Summary Report for INAHOR activity
AFSIS SAFER Project / JAXA ALOS-2 satellite application / APRSAF SAFE Project Cooperative Project
Promote rice planted area and production estimation using space-based technologies in Vietnam

---

1. Introduction

In Vietnam, Agricultural statistics has almost depended on qualitative data generated by "personal statisticians' opinions". These qualitative data have been collected by uncertainty methods, sent and stored mainly by Microsoft Word and Excel installed in private statisticians' desktop computers. Due to the lack of standard "quantitative method", these data can only be used for description statistical reports in separated time points. These data presented in many descriptions in the statistical reports that cannot be linked together to analyze and forecast.

In order to solve this problem, the agricultural statistics system of the Ministry of Agriculture and Rural Development (MARD) has been re-organized and re-located 4 times in the last 20 years, specifically in 2005, 2011, 2023 and 2024. Consequently, at MARD, the name of highest agricultural statistics organizations was changed from the Center for Informatics and Statistics (CIS) to the Digital Transformation and Statistics (DTS) in May 2023. In 2025, the Ministry of Agriculture and Rural Development is going to merge with the Ministry of Natural Resources and Environment (MONRE). Therefore, in 2025, the name and structure of DTS is going to change. This new statistics organization will be located in other places and have new statistical colleagues. All these changes are expected to create a better agriculture statistics system where agriculture statistics data will be collected more and more by standard international quantitative methods and stored in certainty manners.

The INAHOR (**IN**ternational **A**sian **H**arvest m**O**nitoring system for **R**ice) software, developed by the Japan Aerospace Exploration Agency (JAXA), is a high-technology tool to estimate paddy areas by regions. The statistical data of paddy area is a key data for many countries including Vietnam. The INAHOR software combines the artificial intelligence (AI), human experiences or "personal opinions" and field tests, therefore, the INAHOR software can generate much more precise and consistent paddy area data over time.

In the last 8 years, in the total of 63 provinces in Vietnam, the Department of Agriculture and Rural Development (DARD) in Hung Yen province is the only one state organization that provided annually provincial state fund to train their statisticians. Although the provincial fund is limited, the leaders of the DARD are committed to apply new methods and technology to improve agriculture statistics in Hung Yen province. Therefore, DTS and DARD have closely interacted over the last 8 years and the INAHOR software will be applied to estimate district and commune paddy areas in Hung Yen province in 2025 and in subsequent years.

2. Objectives of the INAHOR Activity in Vietnam

The main objectives of the INAHOR activity in Vietnam are as follows;

2.1. Long-term objectives

- The INAHOR software will be used by local state statisticians to estimate paddy areas for all communes, districts and provinces in Vietnam, for example the final year is 2030.
- The data generated by INAHOR will be used officially in all paddy production statistical reports at all administrative level in 2030.

2.2. Short-term objectives

- To introduce participants to advanced technologies for estimating paddy area at the district and provincial levels.
- To enable participants with the skills to apply these technologies effectively in practical scenarios.
- To conduct field surveys for verifying selected locations and gathering additional information for each site.
- To understand more about AI including machine learning and deep learning.

3. Details of INAHOR activity in Vietnam

3.1 INAHOR introduction to local statisticians in Hung Yen province (onsite)

Date: 27 July to 10 August 2024
Lecturer: Dr. Hieu Phan Sy, Deputy Head of Statistics and Forecast Division, DTS, MARD
Participants: 150 persons
Training session: the participants were introduced to see how Google Earth works, how INAHOR software works to create paddy maps and estimate paddy area data for each district in Hung Yen province.

3.2 Checking spots in Google Earth for Thai Binh province to provide Input for INAHOR (online)

Date: 1 to 15 September 2024
Guider: Dr. Hieu Phan Sy, Deputy Head of Statistics and Forecast Division, DTS, MARD
Participants: 10 persons
Training session: the participants were introduced to see how to check spots in Google Earth to decide which spots are paddy and non-paddy.

3.3 Online training via ZALO software

Date: 1-5 October 2024
Objective: Since each officer works in a different location, the online training was conducted to ensure that all trainees are well-prepared for the subsequent offline training.
Lecturer: Dr. Hieu Phan Sy, Deputy Head of Statistics and Forecast Division, DTS, MARD
Participants: 10 persons (The list of participants are shown in ANNEX 1**)**
Training session: the participants were trained to use Google Earth, Quantum GIS (QGIS), ArcView, and INAHOR software to create paddy maps and estimate paddy area data by commune and district for Hung Yen province.

3.4 Offline Workshop

Date: 24-25 October 2024
Objective:

- Introduce AFSIS's related activities and details of SAFER project
- Train participants on how to use Google Earth, QGIS, Google Earth Engine (GEE) and the INAHOR software to draw a paddy map with estimated paddy data area.
- Check a paddy map for Thai Binh province.

Lecturer: Mr. Kimura Shoji, the International Consultant of SAFER project and Dr. Pegah Hashemvand Khiabani, the researcher, the Remote Sensing Technology Center of Japan (RESTEC)
Participants: 18 persons (the list of participants is shown in ANNEX 2**)**
Training session:

- Mr. Miyake Yasuhiro, an AFSIS expert, presented Training guidance and Introduction to AFSIS's SAFER Project
- Dr. Kei Oyoshi, Researcher from JAXA, presented about Introduction to Japanese Earth Observation Satellites, the SAFE Project, and Overview of INAHOR software
- Mr. Kimura Shoji, the International Consultant of SAFER project, provided an overview of INAHOR estimation results of 2024 rainy season rice planted area in Thai Binh, Vietnam. According to Thai Binh statisticians' confirmation, the results showed that INAHOR data are very good. The results estimated using INAHOR showed that the paddy cultivated area is 65,645 ha from the total of 173,010 ha or average 37.9%

- Dr. Pegah Hashemvand Khiabani, Researcher from RESTEC, trained on how to practically use INAHOR for all participants.

3.5 Field Survey

Date: 23-25 November 2024

Objective: To How to draw a paddy map and estimated commune paddy area by INAHOR software, An Vien commune such as;

- The commune statisticians review the commune paddy to check "how good" the map is.
- The commune statistician's exam a commune paddy data to check "how good" the data is
- During the survey, commune statisticians provide more information to paddy spots generated in Google Earth

Location: An Vien commune, Tien Lu district, Hung Yen province

Participants: 9 people jointed the field survey (The list of participants is shown in ANNEX 3)

Activity: An Vien commune was selected randomly to draw a commune paddy map. JAXA already provided 12 spots for An Vien commune and DTS added more spots until the paddy map looks good (The outcomes of this activities are shown in ANNEX 4 and 5)

4. Overall Outcomes and Achievements

- Participants in Hanoi, Hung Yen and Thai Binh provinces were well informed INAHOR and imagined how INAHOR generates paddy maps and paddy area data.
- Participants learned well to apply INAHOR tools to estimate district and provincial paddy area.
- Participants can apply INAHOR to real scenarios.
- One commune-level paddy map with paddy area data is generated.
- A great opportunity to learn more about paddy production, cultivation, processing, storing (Annex 6).
- A great opportunity to learn more knowledge related to INAHOR, AI including machine learning and deep learning (Annex 6).

5. Challenges and Lessons Learned

- **Issues Encountered During the Activity**
  - Limited English proficiency among participants, with only 3 participants being able to read English-training documents.
  - Limited availability of functional laptops, with only a few in good working condition.
  - Poor internet connectivity in several field locations, affecting real-time data access and communication.
  - Limited technological proficiency among local participants, which is cause challenges for effective implementation.
- **Lessons Learned**
  - A simplified guide document should be developed for local officers to facilitate the use of INAHOR and other relevant technologies. This would help address challenges related to limited hardware capabilities and unreliable internet connections.

6. Conclusion

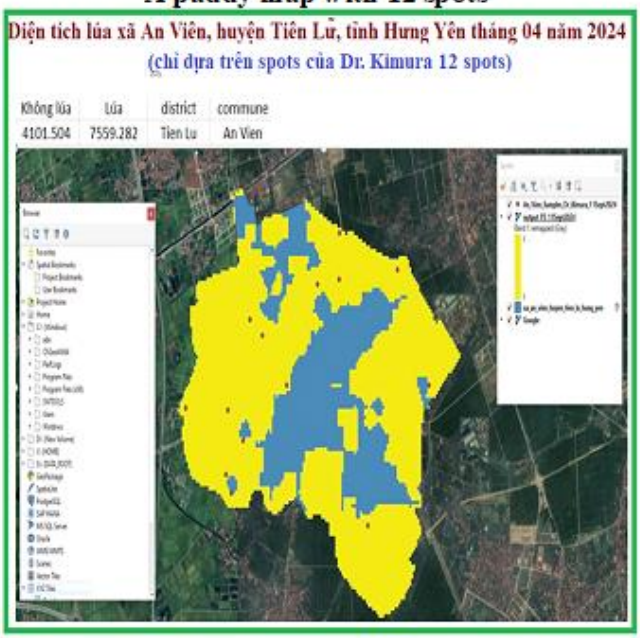
In July and August 2024, the INAHOR system was introduced to more than 150 local statisticians (40% male and 60% female), generating significant interest and enthusiasm among participants. Recognizing the potential of this technology, the leadership of Hung Yen has expressed a keen interest in adopting innovative methods for estimating paddy area at the commune level over the past three years. As a result, INAHOR is being considered for application in 2025 to further strengthen and modernize agricultural statistical practices in the province.

In October 2024, with offline and online trainings, about 8 participants were well trained INAHOR. In November 2024, with the field trip, 2 participants are able to test INAHOR outputs. These participants can become a junior INAHOR trainers for Hung Yen province in 2025. Expectation ally, INAHOR will be applied at Ministry level and other provinces after 2025.

Participants of the offline training 24-25 October 2024

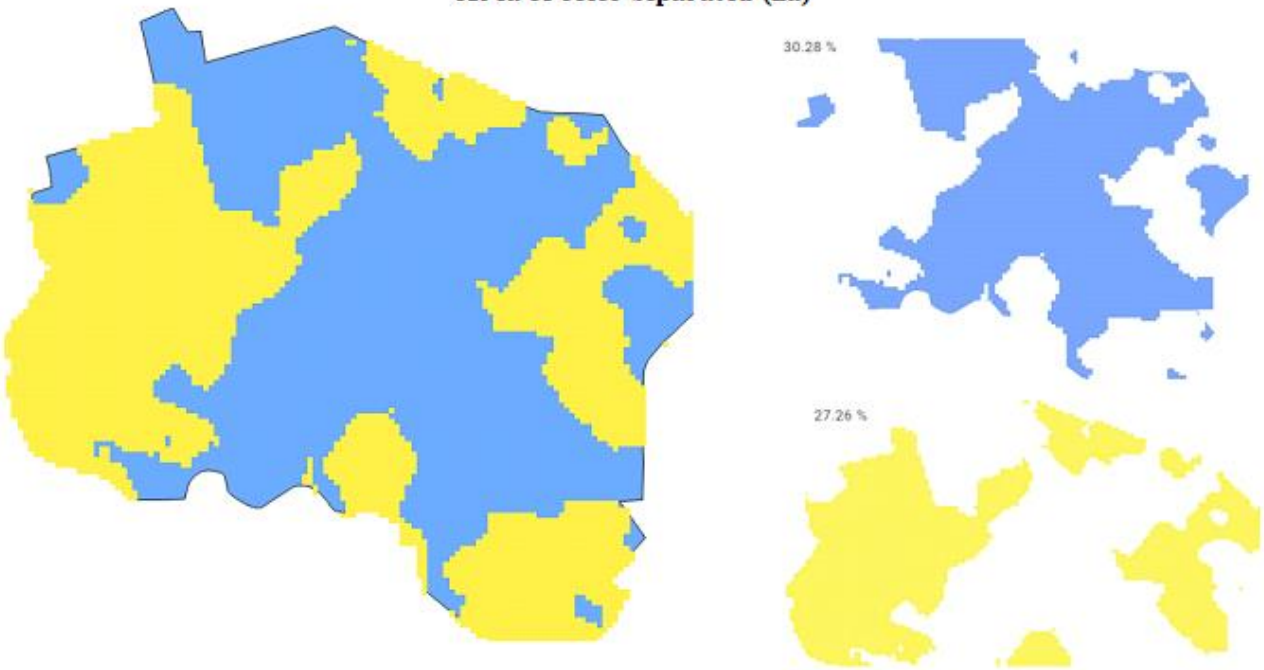
Annex 4. An Vien commune paddy maps generated by INAHOR with different number of spots

A paddy map with 12 spots



A paddy map with 24 spots

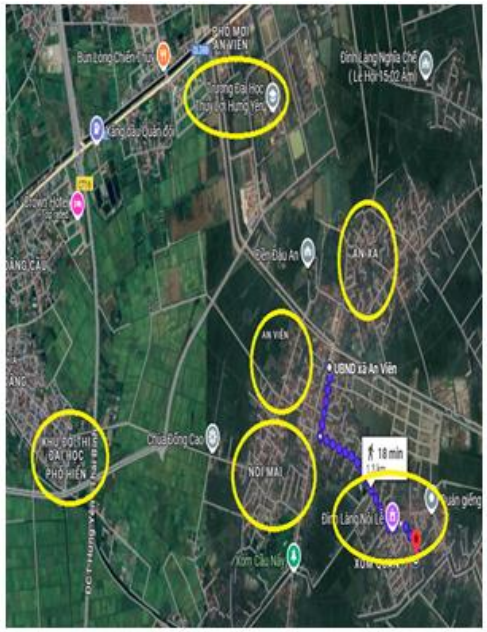
Area of color separated (ha)



	Others	Paddy	Total	% paddy area
Scenario 1 (unclear unit)_Dr. Hieu	4101.5	7559.3	**11660.8**	64.83
Scenario 2 (unclear unit)_Dr. Kimura + Dr. Hieu	7578.6	4082.2	**11660.8**	35.01
Scenario 3 (unclear unit) _ Dr.Kimura	9185.0	2475.7	**11660.8**	21.23
Administrative statistics data (ha)	287	270	557	48.47
Theo color calculate	290.7	261.7	**552.4**	47.38

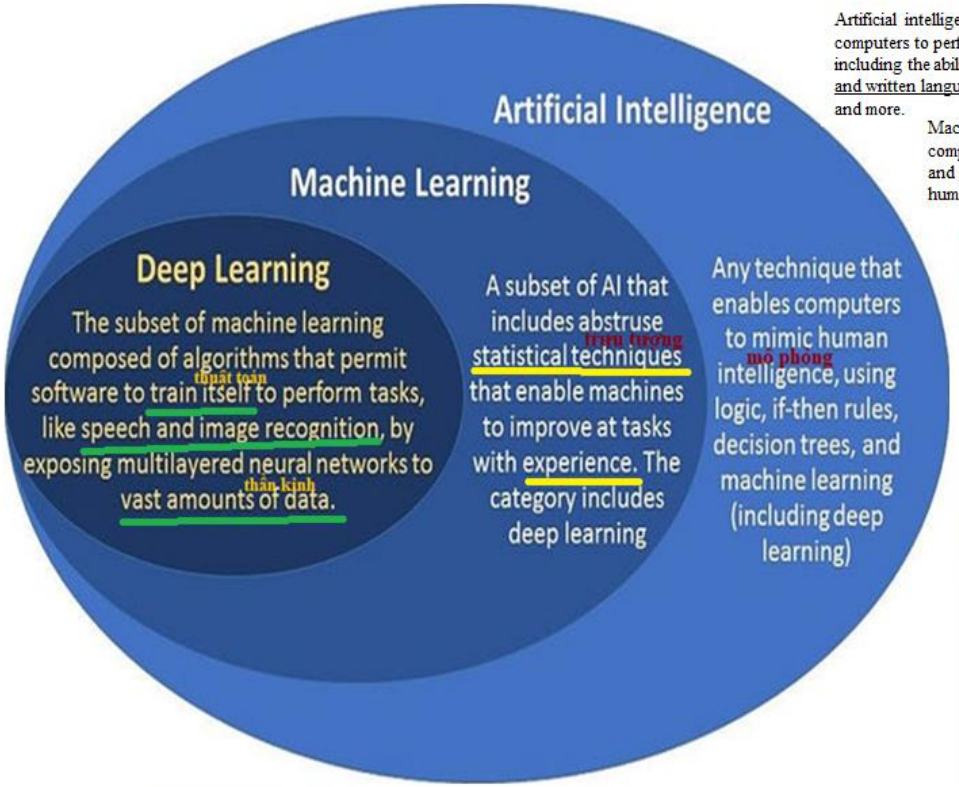
Annex 5. Selected pictures of the training and field survey

Field survey places and activities 23-25 November 2024



Annex 6. Selected knowledge related to AI including machine learning and deep learning

Understanding AI



Artificial intelligence (AI) is a set of technologies that enable computers to perform a variety of advanced functions, including the ability to see, understand and translate spoken and written language, analyze data, make recommendations, and more.

Machine learning (ML) is a branch of and computer science that focuses on the using data and algorithms to enable AI to imitate the way that humans learn, gradually improving its accuracy.

Deep learning is a subset of machine learning that uses multilayered neural networks, called deep neural networks, to simulate the complex decision-making power of the human brain. Some form of deep learning powers most of the artificial intelligence (AI) applications in our lives today.

What is a neural network?

A neural network is a machine learning program, or model, that makes decisions in a manner similar to the human brain, by using processes that mimic the way biological neurons work together to identify phenomena, weigh options and arrive at conclusions.

Tự tóm tắt về Deep Learning (học sâu)

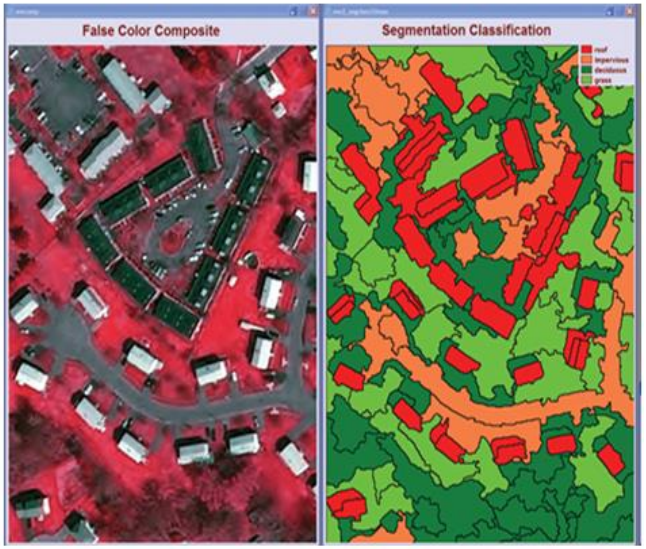
Người ta lắp đặt nhiều con chip nhỏ ở các bộ phận khác nhau của bộ não. Khi con người thao tác với 1 hoạt động nào đó, các bộ phận khác nhau của bộ não sẽ phát sóng ra ngoài khác nhau. Các con chip sẽ ghi lại các sóng đó, mỗi sáng tương ứng với 1 layer. Các thuật toán sẽ tự động nối các layer lại với nhau và tự động đưa ra kết quả công việc.

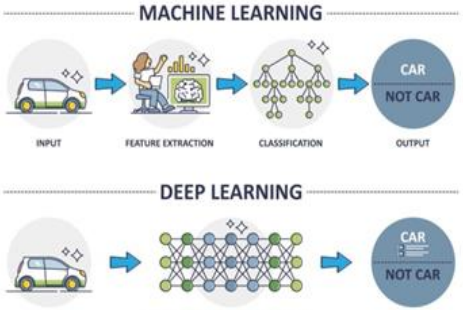
Tự tóm tắt

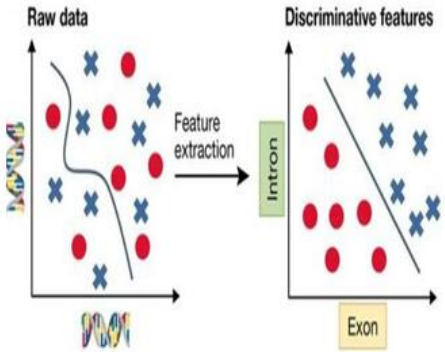
Thứ nhất, AI (trí tuệ nhân tạo) ở bậc chung nhất là bất kỳ dạng nào để máy tính thao tác theo suy nghĩ của con người để làm loại công việc nào đó: ví dụ như viết program, code để tổng hợp số liệu thống kê.

Thứ hai, Machine learning (Máy đang học), là ở bậc dùng rất nhiều toán thống kê, xác suất, có nhiều thứ trừu tượng hơn...kết hợp với kinh nghiemehj của người để cải thiện nhiệm vụ loại việc nào đó. Ví dụ việc xác định cỡ mẫu, địa bàn mẫu và kinh nghiệm của người khi kiểm tra các spot là lúa hay không phải lúa.

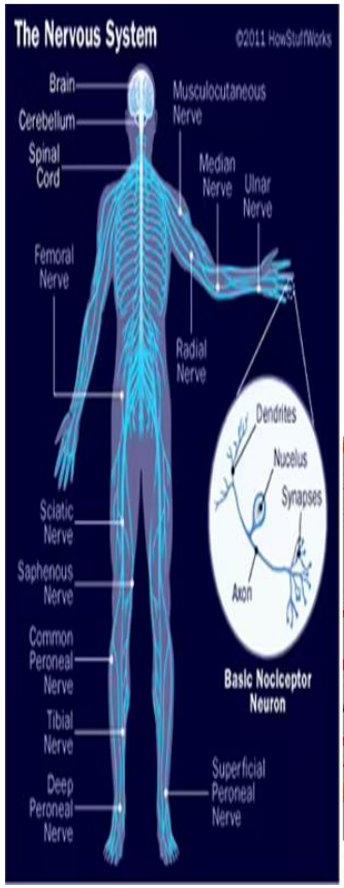
Thứ ba, deep learning (học sâu), với số lượng số liệu vô cùng lớn, các phần mềm được trang bị các thuật toán để phần mềm máy tính tự học, tự nâng cao mức độ chính xác của công việc.

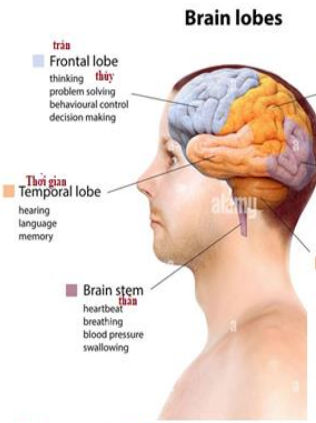




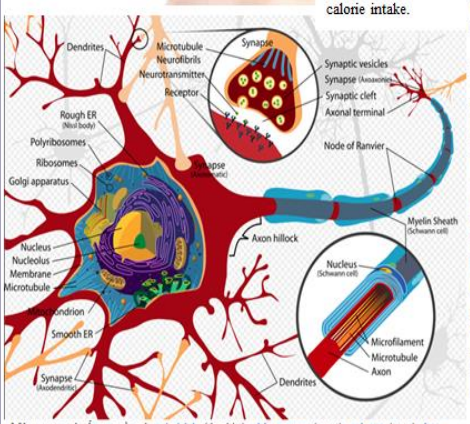


How brain and AI operates

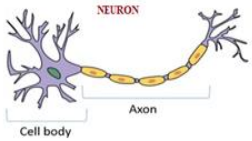




The brain accounts for about 2% of total body weight, yet receives 20% of our blood supply and uses 20% of our total calorie intake.



Một neuron và cấu tạo của nó: sợi nhánh (dendrite), thân neuron (soma), sợi trục (axon), bao mi-ê-lin (myelin sheath), eo răng-vi-ê (node of ranvier), synapse (synapse)



The average adult brain contains around 100 billion brain cells. Each is connected to around 1,000 others. That's 100 trillion connections.

The brain performs an incredible number of tasks including the following:

- It controls body temperature, blood pressure, heart rate and breathing.
- It accepts a flood of information about the world around you from your various senses (seeing, hearing, smelling, tasting and touching).
- It handles your physical movement when walking, talking, standing or sitting.
- It lets you think, dream, reason and experience emotions.

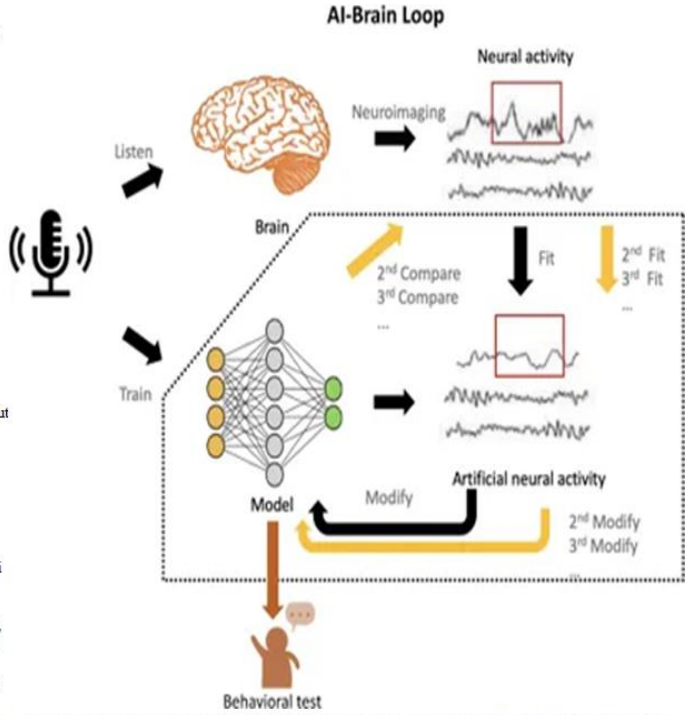
Your brain contains billions of brain cells, called 'neurons.' Each brain cell has a cell body and axons. The cell bodies appear grey-ish in colour, so they're known as **'grey matter'**. They control all of the brain's functions.

Axons radiate out from the cell bodies, forming connections – like wiring – between brain cells.

Every animal you can think of -- mammals, birds, reptiles, fish, amphibians -- has a brain. But the human brain is unique. Although it's not the largest, it gives us the power to speak, imagine and problem solve. It is truly an amazing organ.

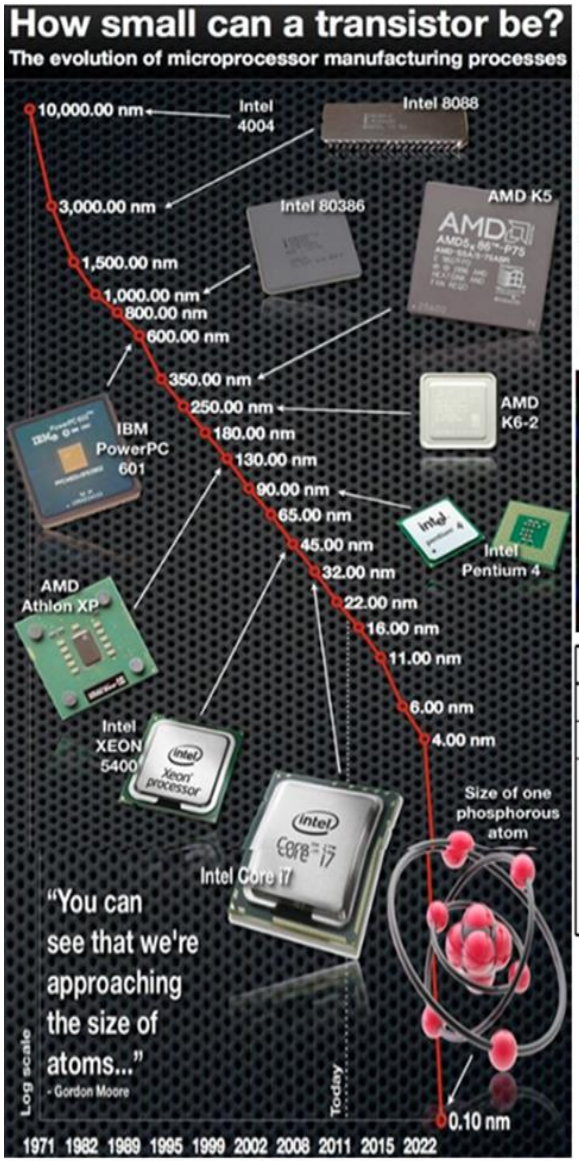
Neuron là tế bào thần kinh có khả năng cảm ứng, phát sinh xung thần kinh và dẫn truyền xung điện này. Neuron là đơn vị cơ bản cấu tạo hệ thần kinh của hầu hết các loài động vật và là thành phần quan trọng bậc nhất của não. Thân và sợi nhánh của các neuron tạo thành chất xám. Sợi trục (nếu đi chung với nhau thành bó gọi là dây thần kinh) cấu tạo chất trắng trong não Ước tính có khoảng 100 tỷ (10^{11}) neuron và 100 nghìn tỷ (10^{14}) synapse trong não người. Các tế bào thần kinh được hỗ trợ bởi microglia và tế bào hình sao (các tế bào thần kinh đệm). Neuron là những tế bào dài nhất trong cơ thể, biệt hóa cao độ nên mất trung thể và khả năng phân chia, nhưng có khả năng tái sinh phần cuối sợi trục nếu bị tổn thương.

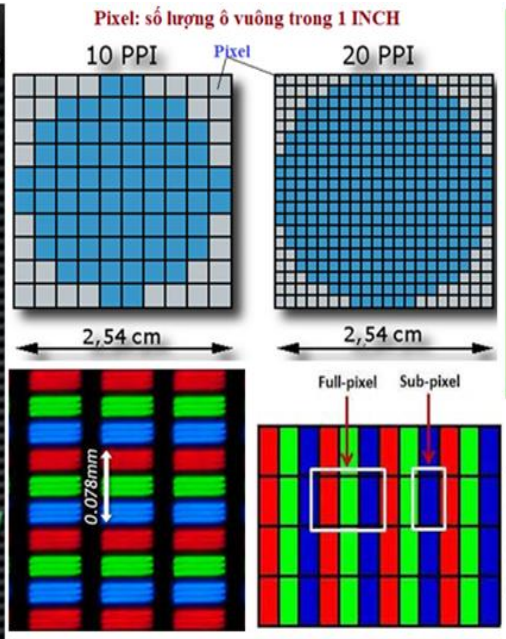
Tủy gai hay **tủy sống** là một cơ quan quan trọng trong cơ thể, chạy dọc bên trong xương sống, từ tủy sống phát sinh các thần kinh gai tạo liên hệ từ não đến toàn bộ cơ thể. tuỷ gai bao gồm chất xám ở giữa và bao quanh bởi chất trắng.
Tủy gai không chỉ là phần nối dài của bộ não, nó còn giữ một chức năng vô cùng quan trọng, đó là sự phản xạ. Có thể nói rõ là sự phản ứng tức thời của cơ thể mà không cần não xử lý, nó bảo vệ chúng ta khỏi những mối nguy hiểm hằng ngày. Ví dụ khi bạn lỡ tay chạm vào nồi canh nóng, lập tức cơ thể bạn sẽ tự xử lý bằng cách ngón tay rụt phắt lại.
Tủy gai nằm xuyên bên trong ống cột sống, nó được bao bọc bởi ba lớp màng: mành ngoài gọi là màng cứng, màng giữa gọi là màng nhện, màng trong gọi là màng nuôi.



Schematic of the AI-Brain loop. First, present the same task (e.g., audio recognition) to human subjects and AI models, which are subsequently trained for this task. Second, record the neural activities in the brain by neuroimaging techniques (e.g., fMRI., EEG., MEG., ECoG), and predict neural responses with these trained AI models. Third, compare the recorded neural activity and the artificial neural activity generated by models. Fourth, use the artificial neural activity to fit the neural activity by modifying the corresponding layers or parameters. Fifth, implement the behavioral evaluation of model and see whether the performance achieve the human-like level. If not, implement the continuous fit until it achieves the both human-like and brain-like level.

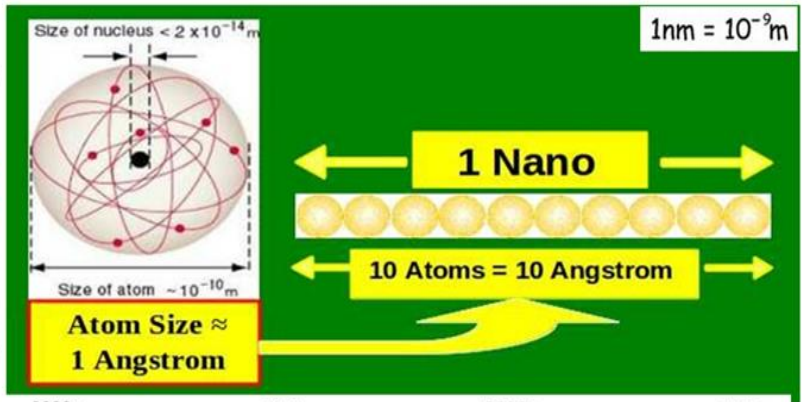
Sizes of relevant components related to AI

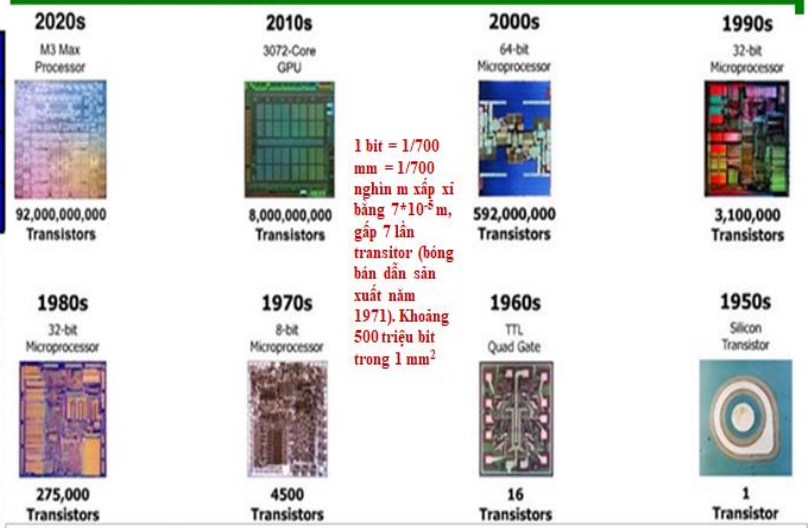




	1.2µm	0.35µm	0.18µm
cell size	112 x 81.4 µm²	21.5 x 19.6 µm²	9.22 x 7.56 µm²
chip size	11.7 x 8.6 cm²	2.4 x 2.2 cm²	1.1 x 1 cm²
photograph			

Transitor (bóng bán dẫn) đầu tiên năm 1947 có chiều dài khoảng 4 cm, 0,04 m hay $4x10^{-2}$ m. Bóng bán dẫn hiện nay là $4x10^{-9}$ m, tức là bóng đầu tiên gấp khoảng 10 triệu lần bóng gần đây





1 bit = 1/700 mm = 1/700 nghìn m xấp xỉ bằng $7*10^{-5}$ m, gấp 7 lần transitor (bóng bán dẫn sản xuất năm 1971). Khoảng 500 triệu bit trong 1 mm^2

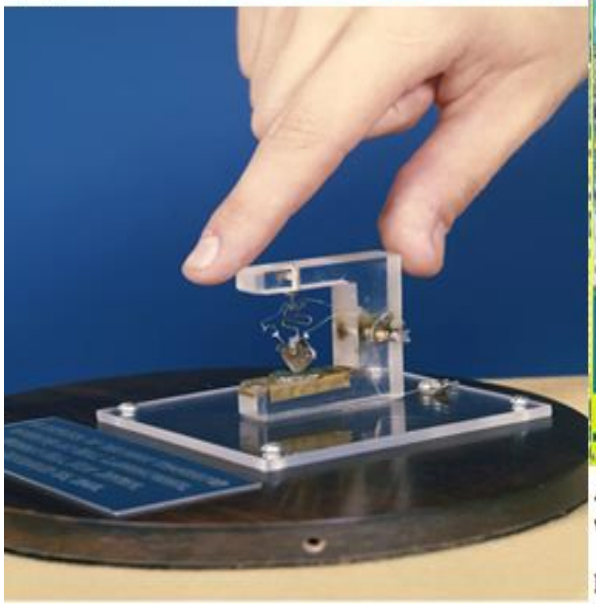
Một số ghi chú về kích cỡ:

1 nm = 1 nano = 1 phần 1 tỷ của 1 mét
Đường kính của một nguyên tử bằng 10 nm = 1 phần 10 tỷ của 1 mét.
Đường kính của một hạt nhân nguyên tử nhỏ hơn 1 phần 20 nghìn tỷ của 1 mét
Đường kính của 1 bóng bán dẫn hiện đại nhất hiện nay bằng 4 nm = 1 phần 4 tỷ của 1 mét.
- Một con CHIP có 1 bóng bán dẫn những năm 1950 tăng lên 92 tỷ bóng bán dẫn những năm 2020
Đường kính của 1 pixel hiện đại bằng khoảng 0,08 mm = bằng 8 phần 10 nghìn của 1 mét
- Đường kính của 1 sub-pixel hiện đại bằng 1/3 của 1 pixel = 8 phần 30 nghìn của 1 mét.
Một màn mình như điện thoại Samsung có diện tích bằng 2x4 = 12 $inch^2$:
- Có ít nhất 1200 pixel, lên đến 8000 pixel, thậm chí nhiều hơn.

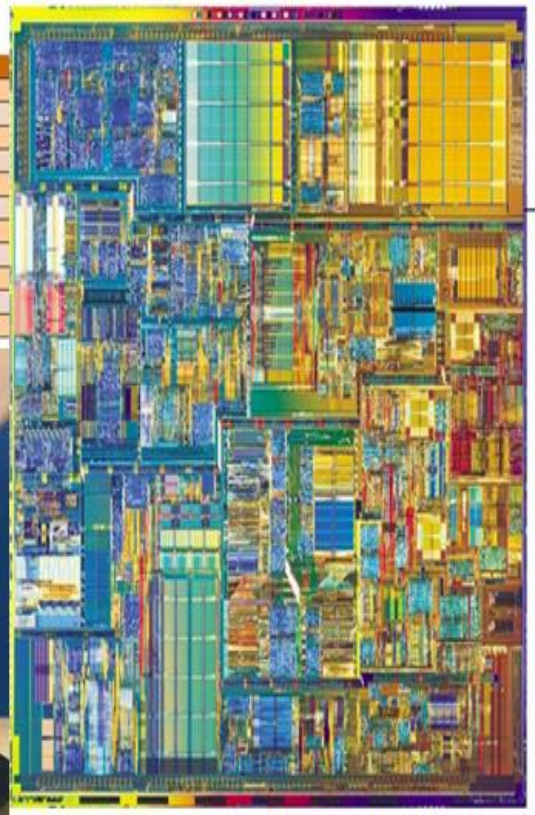
Understanding bit, byte, Transistor, chip, CPU

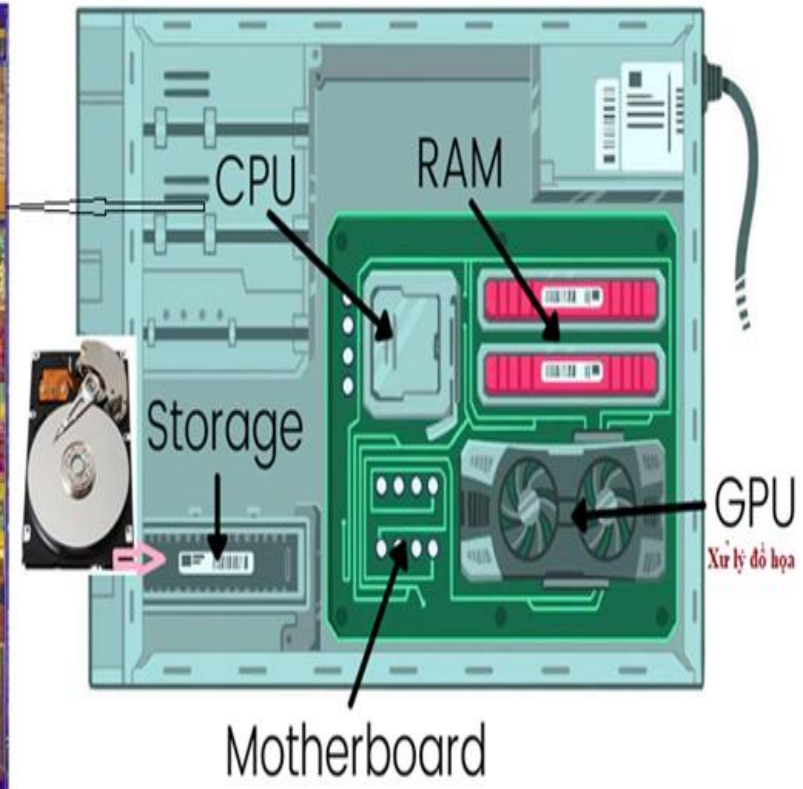
ĐƠN VỊ ĐO LƯỜNG LƯU TRỮ TRÊN MÁY TÍNH

Đơn vị	Ký hiệu	Tương đương	Tương đương với số byte
byte	b	8 bits	1 byte
kilobyte	Kb	1024 bytes	1 024 bytes
megabyte	MB	1024 KB	1 048 576 bytes
gigabyte	GB	1024 MB	1 073 741 824 bytes
terabyte	TB	1024 GB	1 099 511 627 776 bytes
Petabyte	PB	1024 TB	1 125 899 906 842 624 bytes
Exabyte	EB	1024 PB	1 152 921 504 606 846 976 bytes
Zetabyte	ZB	1024 EB	1 180 591 620 717 411 303 424 bytes
Yottabyte	YB	1024 ZB	1 208 925 819 614 629 174 706 176 bytes
Brontobyte	BB	1024 YB	1 237 940 039 285 380 274 899 124 224 bytes
Geopbyte	GB	1024 BB	1 267 650 600 228 229 401 496 703 205 376 bytes



November 17, 1947 American scientists John Bardeen and Walter Houser Brattain observe the basic principles of the transistor, a key element for the electronics revolution of the 20th century. Depicted is a replica of the first transistor





A CPU is a silicon integrated circuit, sometimes called a computer chip. This is because it consists of a square piece of crystallized silicon that is very, very thin.

It has a very intricate pattern on top, which consists of the transistors and wires that have been processed into the silicon: a silicon chip contains millions of tiny transistors connected by very tiny wires. The chip is also connected to the rest of the computer with very tiny strands of wire.

Modern processors often have a heat sink added on top of the CPU: a heat sink is a set of metal fins that quickly dissipate the heat from the CPU into the surrounding air. The heat sink is necessary for some computer chips because they run so fast that the heat they generate can cause damage. Most heat sinks also have a small fan to improve the cooling: they are like a radiator for the CPU.

Understanding paddy production, processing, products and store

Người dân trực tiếp dùng tay để thu hoạch lúa.

Lúa được bỏ vào bao sau khi tuốt.

Lúa được người dân phơi khô, sàn lọc bỏ những hạt lép trước khi cho lúa vào kho.

Sau khi khô, lúa được đưa vào kho cất. Các hộ dân nơi đây nhà nào cũng có một kho lúa để dự trữ

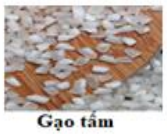
Hạt thóc

Hạt lúa

Vỏ trấu

Gạo thường

Gạo lứt



Gạo tấm

Gạo Xát dối hay còn gọi là **GẠO CÒN CÁM, GẠO HẨM** (các địa phương vùng đồng bằng bắc bộ) là gạo mà khi xay xát còn giữ lại 1 tỷ lệ cám nhất định, chính là loại gạo mà mọi người ăn từ những năm 1990 về trước. Như vậy Gạo Xát dối khác gạo lứt, gạo trắng là ở tỷ lệ cám còn giữ lại. Cụ thể:

- Gạo lứt: 100% cám
- Gạo Xát dối của Gạo Tươi Mỗi Ngày giữ lại khoảng: 50%-70% cám
- Gạo trắng: 0% cám (đã chà sạch lớp cám và đánh bóng gạo)

--> Như vậy, có thể hiểu gạo Xát dối là loại gạo nằm giữa gạo lứt và gạo trắng.

Hạt lúa nảy mầm

Hạt thóc lép chuyển mầu

Gạo tấm là gì?
Gạo tấm là những mảnh vụn của gạo bị vỡ trên đồng lúa, khi phơi khô, khi vận chuyển hoặc khi xay sàng gạo. Máy móc sau đó sẽ phân tách những hạt gạo tấm khỏi gạo nguyên hạt và phân loại theo kích cỡ.

Hạt lép là hạt thóc không có lõi. - Hạt lẫn loại là những hạt thóc khác giống, có kích thước và hình dạng khác với hạt thóc theo yêu cầu.

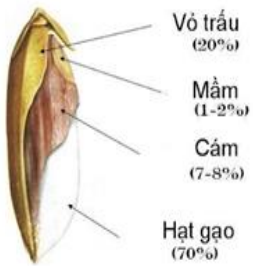
1. Rạ là phần thân dưới của cây lúa.
2. Rơm là phần ngọn mang trọn vẹn hạt lúa trên đó.
3. Rạ thì được phơi lại ngoài đồng đến khô mới gánh về.
4. Rơm có màu vàng.
5. Rạ có màu xám.
6. Rơm được phơi khô để làm thức ăn cho trâu bò khi mùa đông (tháng giá cỏ tươi cao).
7. Rạ dùng để lợp mái nhà hoặc làm chất đốt cho việc đun nấu thức ăn.

Rơm ơ nông thôn Việt Nam

Rạ lúa

Bột cám gạo

Cối giã gạo



Gạo 5%, 10%, 25% tấm
Gạo 5% tấm, gạo 25% tấm là quy cách số lượng tấm trong gạo; Gạo 5% tấm là bình quân 100 hạt gạo chỉ lẫn 5 tối đa hạt tấm, 25% tấm là chỉ được lẫn tối đa 25 hạt tấm trong bình quân 100 hạt gạo, đây là cách xác định quy cách sản phẩm trong hợp đồng mua bán trao đổi, giao dịch trong thương trường.