

# Intelligence artificielle, apprentissage profond: état de l'art et éthique

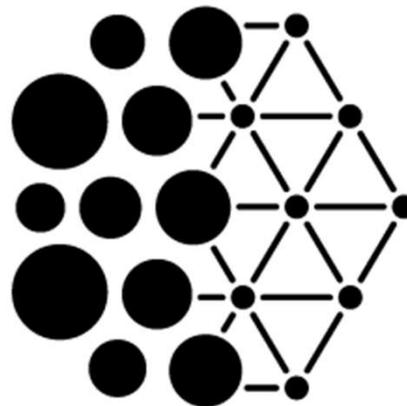
**Yoshua Bengio**

21 septembre 2018

Congrès de l'APQBQ, CHU Sainte-Justine, Montréal

**ICRA**  
INSTITUT  
CANADIEN  
DE  
RECHERCHES  
AVANCÉES

Université   
de Montréal



**Mila**



# Les voitures commencent à se conduire toutes seules...

(mais il y a encore des défis...)



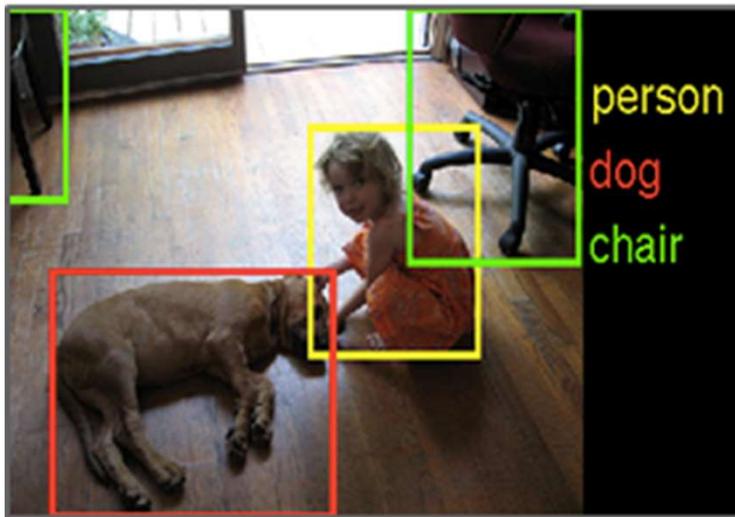


Parler à son téléphone devient possible...

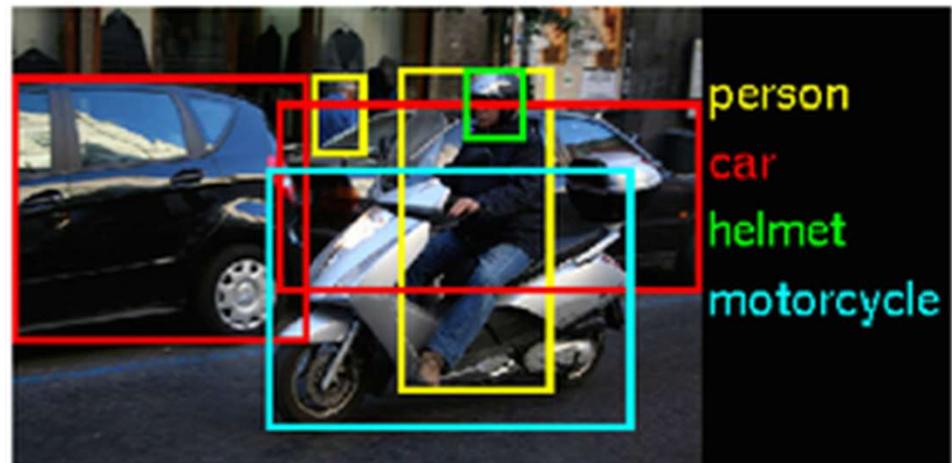
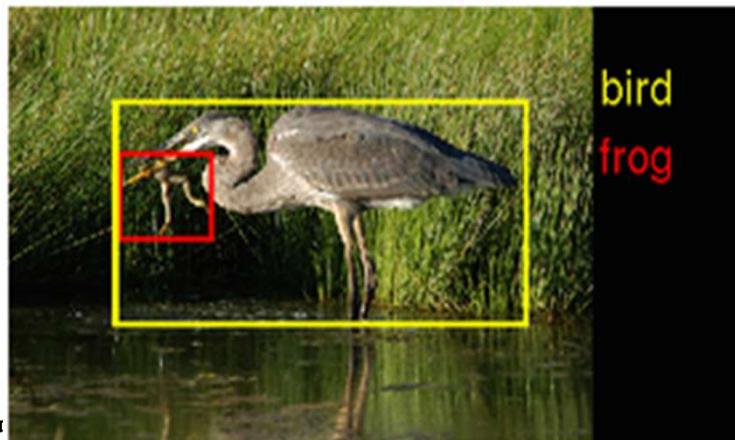
Mars 2016 :  
Champion du monde de Go  
battu par un ordinateur



# Reconnaissance des objets dans les images depuis 2012



- 1000 catégories d'objets, visages
- 2015 : taux d'erreur comparable à l'humain



# Classification d'images médicales

## Détection de cellules cancéreuses par apprentissage profond

	Précision
<b>Imagia</b>	<b>&gt; 90%, real time</b>
<i>GI Experts (Key Opinion Leaders)*</i>	~ 90%
<i>GI Doctors Trained by KOLs*</i>	~ 75%

\*(D. Rex, 2015)

Cadens - Imagia - Satis  
© 2015 - all rights reserved

Confidence:



Probability:

**95%**

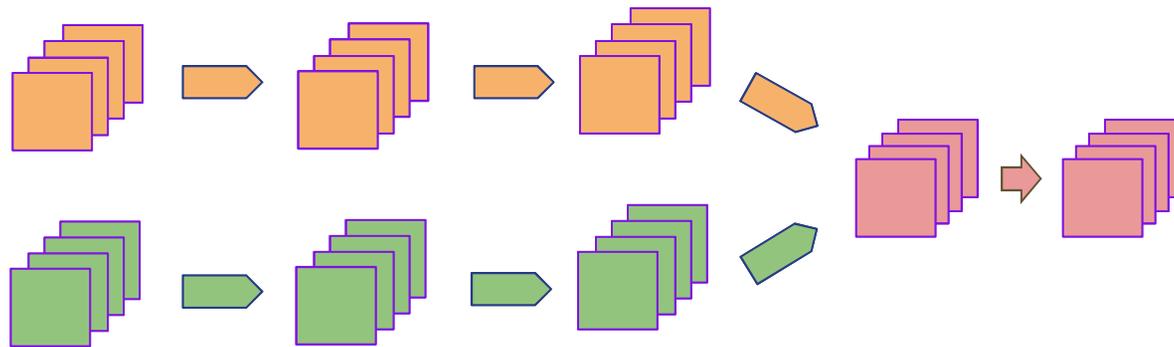
NICE Classification:

**Type 2**



# Deep Data Fusion

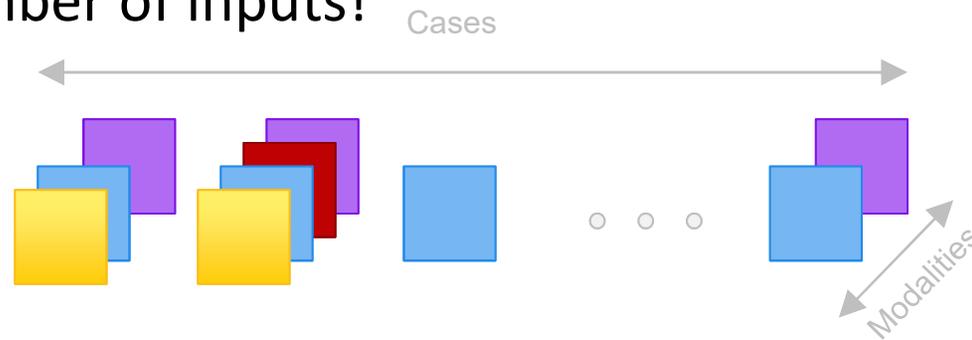
- Deep nets are very good at combining multiple sources of data, multiple sensors or modalities
- With convnets, can have separate pre-processing stages for each modality, then CONCATENATE the representations before continuing processing



Need to map to the same spatial scale, or 'copy' a non-spatial modality at all positions.

# Combining Heterogeneous Sources with Missing Modalities

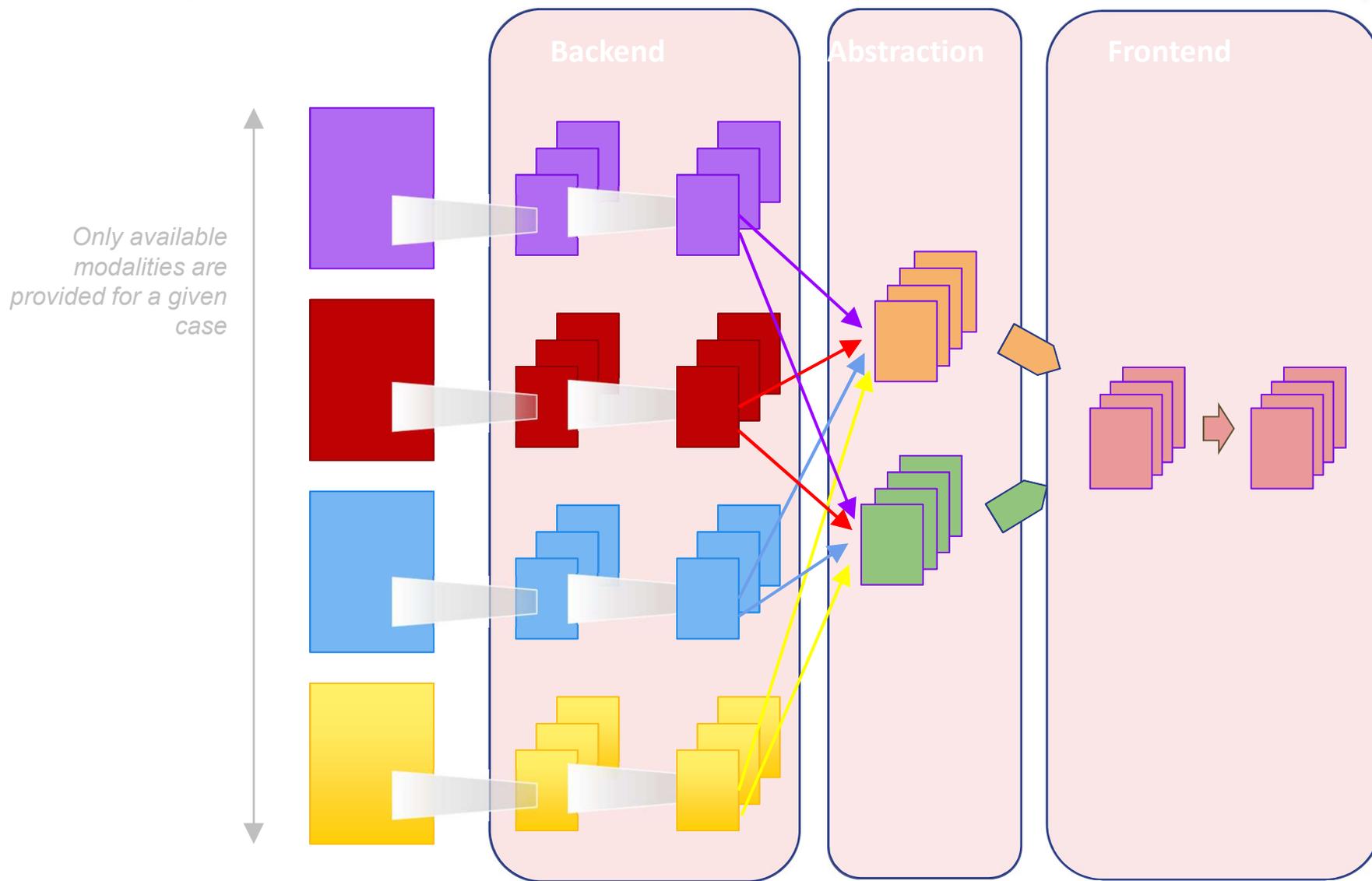
- Different measurements may have been made for different patients, each with a different subset of modalities
- Variable number of inputs!



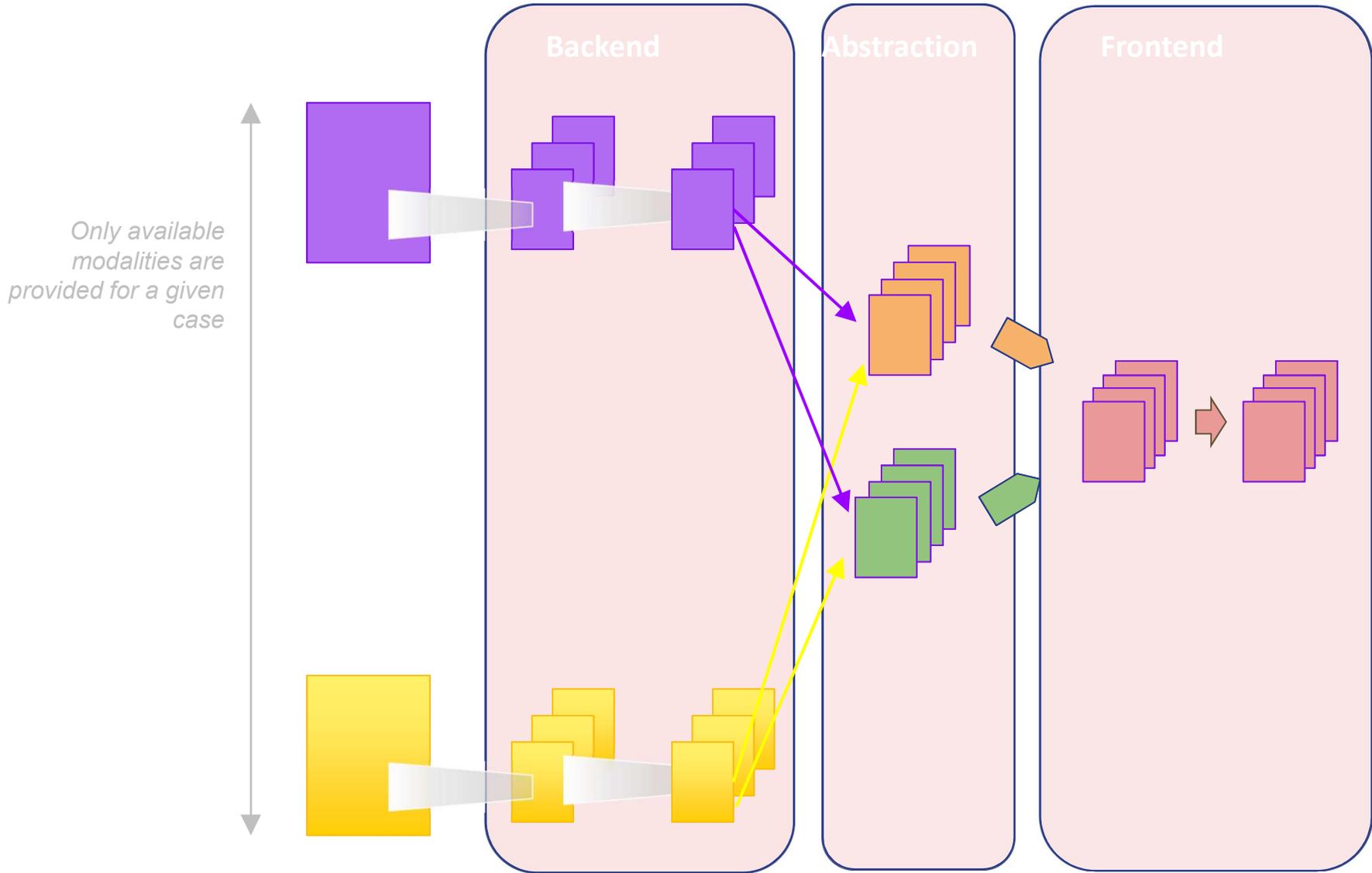
- Handled by HeMIS: Havaei et al MICCAI 2016



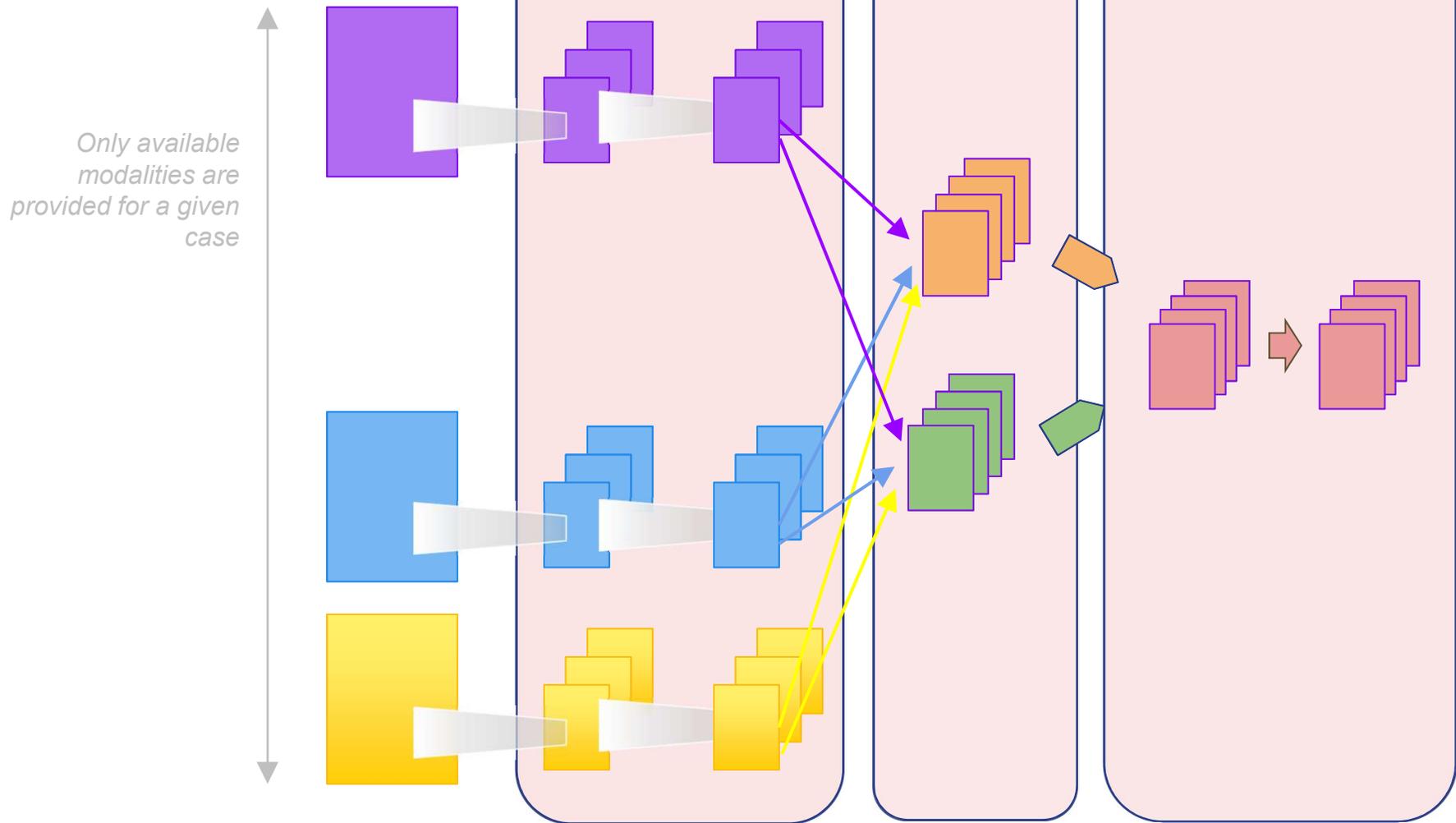
# HeMIS



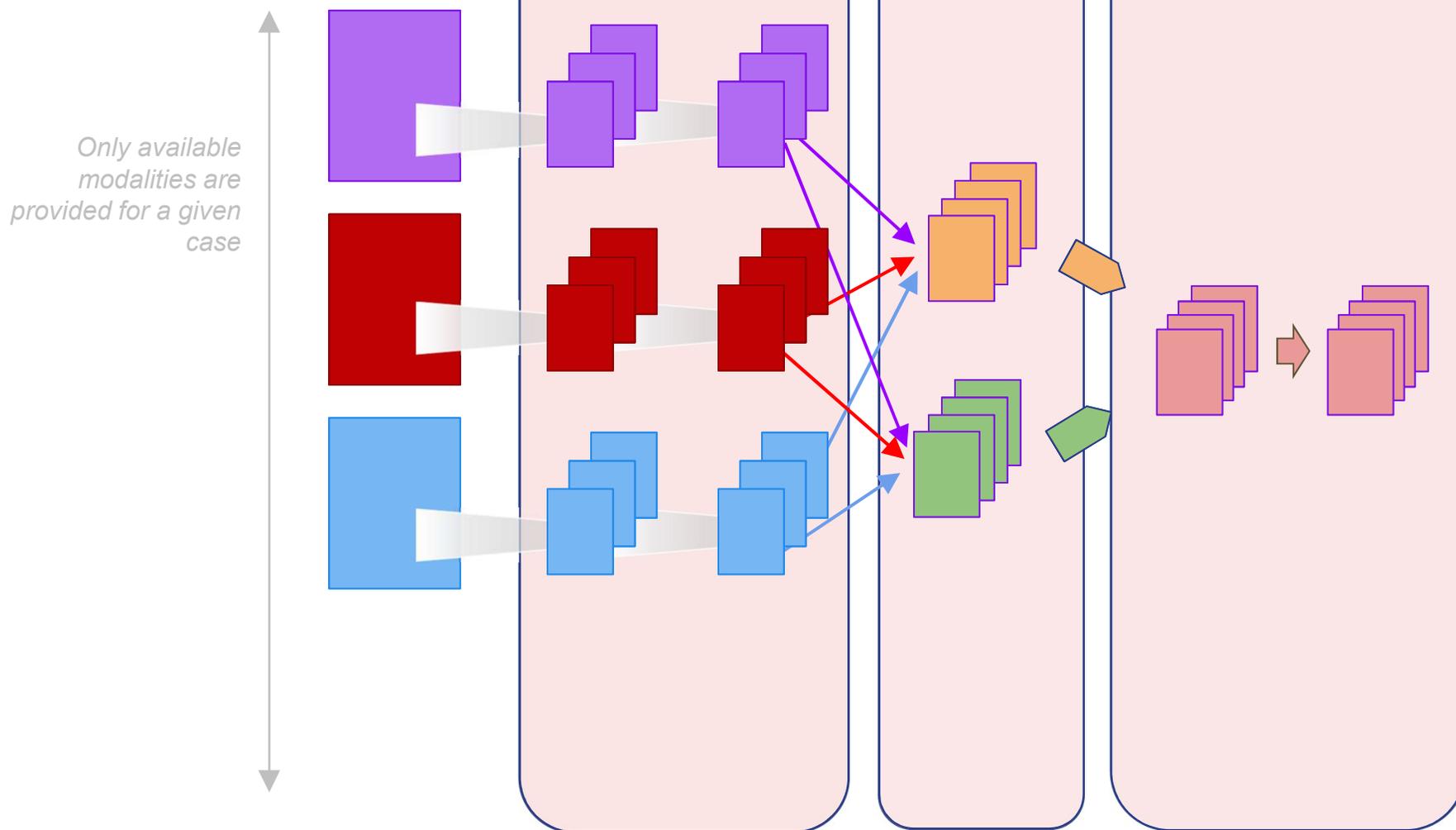
# HeMIS



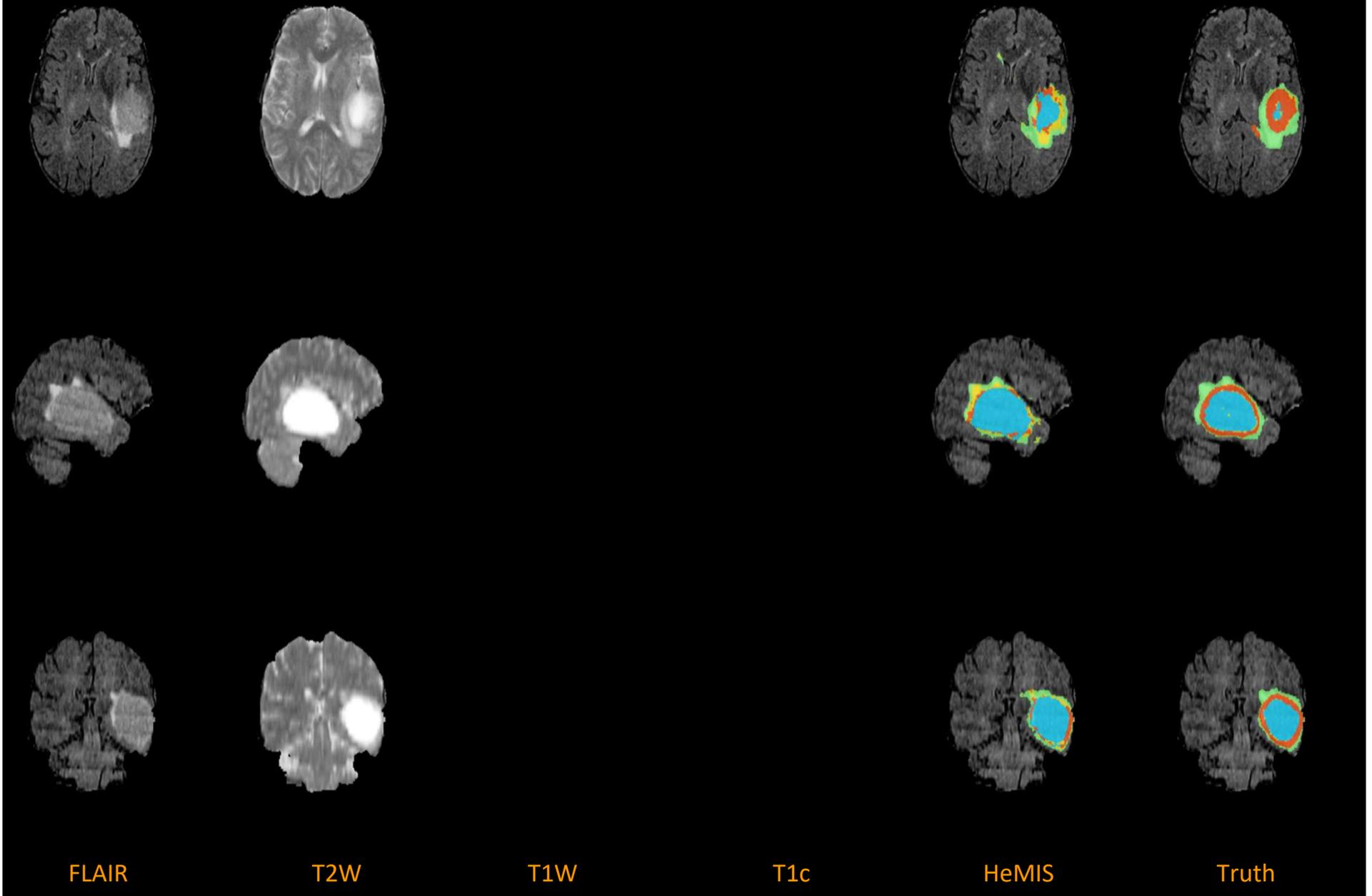
# HeMIS



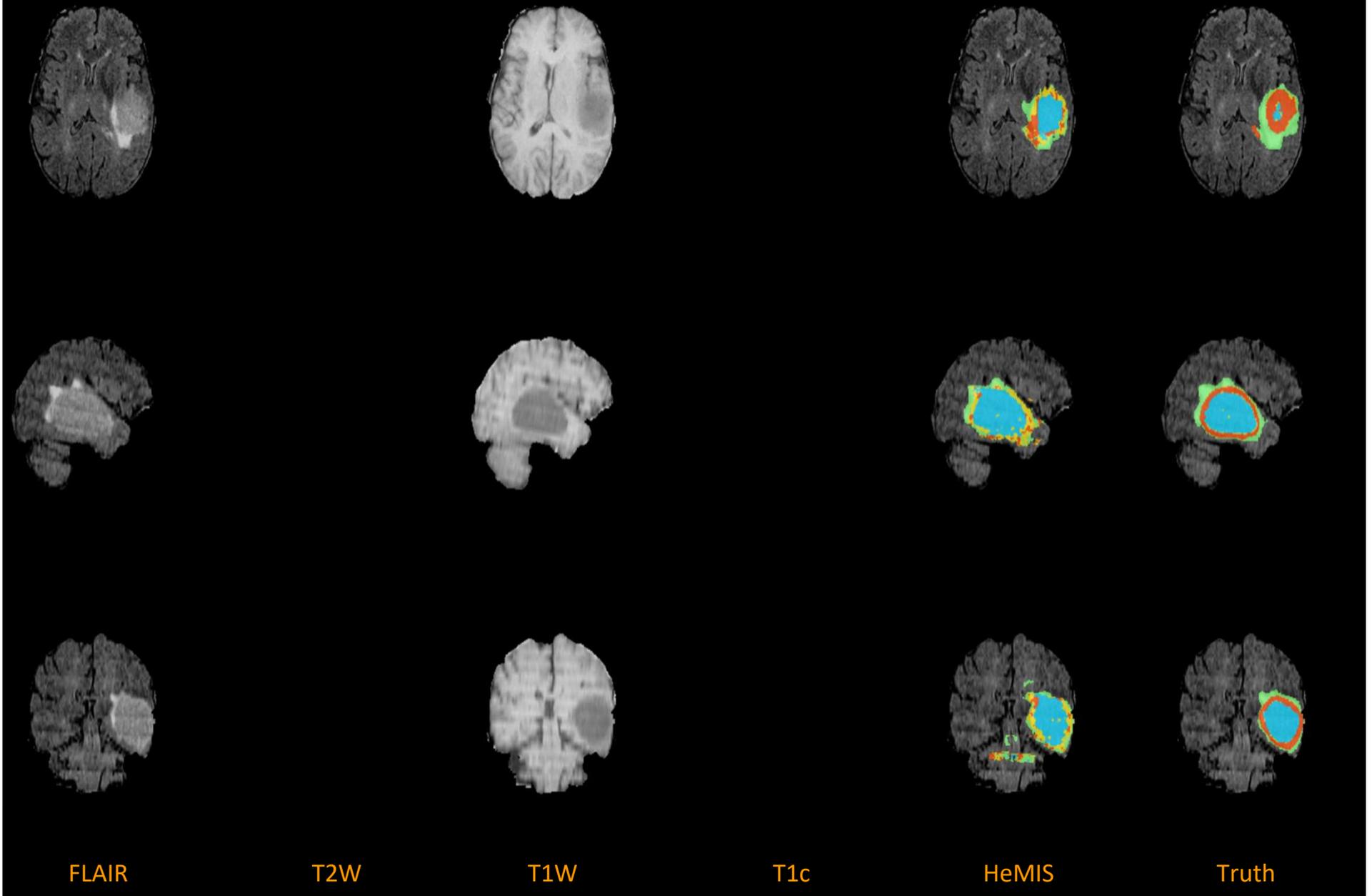
# HeMIS



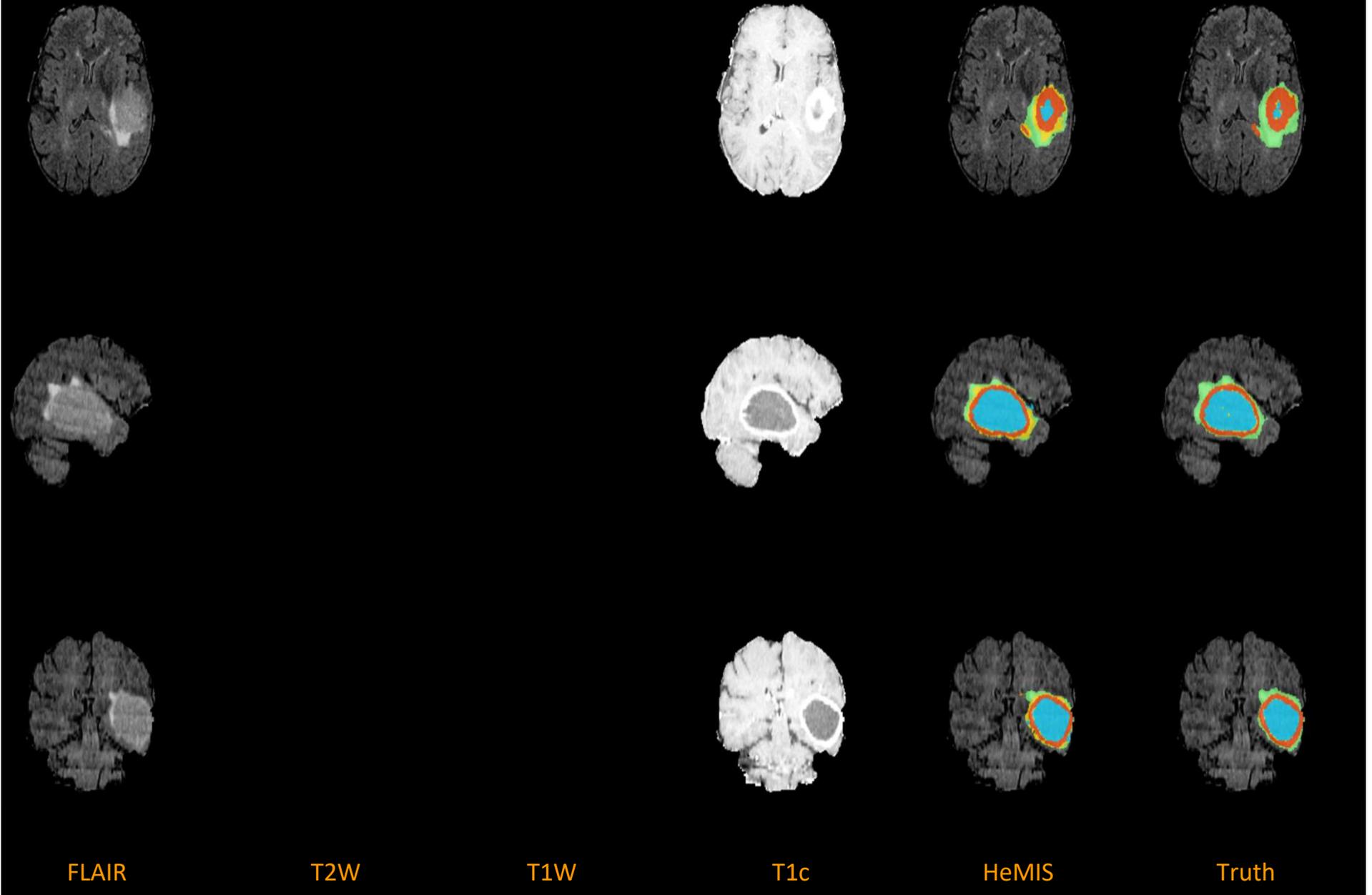
# Results (BRATS)



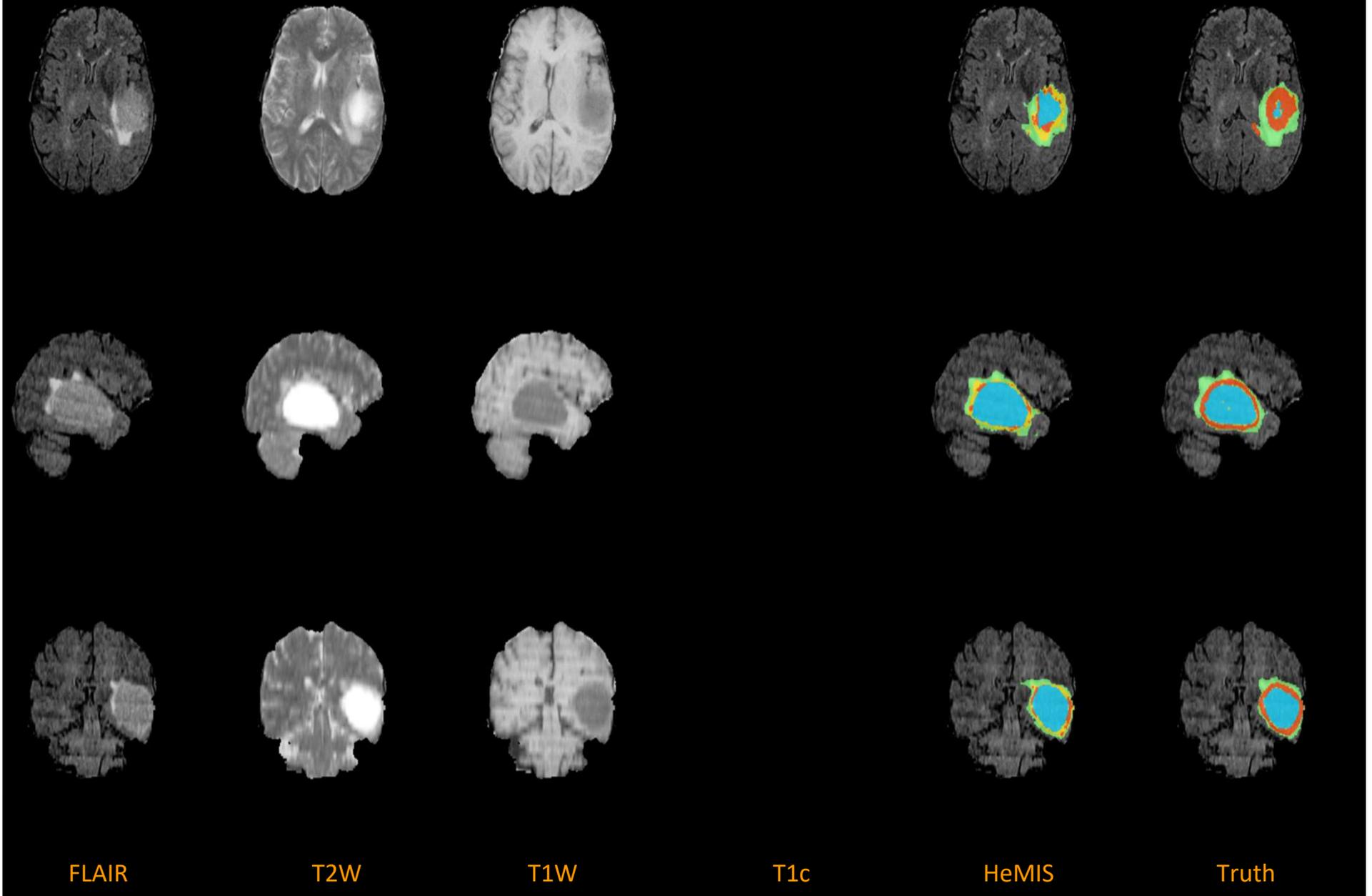
# Results (BRATS)



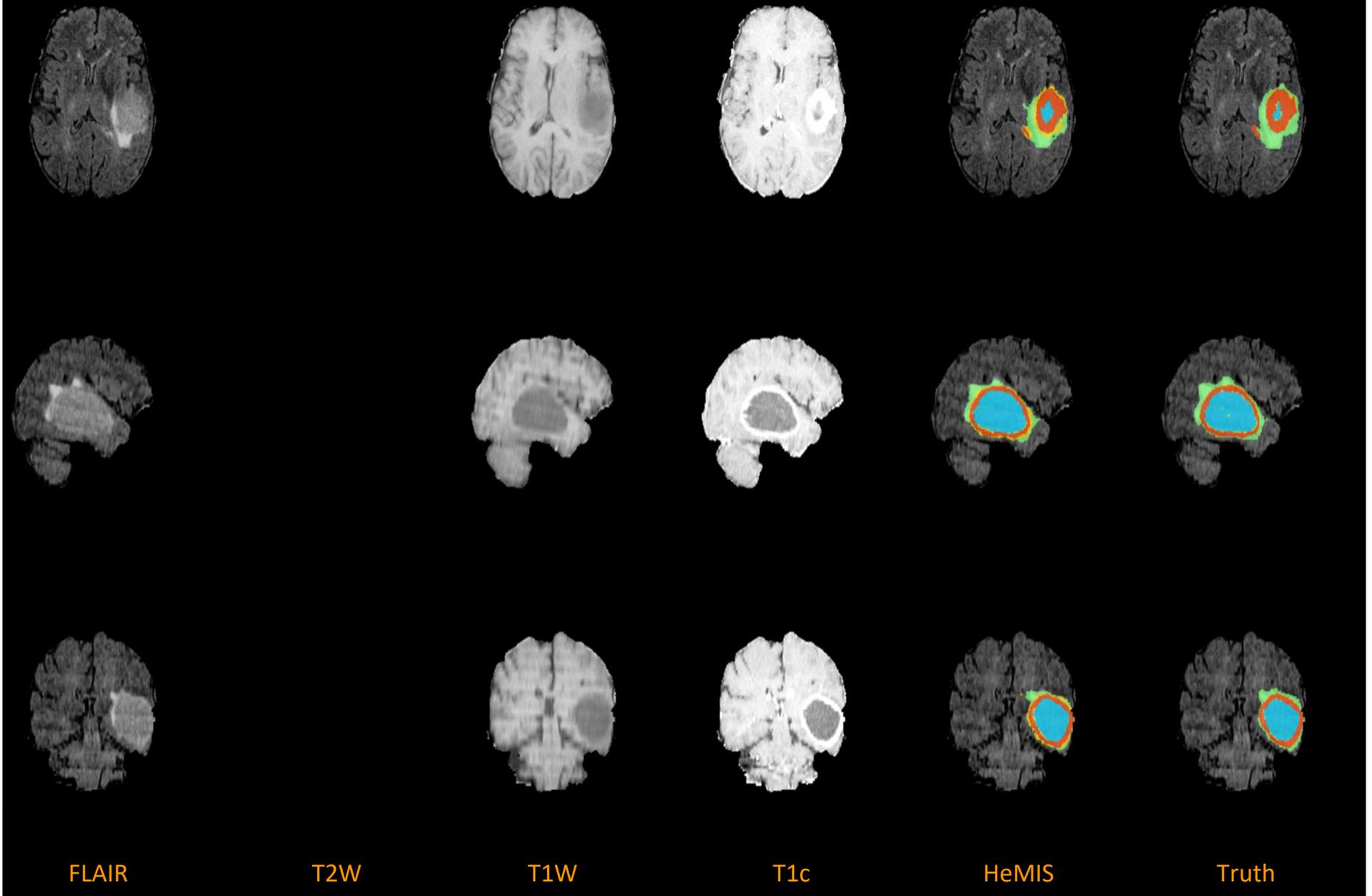
# Results (BRATS)



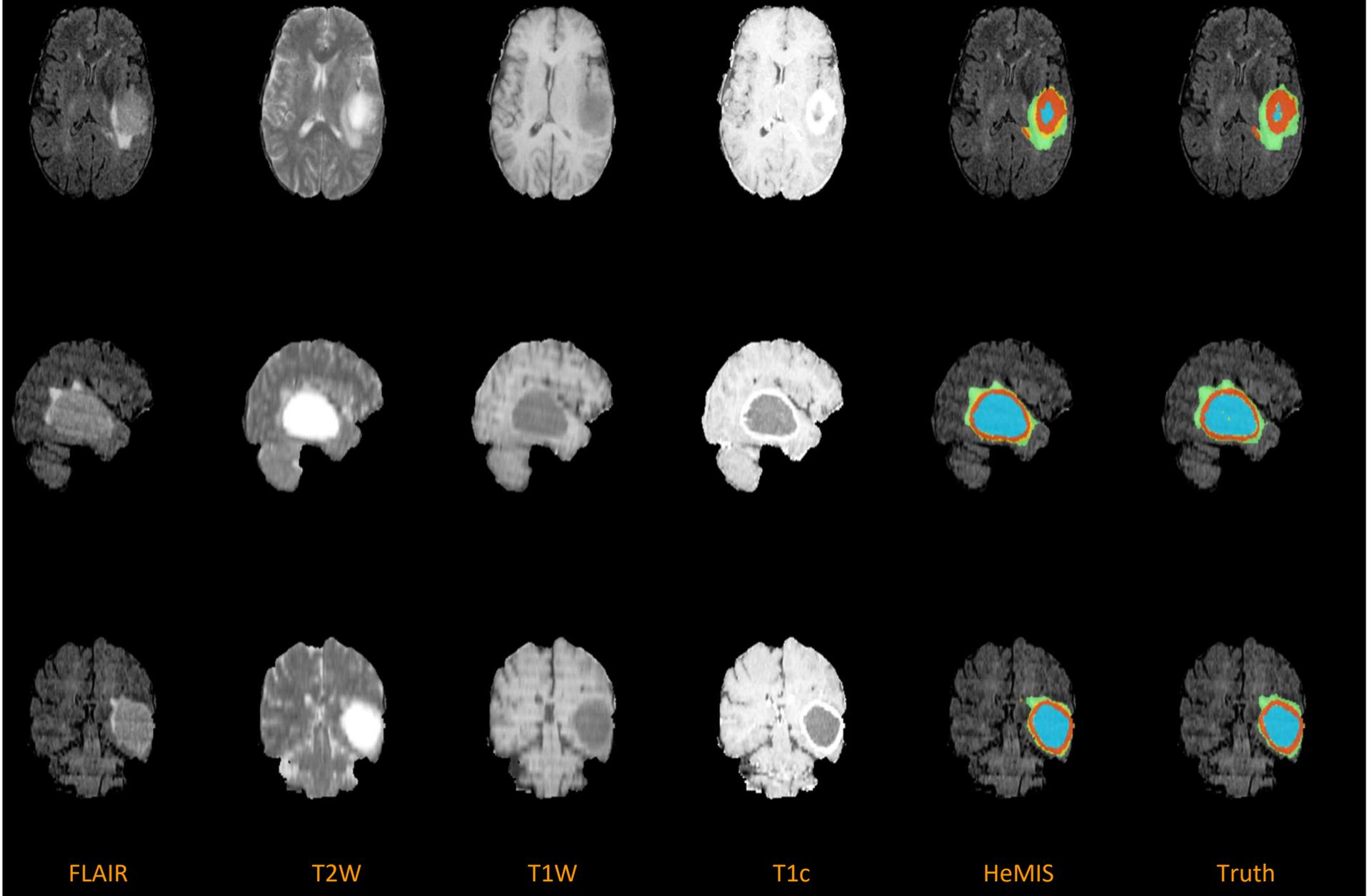
# Results (BRATS)



# Results (BRATS)



# Results (BRATS)



# L'ordinateur peut apprendre à être créatif, inventer de nouvelles images



2014



2015



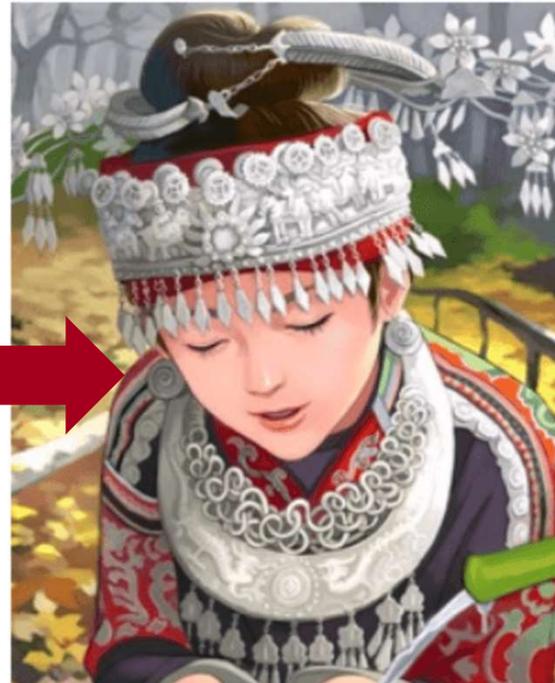
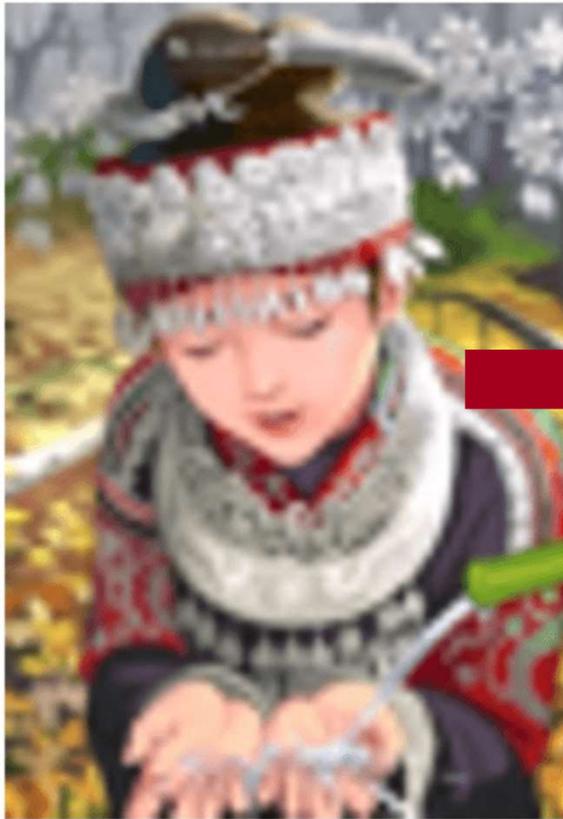
2016



2017



# Inventer dans le contexte



[Ledig et. al., 2016]

**Given text**

This bird has a yellow belly and tarsus, grey back, wings, and brown throat, nape with a black face

This bird is white with some black on its head and wings, and has a long orange beak

This flower has overlapping pink pointed petals surrounding a ring of short yellow filaments

**Generated image**



[Zhang et. al., 2016]

# Qu'est-ce que l'intelligence artificielle?

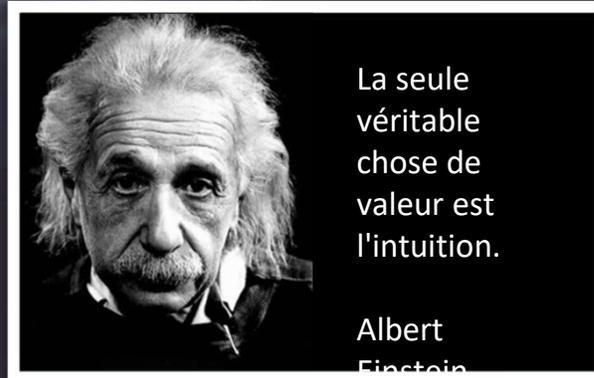
- Intelligence : capacité à comprendre et à prendre de bonnes décisions
- Mes recherches : comprendre les principes mathématiques sous-jacents à l'intelligence, que ce soit la nôtre, celle des animaux, ou celle qu'on est en train de construire
- La clé de l'intelligence : la connaissance



## CONNAISSANCES INTUITIVES & APPRENTISSAGE AUTOMATIQUE

- Les premières décennies de recherches en IA dominées par l'idée de donner directement nos connaissances à l'ordinateur, sous forme de règles, faits, procédures
- Un échec à cause de toutes les connaissances intuitives que nous avons mais qui ne sont pas accessibles à notre conscience, donc incommunicables

La solution :  
l'apprentissage automatique

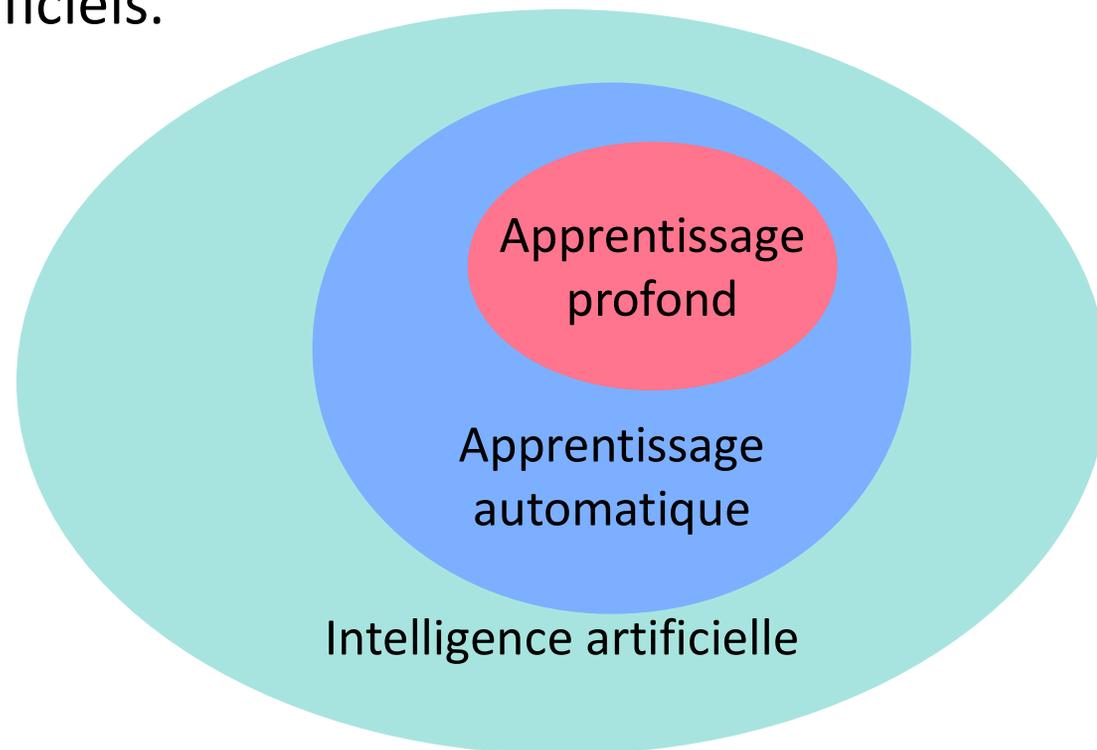


La seule véritable chose de valeur est l'intuition.

Albert Einstein

# Apprentissage profond, apprentissage automatique, intelligence artificielle

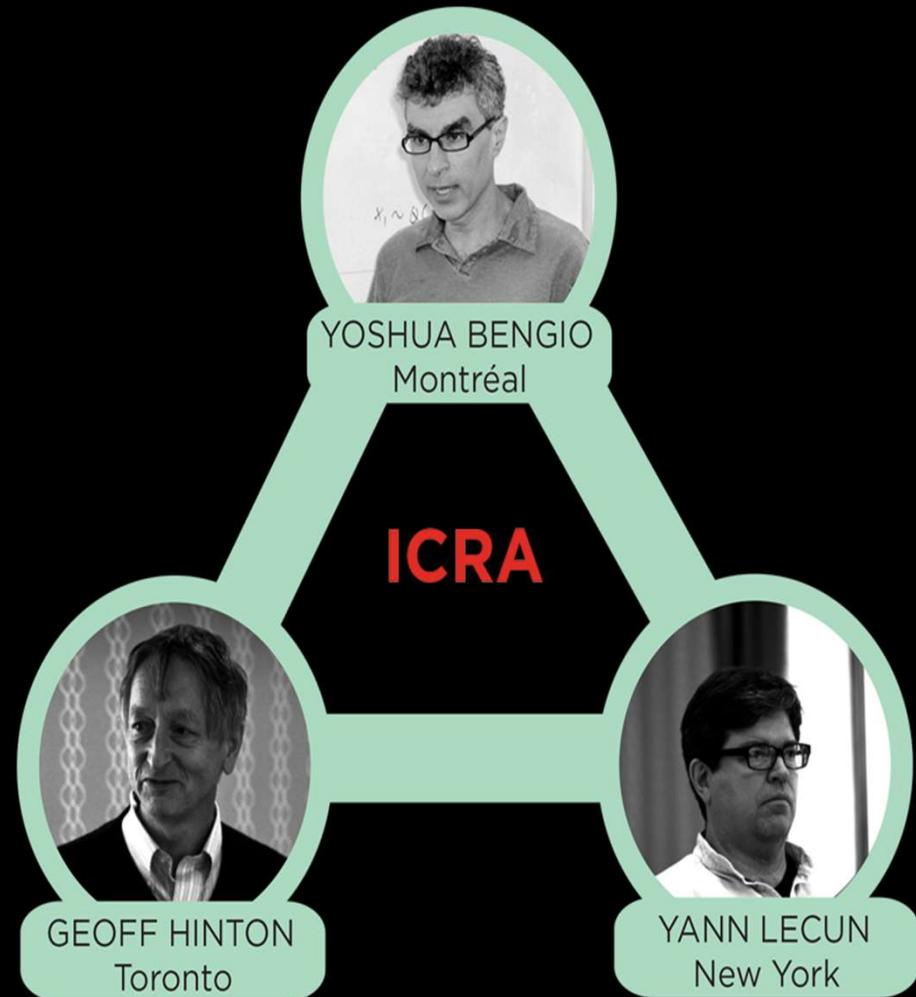
L'apprentissage profond fait suite à des décennies de recherches sur les réseaux de neurones, inspirés du cerveau, et y ajoute entre autre l'importance d'avoir plusieurs niveaux ou plusieurs couches de neurones artificiels.



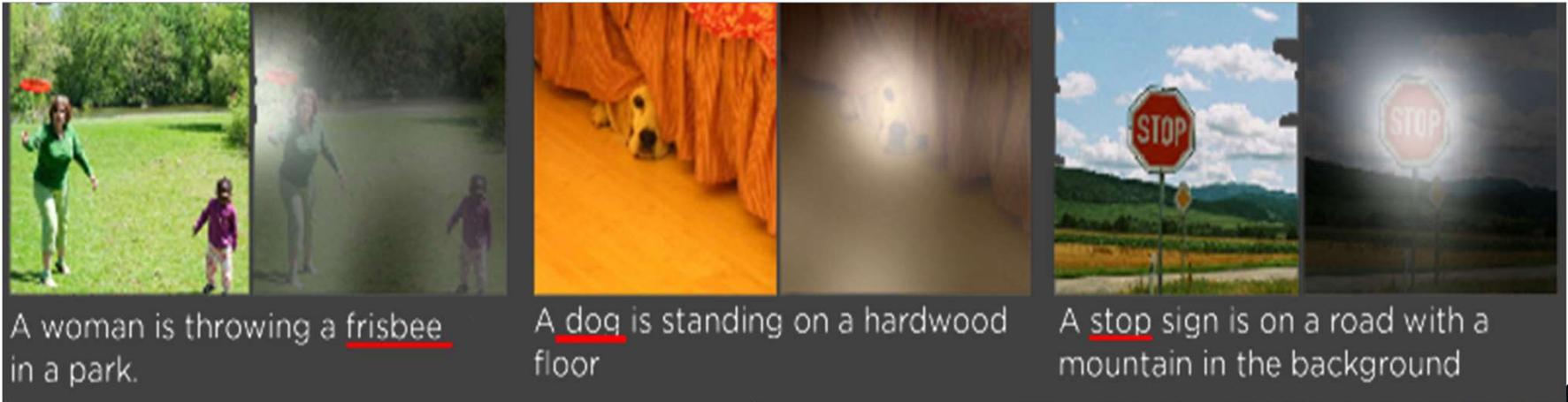
# Percée en apprentissage profond

Un trio de l'ICRA,  
au 2/3 canadien,  
a concrétisé  
l'apprentissage  
profond à partir  
de 2006

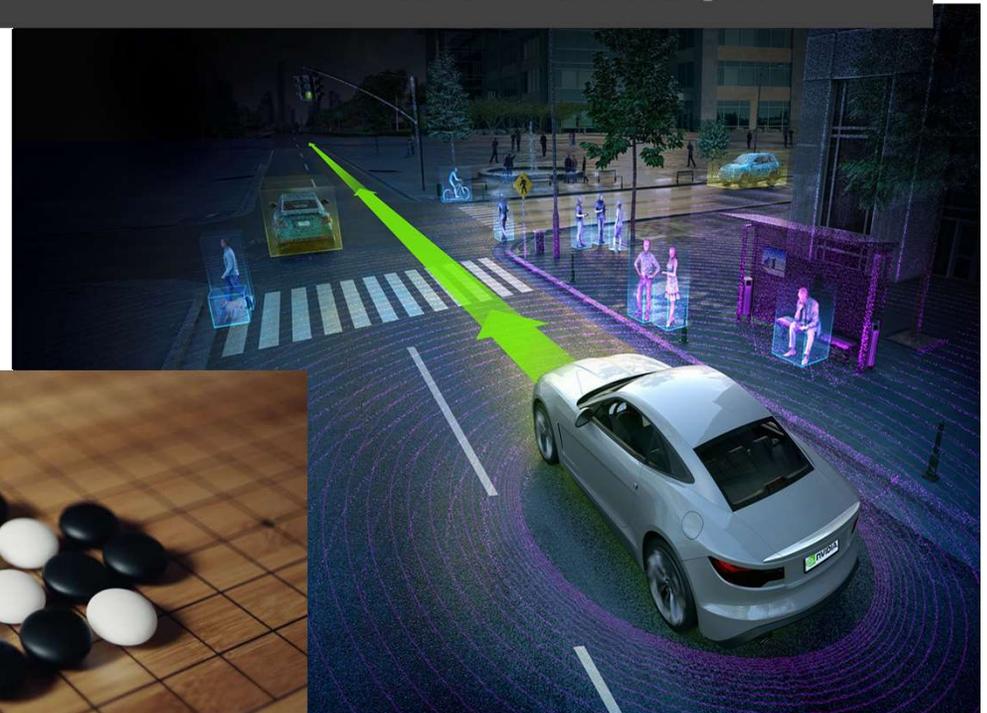
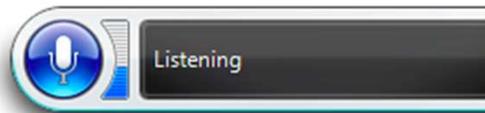
Grâce au  
financement  
gouvernemental  
de la recherche  
exploratoire



# Percées en IA grâce à l'apprentissage profond



Avancées majeures dans la capacité de l'ordinateur à percevoir, planifier, utiliser et comprendre le langage, raisonner...



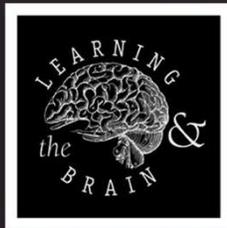
# IDÉE MAÎTRESSE: L'HYPOTHÈSE DE PRINCIPES SIMPLES DE L'INTELLIGENCE

## Hypothèse 1:

il existe un petit nombre de principes relativement simples, comme les lois de la physique, mais expliquant notre intelligence et nous permettant aussi de fabriquer des machines intelligentes

## Hypothèse 2:

**ces principes sont ancrés dans l'apprentissage.**



*Mon but personnel dès 1985:*

*Contribuer à la découverte et la compréhension de ces principes*

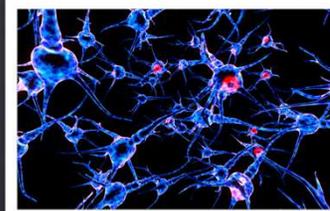
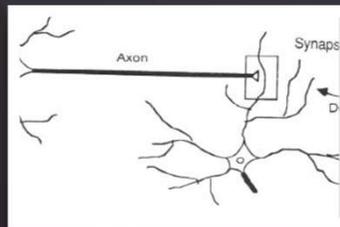
# LA PLUS BELLE MACHINE À APPRENDRE: LE CERVEAU

Exemple vivant de machine à apprendre qui fonctionne bien: le cerveau

**Depuis les années 40, le cerveau est une inspiration pour les chercheurs en intelligence artificielle**

Calcul: neurones

Apprentissage: synapses



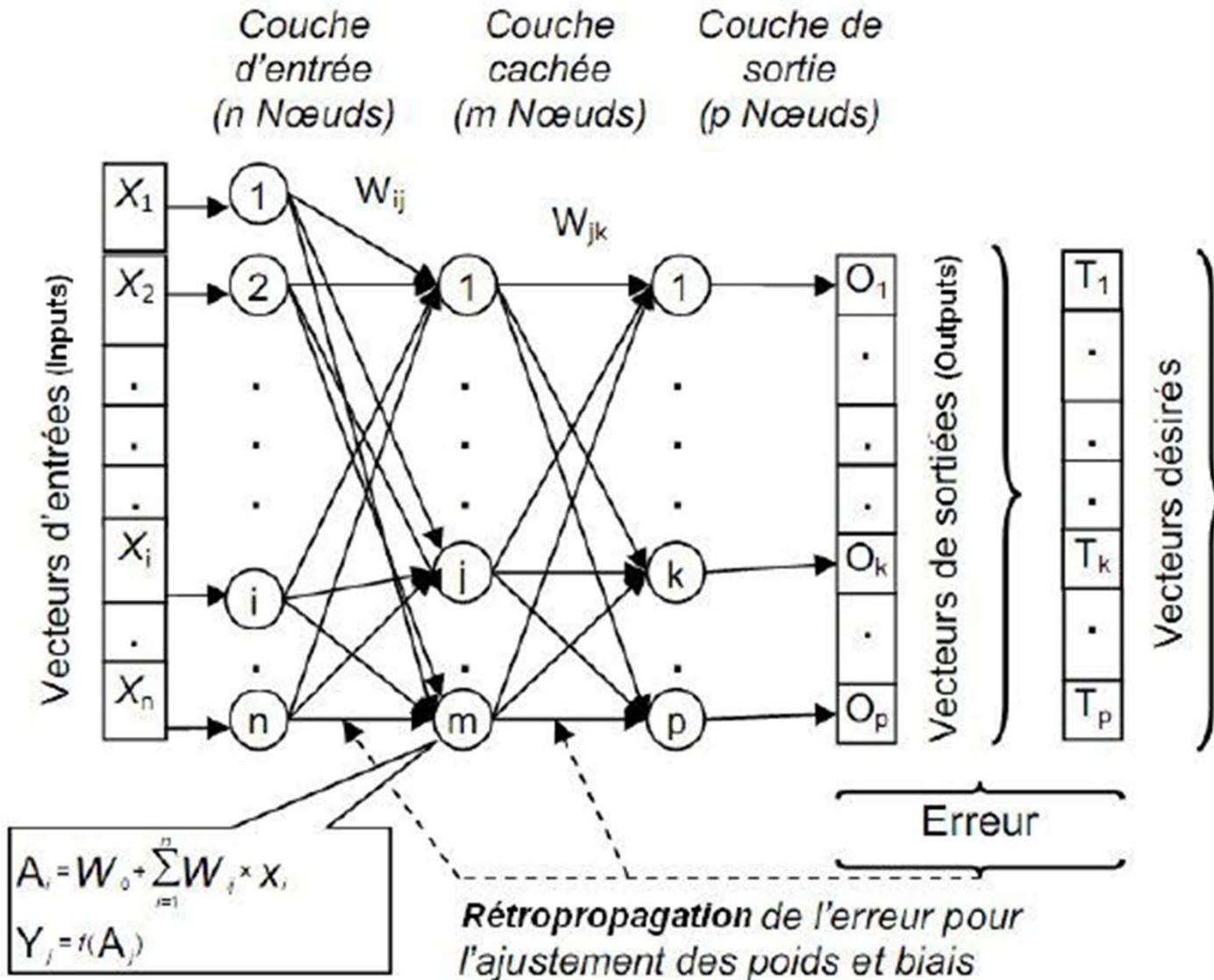
## INSPIRATION BIOLOGIQUE POUR L'IA

Pas suffisant d'imiter les détails,  
Il faut comprendre les principes.

L'échec de l'Avion III de Clément Ader

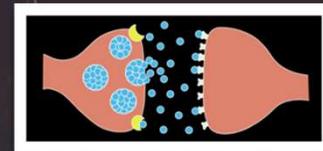
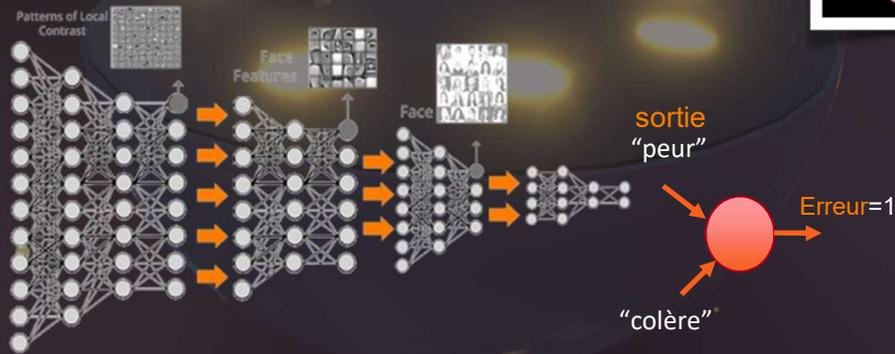


# Réseau de neurones artificiels



# APPRENTISSAGE PAR L'EXEMPLE

- On présente une série d'exemples
- Après chaque exemple on fait une légère correction pour obtenir une réponse un peu meilleure la prochaine fois
  - Petite modification des synapses (force des connexions)



# APPRENTISSAGE SUPERVISÉ : L'ORDINATEUR APPREND À IMITER L'HUMAIN

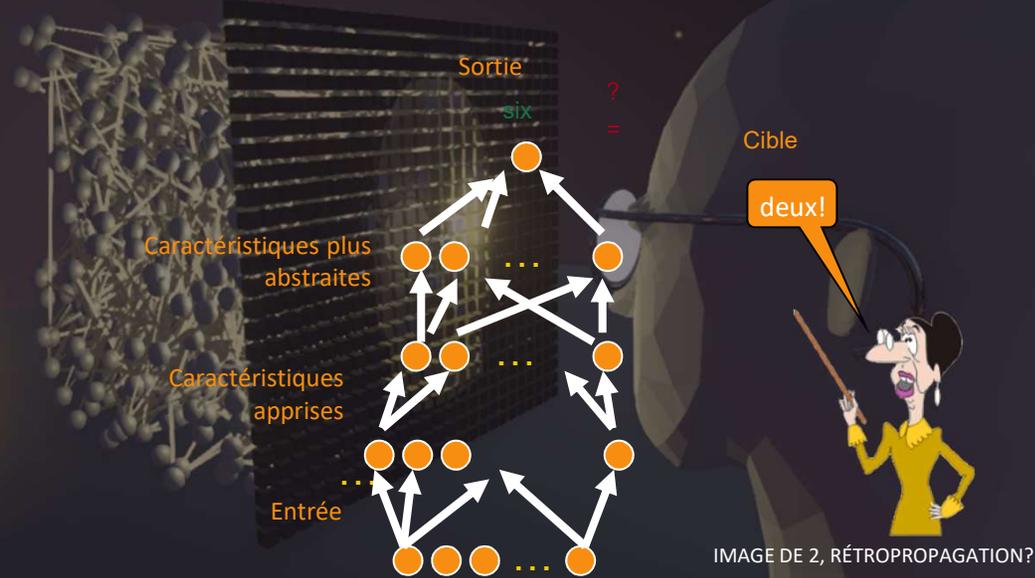


IMAGE DE 2, RÉTROPROPAGATION?

*Nécessite un grand nombre de paires (entrée, cible) comme exemples pour entraîner l'ordinateur à acquérir un comportement*

# COMMENT GÉNÉRALISER CORRECTEMENT À DE NOUVEAUX CAS ?

Ensemble de données  
d'entraînement

The diagram shows a training dataset with 10 rows of handwritten digits. The first five rows contain the digit '2' and are highlighted with a blue border. The last five rows contain the digit '3'. An arrow labeled 'Entrée' points to the first '2' in the third row, and an arrow labeled 'Étiquette' points to the second '2' in the same row. Below the training set, a test case is shown: a white box containing the digit '3' with an arrow pointing to the text '2 ou 3?'.

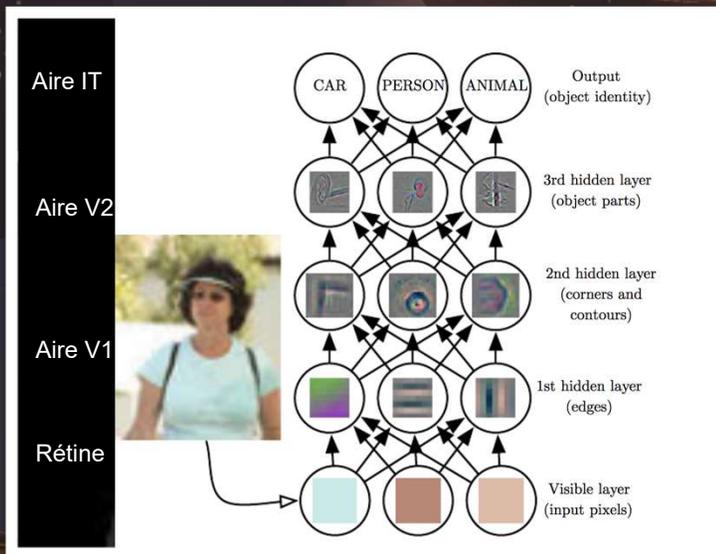
2	2
2	2
2	2
2	2
2	2
3	3
3	3
3	3
3	3
3	3

image de test:  ⇒ 2 ou 3?

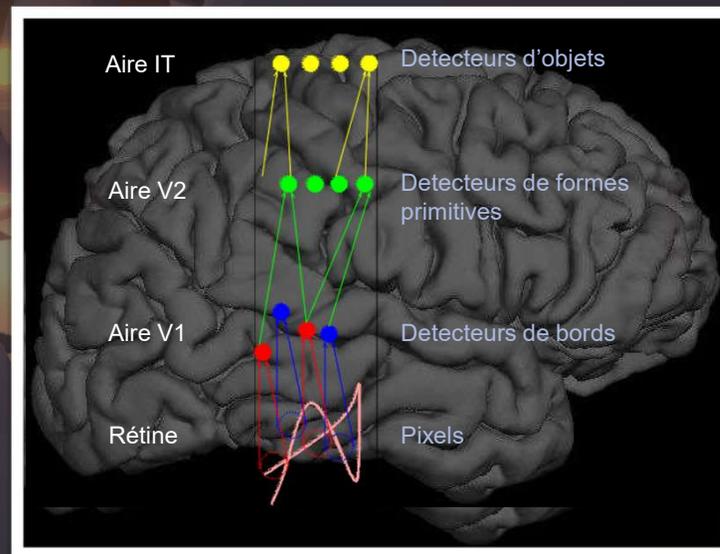
*Apprendre n'est pas  
simplement mémoriser...  
(ce qui est facile, il suffit d'un  
fichier)*

*C'est être capable de  
généraliser à de nouveaux  
cas!*

PLUSIEURS NIVEAUX  
D'ATTRIBUTS RECOMBINABLES  
PERMETTENT DE GÉNÉRALISER  
À DE NOUVELLES CONFIGURATIONS



RÉSEAU DE NEURONE  
PROFOND DANS LE CERVEAU:  
CORTEX VISUEL



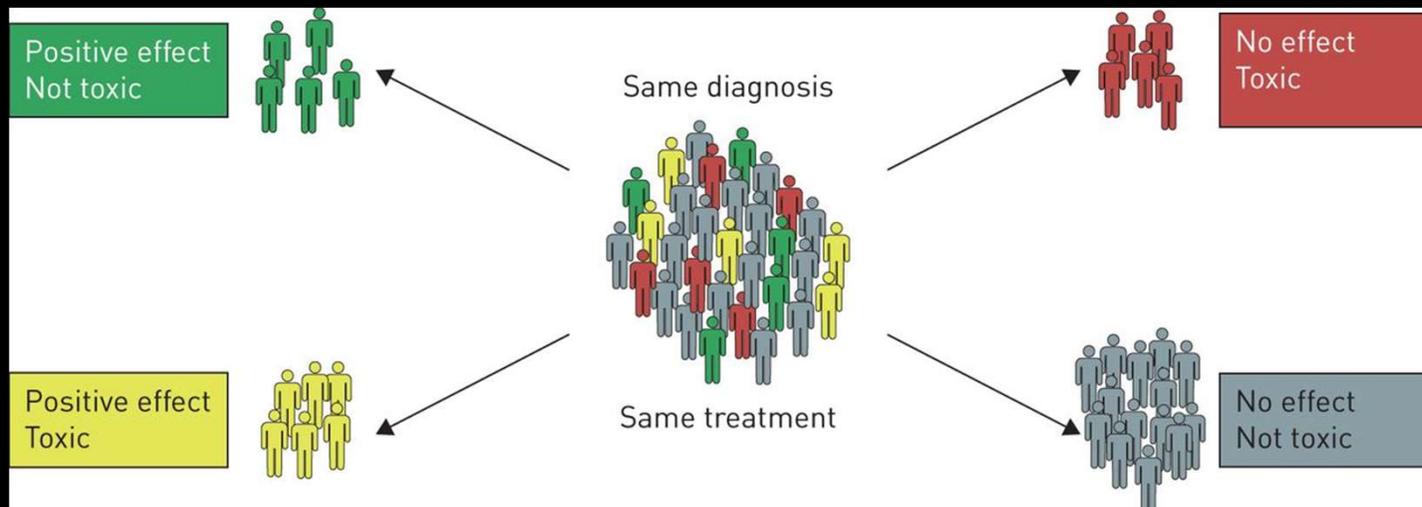


# data is the new oil

- L'apprentissage automatique requiert beaucoup de données
- La 1ere question dans tout projet:
  - Quelles données sont disponibles et pourra-t-on recueillir? Sont-elles représentatives et indiquent-elles à l'ordinateur les bonnes réponses?

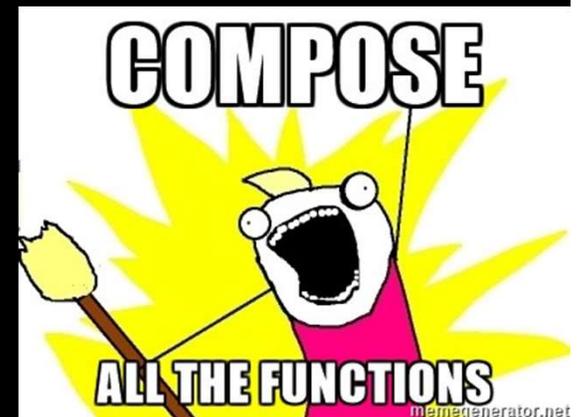
# Deep learning for precision medicine

- Control surgical robots
- Handle high-dimensional inputs which doctors do not have time to look at
  - Genome, mRNA expression levels
  - 3-dim images (e.g. brain scans) or videos (e.g. intestine)
- Handle patient history data
- Handle textual data (e.g. doctors' reports)
- Handle missing data and censored data
- Predict future outcomes or imitate doctors' decisions
- Estimate probabilities of future events
- Predict properties of molecules, for drug discovery
- Generatively design molecules with desired properties

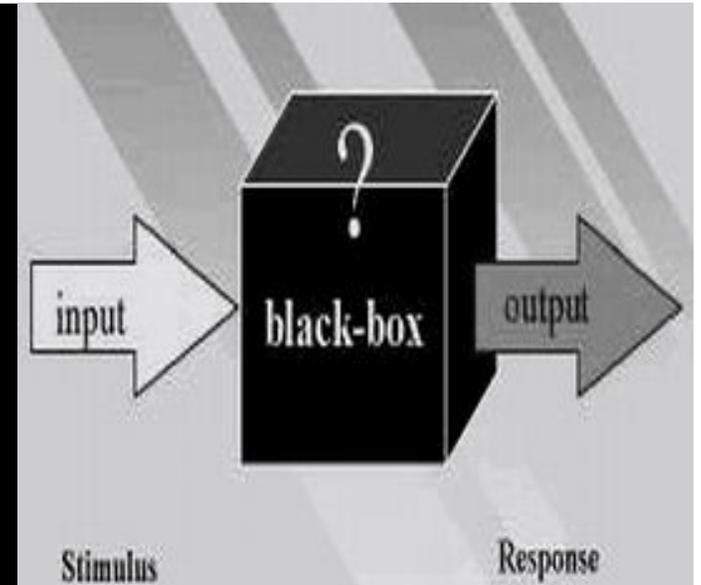


# Exponential Statistical Advantage

- Classical statistical studies (ANOVA) with multiple factors: need number of datapoints exponential in # factors
  - Curse of dimensionality
- Deep learning makes combinatorial assumptions of underlying features and hierarchy of features → theory-proven exponential advantage



# The So-Called Black Box



- Problem with complex machine learning: no simple interpretation of the learned predictor's mechanism
- But we can identify leading factors for each particular case
- Would you rather use a very precise diagnostic which saves your life but is not fully explained in simple terms or one that is very simple but makes wrong predictions and lets you die?
- Doctors are not always able to completely justify their decisions either, because their brain is performing a complex computation

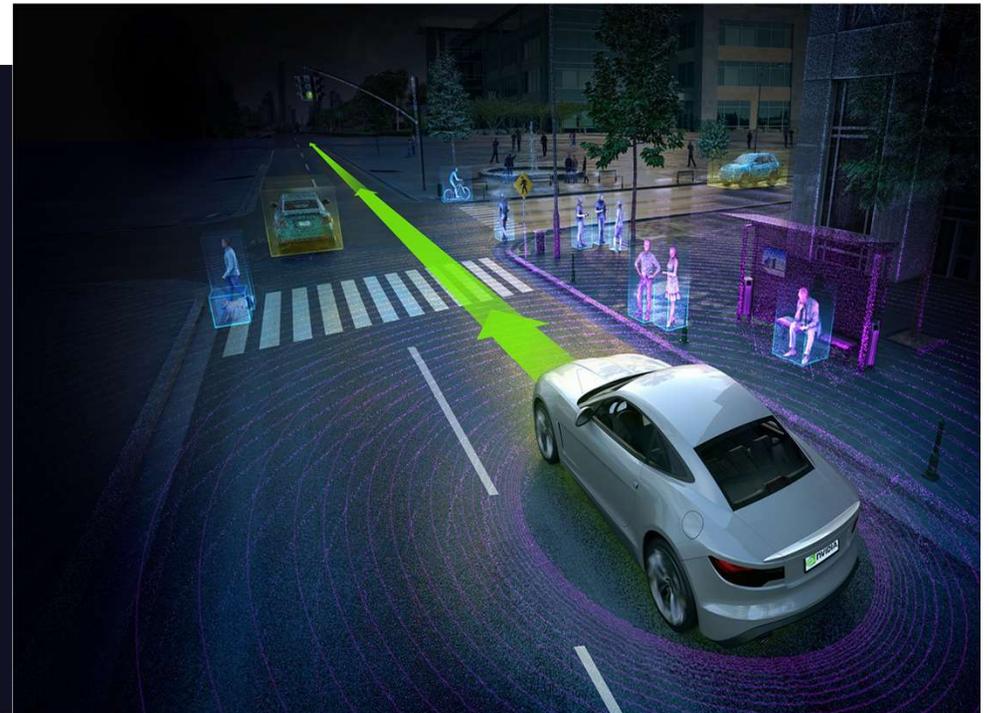


# From Controlled Groups to Population Data and Randomized Interventions

- The traditional control-group setting does not have enough power to allow learning complex dependencies across many variables
- How to avoid confounders?
  - This is an active research area
  - Randomized interventions in principle solve the problem, or Regression Discontinuity Design (RDD)
  - Simulated randomized intervention might be possible if we model the causal structure

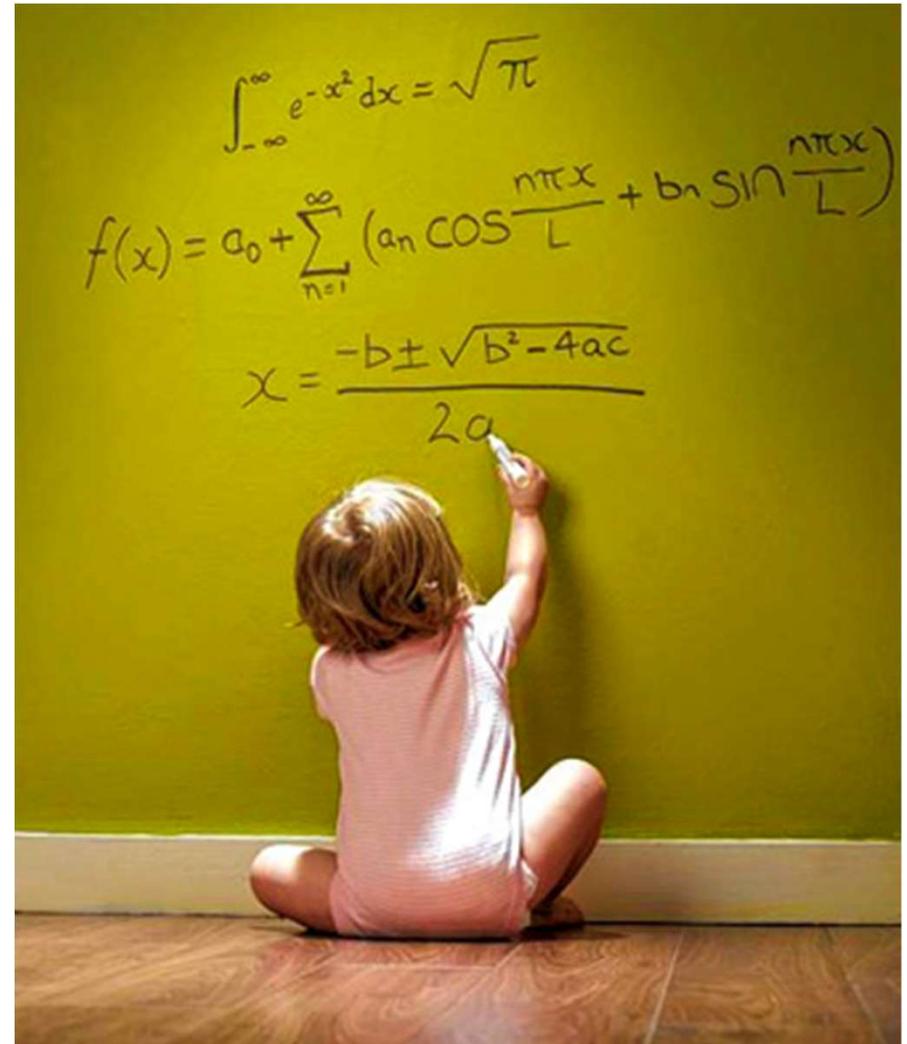
# On est loin de l'IA de niveau humain

- Beaucoup de progrès dans les dernières années
- Grâce à l'apprentissage profond
- Surtout l'apprentissage supervisé
- Surtout pour la perception



# Les êtres humains et les animaux dépassent encore de loin l'ordinateur pour l'apprentissage non supervisé

- Un enfant de 2 ans a découvert la physique intuitive sans que ses parents la lui expliquent
- Les enfants apprennent par l'observation et en interagissant avec leur environnement



# QU'EST-CE QUI NOUS MANQUE?

- Système 1 (intuitif, rapide, inconscient)
- vs
- Système 2 (linguistique, séquentiel, lent, conscient)

- Perception (représentation, classification)
- vs
- Action (raisonnement, planification, causalité)

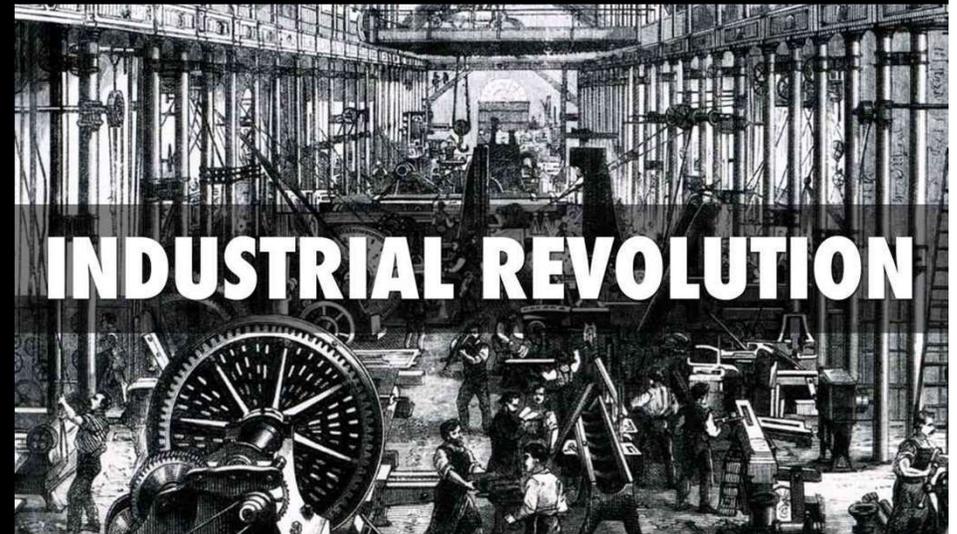


*Outil de recherche privilégié: environnements virtuels (jeux vidéo), où l'apprenant peut explorer pour mieux construire sa représentation du monde (images Atari de DeepMind)*

# IA: la prochaine révolution industrielle

## Première révolution industrielle

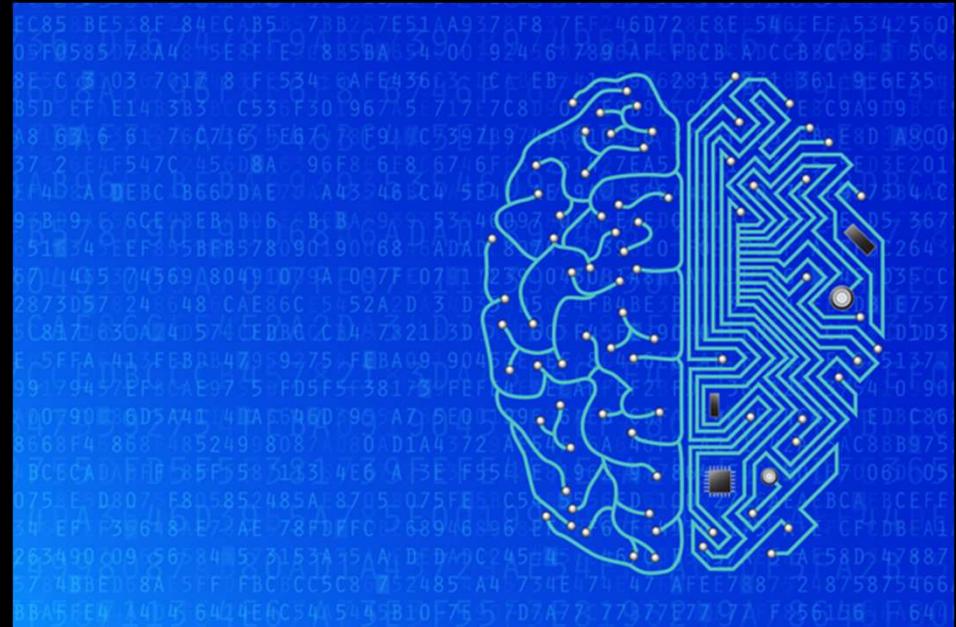
- Les machines augmentent la **puissance mécanique** des humains



# IA: la prochaine révolution industrielle

**Prochaine révolution industrielle:**  
Les machines augmentent la **puissance cognitive** des humains

Croissance rapide du secteur de l'IA



## Tous secteurs de l'économie:

Augmentation du PIB total de 14% d'ici 2030 = 16 trillions \$US

# Impact sur les emplois?

- Estimations de 30 % des emplois touchés en 10 ans
- Le plus: routiniers ou spécialisés, transports, ?
- Le moins: informatique et tâches nécessitant un large bagage général ou un aspect relationnel humain important



# Coût social de la transition

- Faire face à la transition
  - Plus de formation
  - Meilleur filet social, revenu minimum garanti?

**Revenu minimum  
garanti : trois études  
de cas**



**IRIS**

Institut de recherche  
et d'informations  
socio-économiques

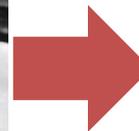
Eve-Lyne Couturier,  
chercheure

progrès  
technologique  
&  
progrès  
humain

- **Applications médicales** (actuellement surtout imagerie médicale)
  - cancer, médecine personnalisée, découverte de médicaments
- **Automatisation et robotique**
  - réduction significative des coûts de production, industrie et services
- **Véhicules autonomes**, mobilité accrue
- **Compréhension du langage**: assistants personnels, interfaces plus accessibles avec l'ordinateur, traduction automatique, démocratisation des services légaux ...
- **Recherche d'information** plus intelligente et soutien à l'éducation personnalisée.

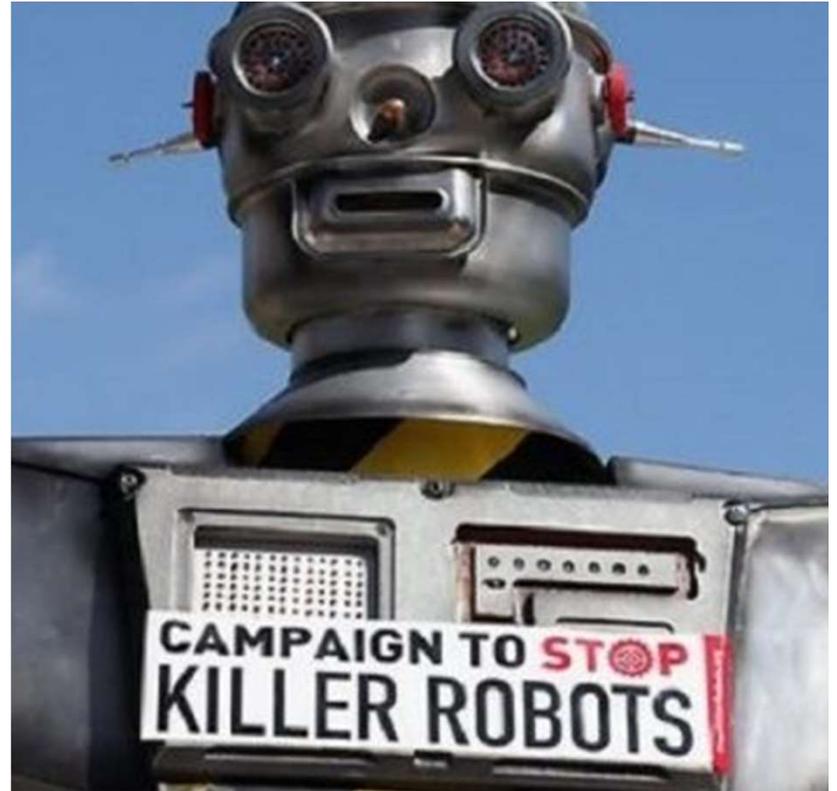
# La fin du travail comme esclavage

- ***Si on fait les bons choix collectifs***
- Les machines vont faire les tâches que les humains ne veulent pas faire (dangereuses, pénibles, répétitives)
- On pourrait éliminer le travail comme esclavage, chacun étant libre de s'épanouir en choisissant comment contribuer de son mieux à la collectivité



# IA: les inquiétudes

- Big Brother, contrôle politique
- Robots tueurs
- Manipulation par la publicité ciblée
- Vie privée et droits individuels menacés
- Misère pour ceux qui vont perdre leur emploi
- Renforcement des discriminations et biais sociaux
- Inégalités accrues et concentration du pouvoir



# Aspects sociaux et éthiques de l'IA: Agir de façon responsable

- Canada/Qc deviennent chefs de file en IA
- Différentiateur de notre marque:  
Nous devons agir de façon responsable en IA
  - Parce que c'est la chose juste à faire
  - Parce que ça attire et retient le talent



# La Déclaration de Montréal pour le Développement Responsable de l'IA

7 principes

**BIEN-ÊTRE**

**AUTONOMIE**

**JUSTICE**

**VIE PRIVÉE**

**CONNAISSANCE**

**DÉMOCRATIE**

**RESPONSABILITÉ**

# La Déclaration de Montréal pour le Développement Responsable de l'IA

## Bien-être

Le développement de l'IA devrait ultimement viser le bien-être de tous les êtres sentients.

## Justice

Le développement de l'IA devrait promouvoir la justice et viser à éliminer les discriminations, notamment celles liées au genre, à l'âge, aux capacités mentales et physiques, à l'orientation sexuelle, aux origines ethniques et sociales et aux croyances religieuses.

## Connaissance

Le développement de l'IA devrait promouvoir la pensée critique et nous prémunir contre la propagande et la manipulation.