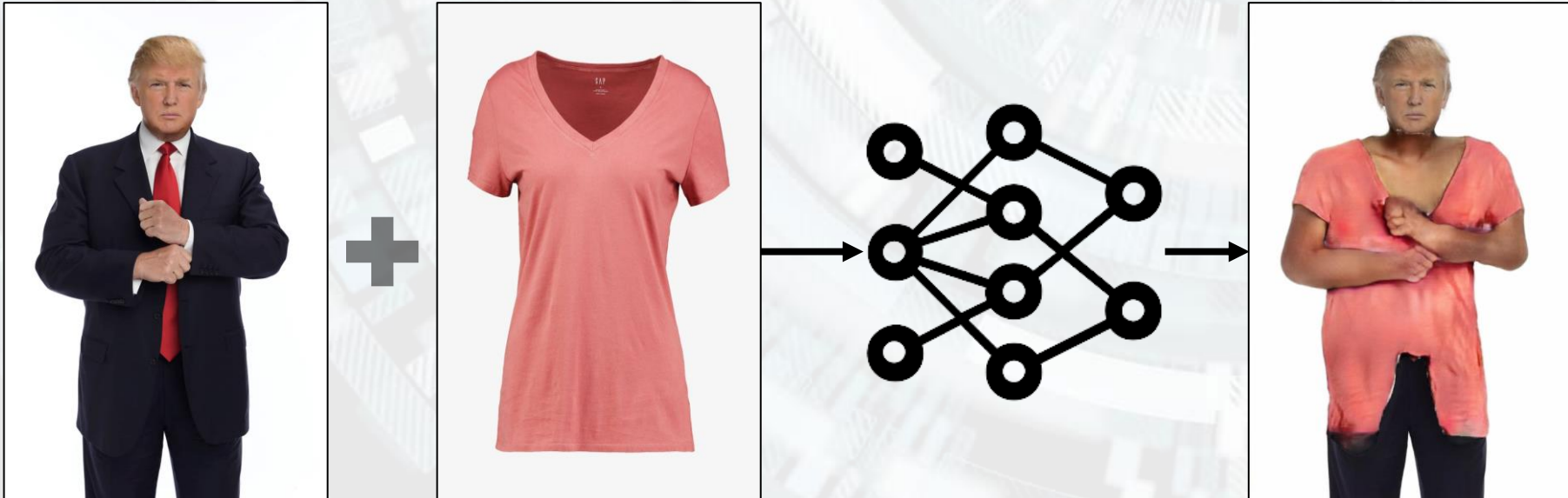
VS



Aplicații concrete

7. Transfer de stil

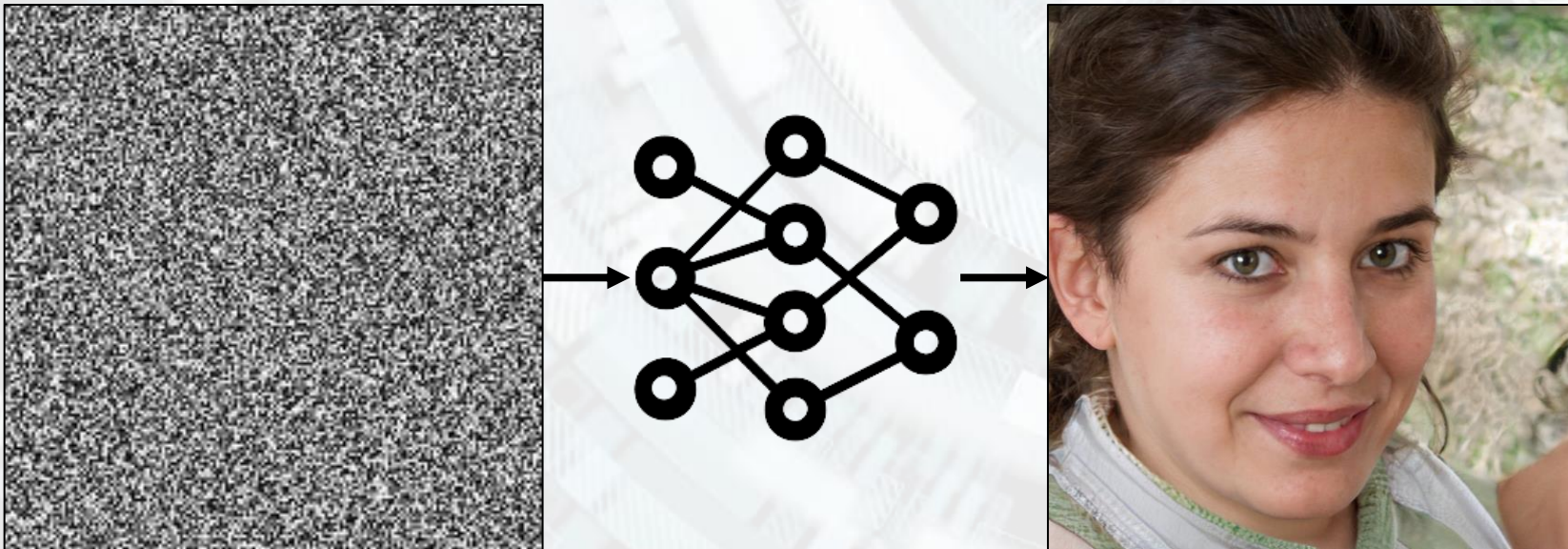
- input: imagine țintă + stil nou
- output: imagine cu noul stil



Aplicații concrete

8. Generare de imagini

- input: zgomot alb
- output: imagine HD



Utilitare

Deep Learning Frameworks:

- PyTorch – Meta AI;
- Tensorflow – Google;
- Keras – Google, front-end pentru Tensorflow;
- MXNet – Apache Software Foundation;
- DeepLearning4J (DL4J) – Konduit;
- Caffe – Berkeley Vision and Learning Center;
- Microsoft Cognitive Toolkit (CNTK) – Microsoft;



obsolete

Sfârșit M1

Bibliografie

- [1] <https://machinelearningknowledge.ai/brief-history-of-deep-learning/>, accesat octombrie 2022.
- [2] Turing, A. M., & Haugeland, J. (1950). Computing machinery and intelligence. The Turing Test: Verbal Behavior as the Hallmark of Intelligence, 29-56.
- [3] Feigenbaum, E. A. (2003). Some challenges and grand challenges for computational intelligence. Journal of the ACM (JACM), 50(1), 32-40.