Artificial Intelligence Glossary:

AI: Key Concepts, Terminology and Geospatial Considerations

Pre-Release Version

Researcher: Abby Heazlewood

Editor: Nathan Heazlewood

Contributors: Vipul Bhatia and Robert Mi

EAGLE TECHNOLOGY

Introduction

Artificial Intelligence (AI) is inherently complex because it spans multiple disciplines computer science, mathematics, cognitive science, ethics, and domain-specific applications like GIS. It involves not just algorithms and models, but also concepts such as learning paradigms (supervised, unsupervised, reinforcement), neural network architectures, probabilistic reasoning, and emerging technologies like transformers and generative models. Each of these areas has its own terminology, workflows, and underlying theory, which can feel overwhelming to professionals whose expertise lies primarily in spatial analysis and geospatial technologies.

Understandably, many GIS professionals approach AI by starting with what they know best: GIS workflows and tools. They often explore AI through practical applications such as automated feature extraction, land cover classification, or predictive modelling in ArcGIS. This GIS-first approach makes sense because it provides immediate relevance and tangible outcomes using AI to enhance familiar tasks like image segmentation or spatiotemporal modelling. However, this method can sometimes limit understanding to tool-specific functionality rather than the broader principles that drive AI innovation.

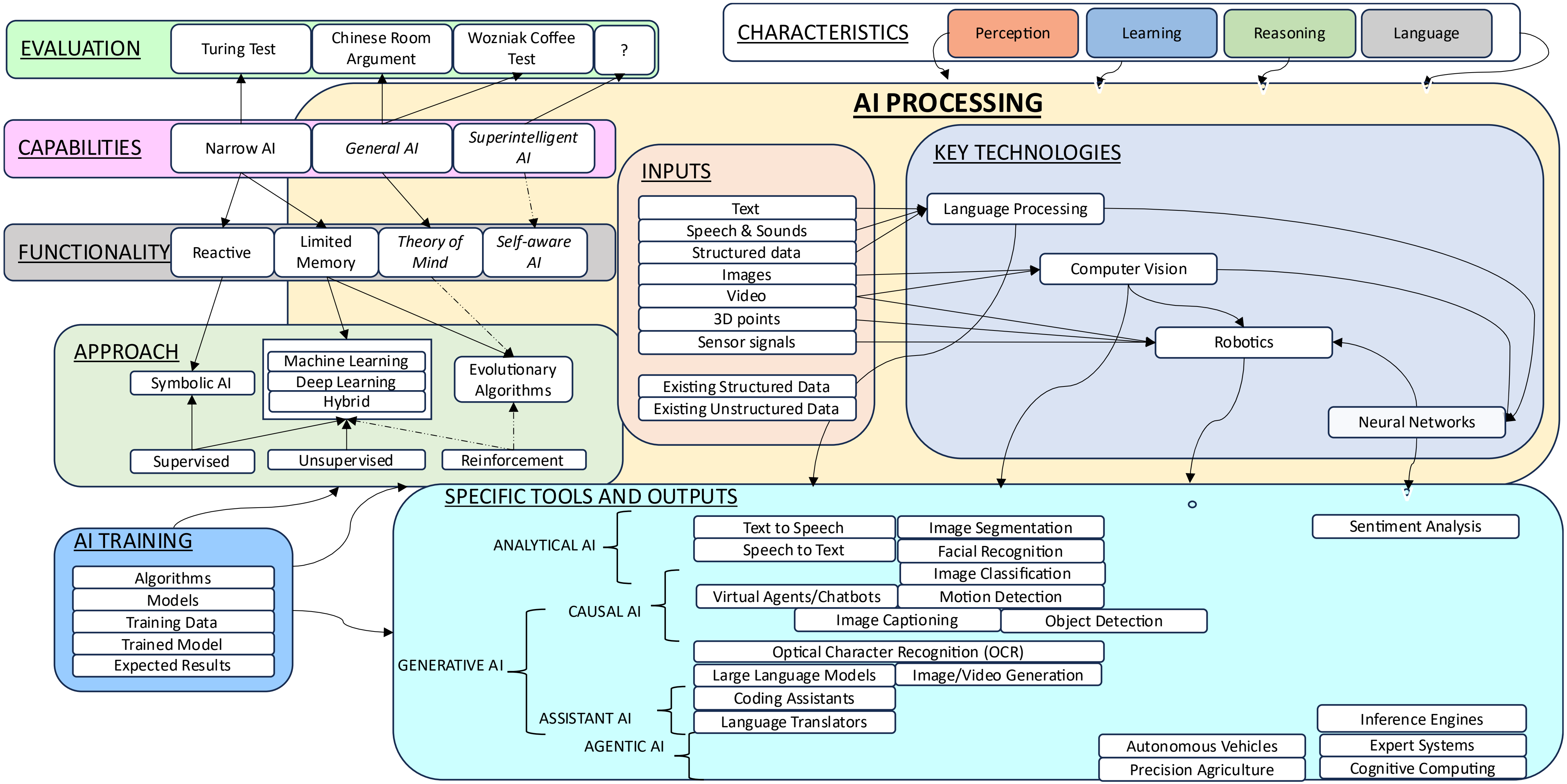
An alternative and arguably more robust approach is to begin with the **core concepts and terminology of AI**. This means understanding foundational ideas like what constitutes a neural network, how machine learning differs from deep learning, what "attention mechanisms" do in transformers, and why concepts like overfitting, bias, and explainability matter. Once these fundamentals are clear, GIS professionals can then map these concepts to geospatial contexts: for example, recognising that convolutional neural networks (CNNs) excel at raster imagery analysis, or that reinforcement learning could optimise routing in dynamic transport networks. This perspective not only deepens technical literacy but also empowers professionals to evaluate new tools critically, adapt to emerging trends, and even design bespoke AI solutions for spatial problems.

This document follows this approach: AI core concepts and key terminology is the starting point, and then the relationship of those items is elaborated on with a geospatial perspective.

This document is intended to be a self-development resource for geospatial professionals to aid their understanding of AI and its impact on the realm of GIS. It could also be used by AI professionals wanting to understand how AI is being used within the geospatial domain.

Text in green boxes focusses on geospatial interpretations, examples or resources. Some readers may want to have a focus on this information.
Text in grey boxes represents highly detailed definitions or concepts that are less important for most geospatial professionals (unless for example you are a GIS Developer or have a similar highly technical role). Many readers could decide to skip the grey boxes.

The diagram on the following pages shows how the groups of concepts have been organised.



EVALUATION
Turing Test
Chinese Room Argument
Wozniak Coffee Test
?
CHARACTERISTICS
Perception
Learning
Reasoning
Language
AI PROCESSING
CAPABILITIES
Narrow AI
General AI
Superintelligent AI
FUNCTIONALITY
Reactive
Limited Memory
Theory of Mind
Self-aware AI
INPUTS
Text
Speech & Sounds
Structured data
Images
Video
3D points
Sensor signals
Existing Structured Data
Existing Unstructured Data
KEY TECHNOLOGIES
Language Processing
Computer Vision
Robotics
Neural Networks
APPROACH
Symbolic AI
Machine Learning
Deep Learning
Hybrid
Evolutionary Algorithms
Supervised
Unsupervised
Reinforcement
AI TRAINING
Algorithms
Models
Training Data
Trained Model
Expected Results
SPECIFIC TOOLS AND OUTPUTS
ANALYTICAL AI
CAUSAL AI
GENERATIVE AI
ASSISTANT AI
AGENTIC AI
Text to Speech
Speech to Text
Image Segmentation
Facial Recognition
Image Classification
Virtual Agents/Chatbots
Motion Detection
Image Captioning
Object Detection
Optical Character Recognition (OCR)
Large Language Models
Image/Video Generation
Coding Assistants
Language Translators
Sentiment Analysis
Inference Engines
Expert Systems
Cognitive Computing
Autonomous Vehicles
Precision Agriculture

PART 1: AI IN THE CONTEXT OF GEOSPATIAL

Artificial Intelligence (AI) is transforming the work of GIS teams by automating data processing, improving accuracy, and enabling advanced analytics such as predictive modelling and feature extraction. This speeds up workflows, reduces some types of errors (while needing to have a human in the loop to watch for other types), and unlocks deeper insights, helping organisations make faster, smarter spatial decisions.

How is AI defined?

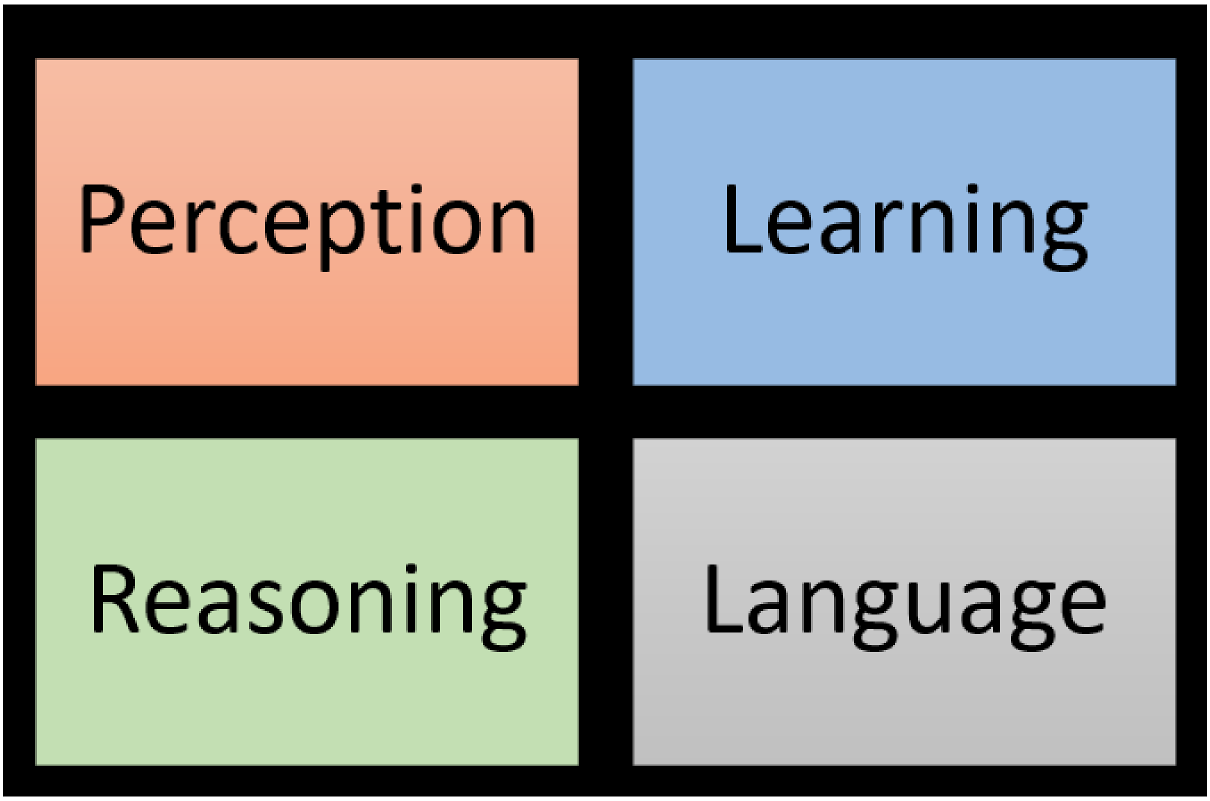


Figure 1- Four key characteristics of AI that are the basis of this framework.

To begin to understand the impact of AI on GIS, it is useful to consider how AI is defined. The most widely accepted definition of Artificial Intelligence (AI) is: **"the field of computer science that focuses on creating systems capable of performing tasks that typically require human intelligence. These tasks include reasoning, learning, perception, and understanding language."** In practice these four key characteristics of AI are often interlinked and overlap. This definition is often attributed to foundational work by John McCarthy, who coined the term "artificial intelligence "in 1956, and is echoed in modern interpretations by organizations like Stanford University and Oxford.

Every organisation is different to the next organisation, therefore one of the complexities determining how a GIS Team's strategy should incorporate AI will differ from organisation to organisation. The four key characteristics of AI described above can then be used to structure key questions to simplify this complexity.

Key Questions for applying AI to GIS within an organisation

The key questions to consider with this analysis are:

1. **Do you use, or would it be beneficial to start using; data from sensors?**
2. **Does your organisation use, or would it be beneficial to start using; images, video or other data from the field?**
3. **Do you use, or would it be beneficial to start using; images, video or other data from flying or orbiting data capture platforms?**
4. **Do you use, or would it be beneficial to enhance your use of existing data assets?**
5. **Do you need to localise or improve the accuracy of image interpretation AI models?**
6. **Do you need to improve services or applications through analysis of user actions, preferences, sentiment, corrections, or trends?**
7. **Do you have processes that could be automated? This can include decision making, or inference from partial data.**
8. **Do you need to broaden the data used for decision making, or make the assessment process more streamlined? Do you want to enhance the accuracy of datasets?**
9. **Do you want wider capabilities to automatically generate maps or create apps?**
10. **Do you want to increase the efficiency and accuracy of basic coding?**
11. **Do you want to be able to automate metadata, documentation, or specifications?**
12. **Do you want to be able to use or implement automated support / training assistants? (including for your own internal systems)**
13. **Do you want to be able to localise or translate components of systems or documentation?**
14. **Do you want to be able to search across all of your databases and documents in natural language, including using approximate location references?**

These questions are further structured and explained in the following sections.

AI Perception

AI perception refers to the ability of artificial intelligence systems to sense and interpret the physical world through various inputs, mimicking human perception. In the context of geospatial technology this primarily involves computer vision, which enables machines to analyze and understand visual data from cameras and imagery. Beyond vision, AI perception extends to other sensory modalities such as sound (via speech and acoustic analysis), vibration (for detecting movement or structural integrity), and data from specialized sensors like LiDAR, radar, temperature, and pressure sensors. By fusing these diverse inputs, AI can build a rich, multi-dimensional understanding of environments, enabling applications in autonomous vehicles, industrial monitoring, smart cities, and robotics.

AI can leverage data from multiple sensors—one example might be a situation where rainfall gauges, water level monitors, and flow speed sensors—aggregate and correlate readings to understand the hydrological state of a region. The spatial location of these sensors is critical because water movement is inherently geographic. Using geospatial models, AI can map sensor data to river networks and catchment areas, enabling predictive modelling of flood risks and estimating how upstream rainfall will impact downstream water levels.

Agentic AI takes this a step further by making autonomous, localised decisions based on sensor inputs and predefined goals. For example, an AI agent controlling a floodgate can monitor nearby sensors, evaluate thresholds, and decide to open gates without waiting for central approval. These agents can also coordinate with others—such as downstream gates—to prevent conflicting actions and ensure system-wide safety. By integrating predictive models, agentic AI can act proactively, such as lowering reservoir levels ahead of forecasted heavy rain, reducing flood risk and improving response times.

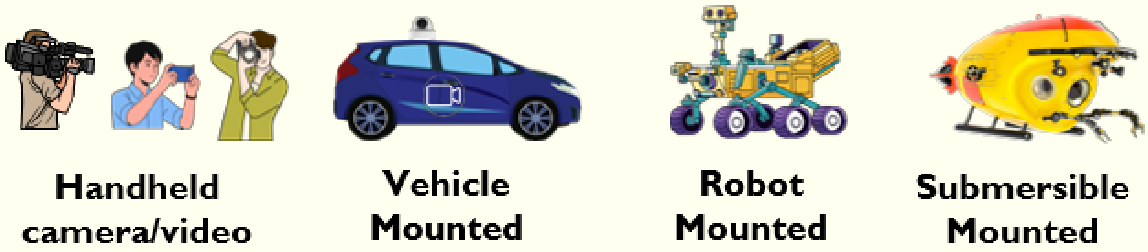
1. Do you use, or would it be beneficial to start using; data from sensors?

Telemetry from sensors GeoAI can identify and locate sensors that are giving specific readings or an alert. Those readings can then be used in subsequent AI reasoning to determine automated responses or actions. The relationships and proximity between different sensors can also be critical. Examples only depicted here: there are multitudes of other sensors.

2. Do you use, or would it be beneficial to start using; images, video or other data from flying or orbiting data capture platforms?

Automated analysis of surface* images, video, or remotely sensed data (such as LiDAR) to extract and identify features present in the image, or to determine changes, or to read text in the image. (e.g. reading road-signs).

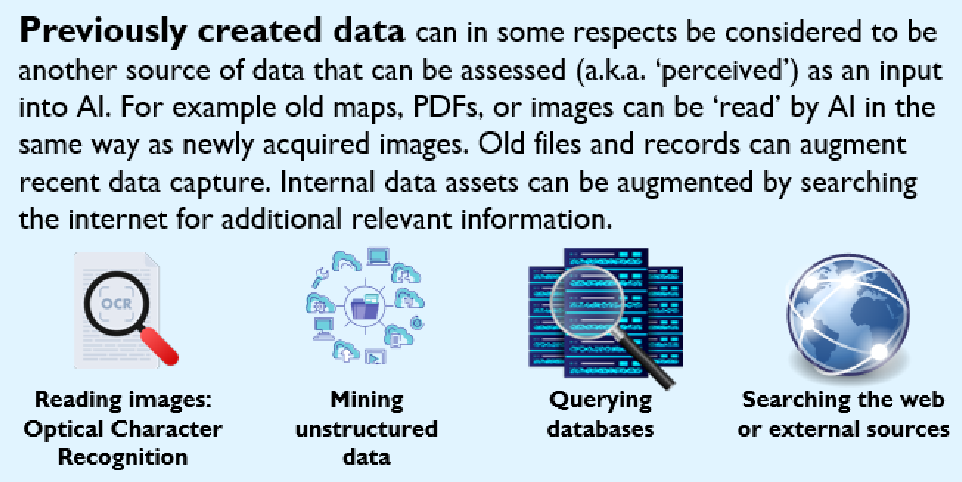
[or sub-surface images such as mines, caves or pipes]*



3. Do you use, or would it be beneficial to start using; images, video or other data from flying or orbiting data capture platforms?

Automated analysis of images, video, or remotely sensed data from satellites, aircraft, or drones to extract the features present in the image, or to determine changes that have occurred over time.

4. Do you use, or would it be beneficial to enhance your use of existing data assets?

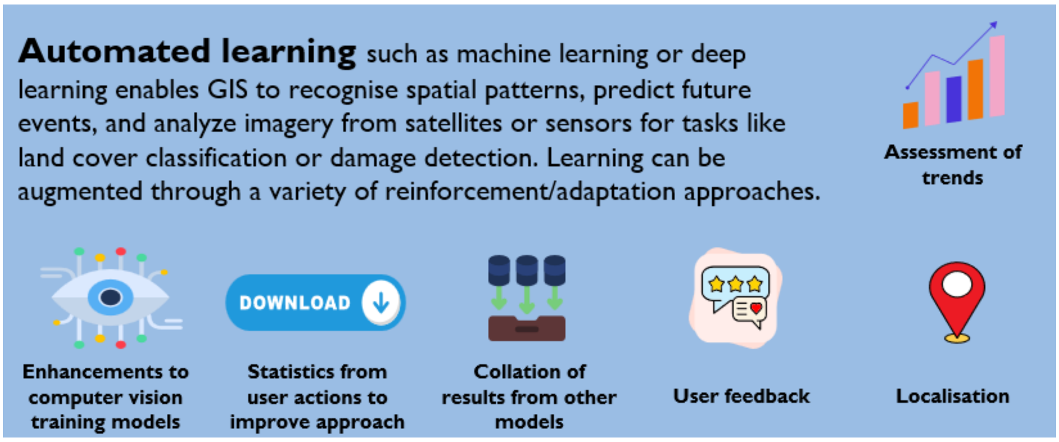


AI Automated Learning

AI Automated Learning is increasingly vital to GIS because it enables systems to continuously refine and adapt without manual intervention, improving both efficiency and accuracy. In geospatial contexts, automated learning enhances image interpretation by localising features more precisely and reducing misclassification errors, which is critical for applications such as land-use mapping or disaster response. Furthermore, by analysing user actions, preferences, sentiment, corrections, and emerging trends, AI can personalise services and optimise workflows, leading to smarter applications that anticipate user needs and deliver more relevant insights. This dynamic feedback loop not only improves data quality but can also drive enhancements and innovation in location-based services and decision-making tools.

5. Do you need to localise or improve the accuracy of image interpretation AI models?

6. Do you need to improve services or applications through analysis of user actions, preferences, sentiment, corrections, or trends?



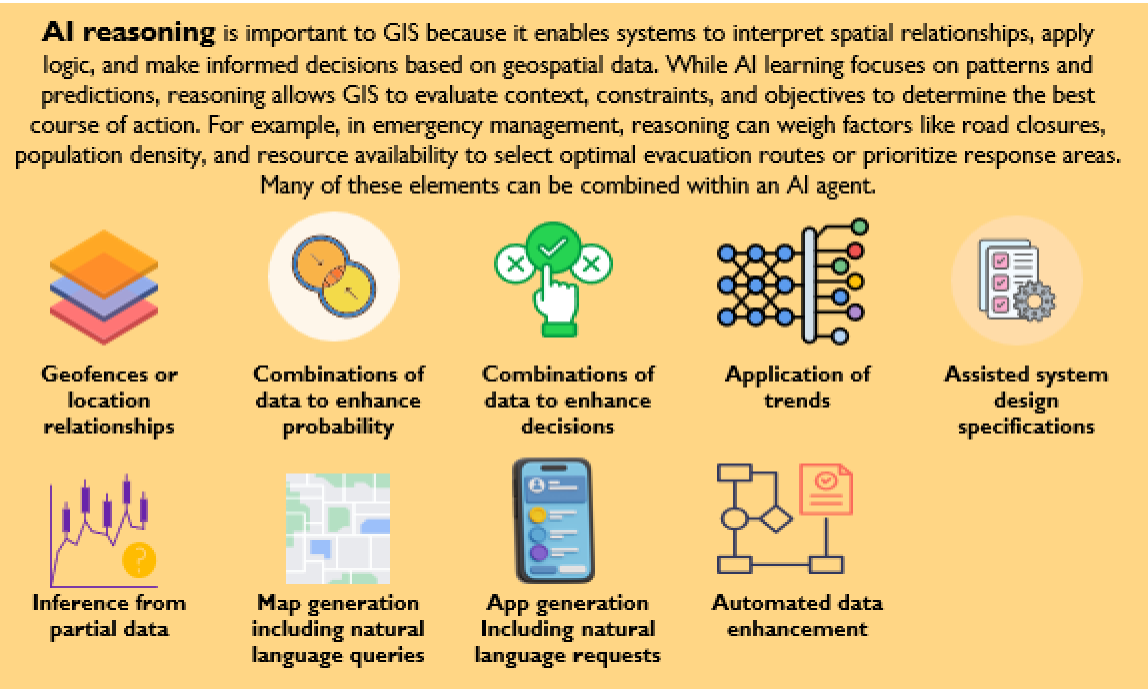
AI Reasoning

AI reasoning and problem solving are crucial to GIS because they enable systems to automate complex processes that traditionally require human expertise, such as decision-making and inference from incomplete or uncertain data. By applying advanced reasoning, AI can rapidly broaden the scope of data used for spatial analysis, integrating diverse sources to make assessments more comprehensive and streamlined. This capability can also be used to enhance the accuracy of geospatial datasets by identifying patterns, correcting inconsistencies, and filling gaps intelligently. AI can also automatically generate maps tailored to specific objectives or even create applications that respond dynamically to user needs, significantly reducing development time and improving usability.

7. Do you have processes that could be automated? This can include decision making, or inference from partial data.

8. Do you need to broaden the data used for decision making, or make the assessment process more streamlined? Do you want to enhance the accuracy of datasets?

9. Do you want wider capabilities to automatically generate maps or create apps?



AI Language Processing

AI language processing capabilities are highly important to GIS because they significantly improve efficiency and accuracy in tasks that involve text and communication. For example, they can streamline basic coding by generating or correcting scripts, automate the creation of metadata, documentation, and technical specifications, and even power intelligent support or training assistants tailored to an organisation's customised systems. Language processing also enables localisation and translation of system components or documentation, making GIS solutions more accessible globally. Furthermore, it allows users to search across all databases and documents using natural language queries, including approximate location references, which simplifies data discovery and enhances usability for non-technical stakeholders.

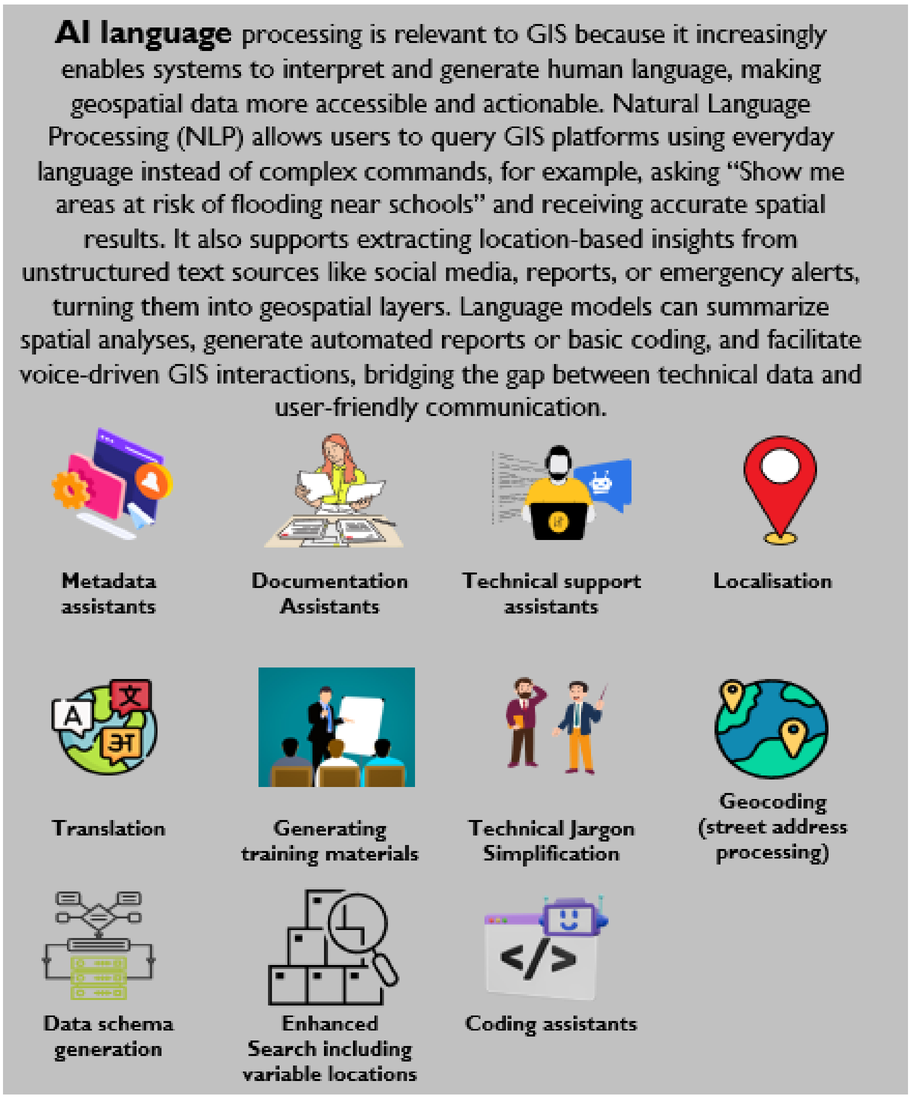
10. Do you want to increase the efficiency and accuracy of basic coding?

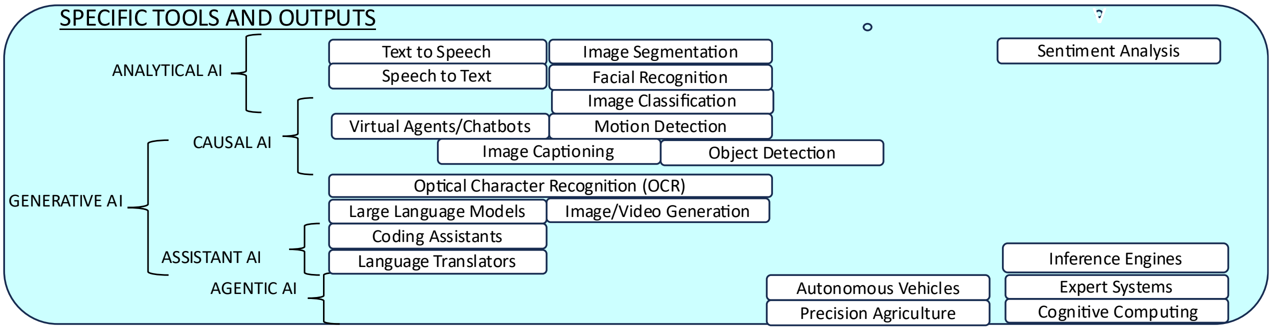
11. Do you want to be able to automate metadata, documentation, or specifications?

12. Do you want to be able to use or implement automated support / training assistants? (including for your own internal systems)

13. Do you want to be able to localise or translate components of systems or documentation?

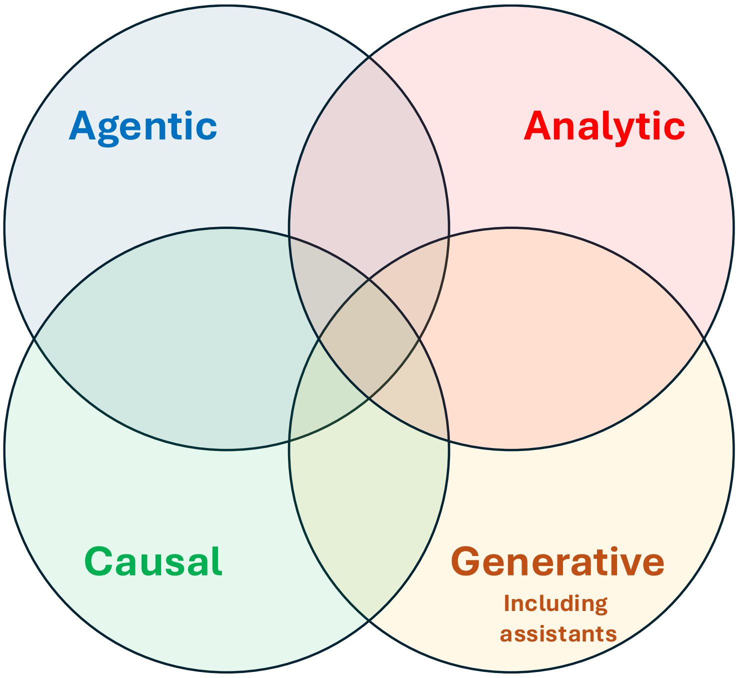
14. Do you want to be able to search across all of your databases and documents in natural language, including using approximate location references?





TOOL FAMILIES

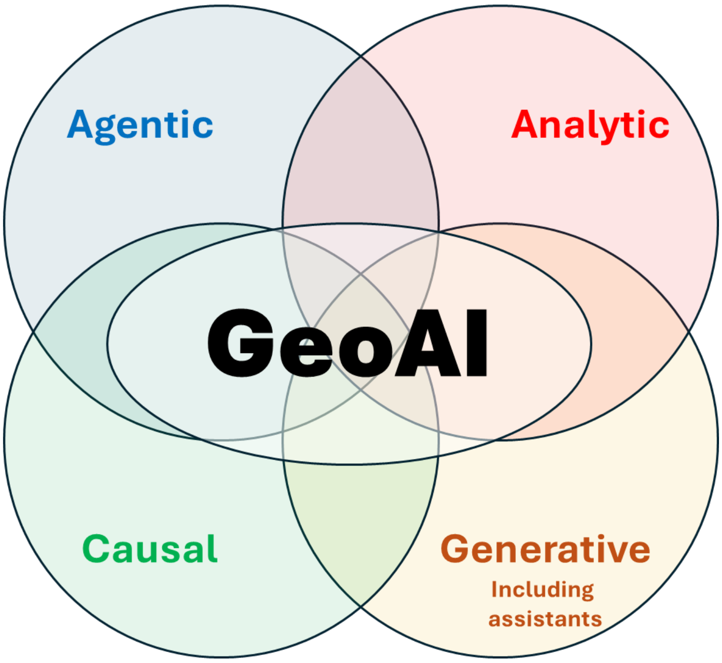
AI tool families represent **broad categories of AI systems** based on their **core capabilities and intended functions**. The four major families are Analytical AI, Generative AI (which includes 'AI assistants'), Causal AI, and Agentic AI. Individual tools often combine elements of more than one of these tool family characteristics.



Esri defines GeoAI as: "the application of artificial intelligence (AI) fused with geospatial data, science, and technology to accelerate real-world understanding of business opportunities, environmental impacts, and operational risks."

Different strands of GeoAI utilise different aspects of the four main AI tool families, and one GeoAI tool might utilise multiple aspects of multiple tool families at the same time.

The following section explains these tool families in more detail and explains how GeoAI relates to those key families.



Analytical AI

Analytical AI focuses on **understanding patterns, making predictions, and supporting decision-making** using data.

Key Characteristics:

- Uses statistical models and machine learning.
- Excels at classification, regression, and forecasting.
- Often used in business intelligence and diagnostics.

Geospatial Examples:

- **Predictive Traffic Modeling**: AI analyzes historical and real-time traffic data to forecast congestion and optimize routing.
- **Urban Growth Analysis**: Machine learning models detect patterns in satellite imagery to predict future urban expansion.
- **Disaster Impact Assessment**: AI processes geospatial data to estimate flood zones, wildfire spread, or earthquake damage.
- **Agricultural Yield Prediction**: Combining satellite imagery and weather data to forecast crop health and productivity.

Analytical AI is demonstrated in many areas across ArcGIS technology.

Generative AI

Generative AI is designed to **create new content,** text, images, audio, code, or video based on learned patterns.

Key Characteristics:

- Can generate human-like language, realistic images, or music.
- Often used in creative and content-driven applications.
- Uses models like GANs, VAEs, and Transformers.

Non-GIS specific Examples:

- ChatGPT (text generation)
- DALL·E (image generation)
- GitHub Copilot (code generation)
-

Geospatial Examples:

1. Automated Map Generation

Generative AI can create detailed and context-aware maps from textual descriptions or satellite imagery. For example:

- Generating land use maps from satellite images.
- Creating maps from user defined criteria stated in natural language.

2. Data Augmentation and Simulation

Generative models can simulate realistic geospatial scenarios, such as:

- Predicting urban growth or deforestation patterns.
- Generating synthetic satellite imagery for training other AI models.

3. Remote Sensing and Image Enhancement

Generative AI improves the quality of satellite imagery by:

- Filling in missing data (cloud removal).
- Enhancing resolution (super-resolution techniques).
- Detecting features like roads, buildings, or water bodies.

4. Scenario Planning and Decision Support

In urban planning or disaster response, generative models can simulate:

- Evacuation routes.
- Environmental impact assessments.

Resources:

- https://architecture.arcgis.com/en/overview/introduction-to-arcgis/geospatial-ai.htm
- https://www.esri.com/about/newsroom/publications/wherenext/nexttech-genai
- https://www.esri.com/content/dam/esrisites/en-us/media/products/arcgis-pro-issues-addressed/ai-assistant-pro.pdf
- https://www.esri.com/arcgis-blog/products/arcgis-online/geoai/whats-new-in-ai-assistants-october-2025
- https://www.esri.com/arcgis-blog/products/arcgis-pro/data-management/metadata-data-interoperabilitys-hidden-talent-part-one

- https://developers.arcgis.com/net/api-reference/ReadMe.html
- Introducing AI components (beta) in the ArcGIS Maps SDK for JavaScript

Causal AI

Causal AI aims to **understand cause-and-effect relationships**, not just correlations.

Key Characteristics:

- Goes beyond prediction to explain *why* something happens.
- Useful for decision-making under uncertainty.
- Often uses causal graphs, counterfactual reasoning, and interventions.

Geospatial Examples:
Causal AI in GIS (Geographic Information Systems) is used to go beyond mapping and spatial analysis by uncovering **cause-and-effect relationships** in geospatial data. This enables more informed decision-making in areas like urban planning, environmental management, disaster response, and public health.

Key Utilizations of Causal AI in GIS:

1. Environmental Impact Analysis

- **Use Case**: Understanding how deforestation affects local climate or biodiversity.
- **Causal AI Role**: Distinguishes between correlation (e.g., tree loss and temperature rise) and causation (e.g., tree loss *causes* temperature rise).

2. Urban Planning and Infrastructure

- **Use Case**: Evaluating how new roads or zoning laws affect traffic congestion or housing prices.
- **Causal AI Role**: Models the *impact* of interventions (e.g., building a new highway) on urban dynamics.

3. Disaster Risk and Resilience

- **Use Case**: Assessing how land use changes influence flood risk.
- **Causal AI Role**: Identifies causal links between human activity and increased vulnerability to natural disasters.

4. Public Health and Epidemiology

- **Use Case**: Studying how environmental factors (e.g., pollution, green space) affect disease spread or mental health.
- **Causal AI Role**: Helps isolate the true drivers of health outcomes from spatial data.

5. Agriculture and Food Security

- **Use Case**: Determining how irrigation practices or soil conditions affect crop yields.
- **Causal AI Role**: Supports precision agriculture by identifying which interventions lead to better outcomes.

How It Works in GIS Context:

- **Causal Graphs**: Represent spatial variables (e.g., land use, rainfall, elevation) and their causal relationships.
- **Counterfactual Analysis**: Answers questions like *"What if we had preserved this wetland?"*
- **Intervention Modelling**: Simulates the effects of policy or environmental changes on spatial outcomes.

Causal AI is demonstrated in many areas across ArcGIS technology.

Agentic AI

Agentic AI refers to systems that use an Agent to **autonomously plan, act, and adapt** to achieve goals in dynamic environments. AI agents are autonomous or semi autonomous systems capable of planning, reasoning, and carrying out tasks to achieve specified goals. They may orchestrate multiple AI models or tools to complete complex workflows.

Key Characteristics:

- Combines reasoning, memory, and action.
- Often uses reinforcement learning, planning algorithms, and multi-agent systems.

Can operate over long time horizons with minimal human input.

GeoAI

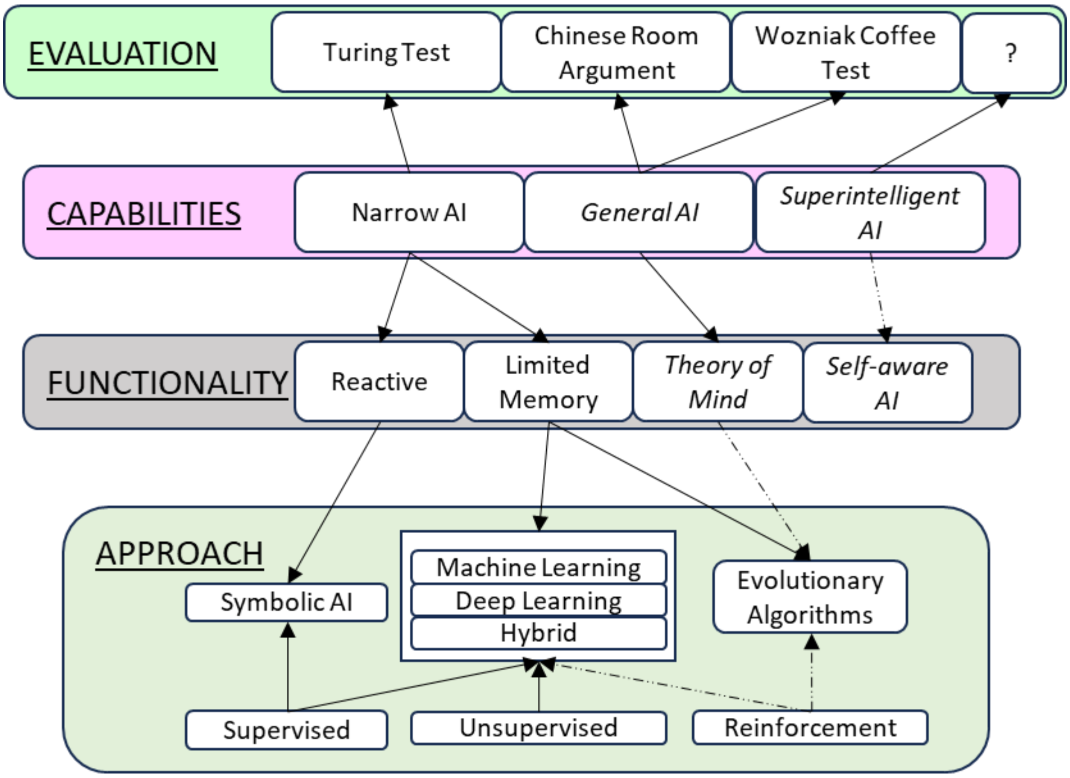
GeoAI refers to the integration of artificial intelligence techniques with geospatial data and GIS workflows. It combines spatial analysis with machine learning and deep learning to solve location-based problems. GeoAI can span or incorporate many of the preceding tool families of AI.

Geospatial examples:

- Automating building footprint extraction from aerial imagery.
- Predicting traffic congestion using historical and real-time sensor data.
- Detecting illegal deforestation from satellite images.
- Analysing social media posts for disaster response by extracting location information.

Resource:
https://pro.arcgis.com/en/pro-app/latest/help/analysis/ai/geoai.htm

PART 2: HIGH-LEVEL CONCEPTS



It is important for geospatial professionals to understand high-level AI concepts because it enables them to cut through the hype and distinguish today's practical, narrow AI capabilities—such as imagery classification and feature extraction—from speculative ideas like General AI or Superintelligent AI. Recognising that current systems lack human-level reasoning, Theory of Mind, and real-world common-sense skills (as illustrated by tests like the Wozniak Coffee Test) helps prevent unrealistic expectations about automation or autonomous spatial decision-making. Awareness of foundational approaches like symbolic AI and evolutionary algorithms also equips professionals to evaluate emerging tools, choose appropriate methods for spatial problems, and understand how hybrid systems combine rules, optimisation, and machine learning. Ultimately, this high-level literacy ensures better technology choices, clearer communication with stakeholders, and a grounded understanding of both the real and emerging capabilities of AI within geospatial workflows.

CAPABILITIES	Narrow AI	*General AI*	*Superintelligent AI*

AI CAPABILITIES

These categories reflect different levels of intelligence and autonomy in artificial systems. The diagram above shows the 3 most accepted types of intelligence.

Narrow AI (Weak AI)

Narrow AI refers to AI systems that are designed and trained for a specific task or a narrow range of tasks. These systems do not possess consciousness, self-awareness, or general intelligence.

General Examples:

- Siri, Alexa (voice assistants) and other *Chatbots*
- Google Translate
- Recommendation algorithms (Netflix, Amazon)
- Facial recognition of passports at an airport
- Car registration plate readers at car-parks

Geospatial Examples (if AI is used in place of manual processes):

- Remote sensing geographic feature extraction
- Predicting spatial phenomena like flood risk, crime hotspots, or disease outbreaks.
- Detecting illegal mining or fishing activities from satellite data.
- Extracting location information from unstructured text (e.g., social media, news).

Resources:

Artificial Intelligence in GIS: Promise, Progress, and Possibilities

General AI (AGI Artificial General Intelligence)

General AI refers to a theoretical form of AI that can understand, learn, and apply intelligence across a wide range of tasks much like a human being. It has yet to be demonstrated.

General Example:

A car parks licence plate recognition AI agent, has decided (of its own volition and without ever being programmed to do so) to do a spectral scan of car's exhaust gases and listen to the car noises to check for mechanical problems, and then contact the owner of the faulty car with a booking at a garage, where it has already ordered the correct parts to be delivered.

Geospatial Examples:

- An emergency management AI agent, detects heavy rain in an upstream area, so takes it upon itself to reach out to the electricity companies AI agent to close certain floodgates, and to the transport authority AI agent to close 7 specific roads.

Superintelligent AI

Superintelligent AI refers to a hypothetical AI that surpasses human intelligence in all aspects: creativity, problem-solving, emotional intelligence, and more. This has yet to be achieved.

AI FUNCTIONALITY

The functionality of AI refers to how an AI system operates and interacts with its environment. This is often categorized into four types based on increasing levels of cognitive sophistication:

Reactive Machines

These are the most basic types of AI systems. They can only respond to specific inputs with pre-programmed outputs.

Functionality:

- No memory or learning capability.
- Cannot use past experiences to inform current decisions.
- Operates purely in the present moment.

Example:

IBM's Deep Blue chess computer, which could evaluate millions of possible moves but had no memory of past games.

Geospatial Example:

- A GIS application that displays real-time traffic conditions operates by reacting to incoming sensor data or live feeds, such as congestion levels, and updating the map to reflect the current situation. It does not analyze historical trends or predict future traffic patterns; instead, it focuses solely on the present state. These systems are considered reactive because they lack internal memory or learning capabilities, responding only to immediate inputs without retaining past information. While some people might argue that this capability is not truly 'AI', nevertheless the AI characteristic of 'perception of the environment' is demonstrated by this capability.

Limited Memory

These AI systems can use past experiences or data to make better decisions in the present.

Functionality:

- Can store and use historical data temporarily.
- Most modern AI systems fall into this category.
- Enables learning from past actions to improve future performance.

Geospatial Examples:

- Retention of a GIS user's previous preferences and searches, and comparing those to other users, to optimise their next experiences with a system.
- Self-driving cars that observe other vehicles' speed and direction to make driving decisions.

Theory of Mind (Conceptual)

This is a theoretical stage of AI that would understand human emotions, beliefs, intentions, and social interactions.

Functionality:

- Capable of modelling the mental states of others.
- Could engage in meaningful social interactions.
- Would require a deep understanding of human psychology.

Status:
Still under research and development; not yet realized.

Self-Aware AI (Hypothetical)

This is the most advanced and speculative form of AI. It would possess consciousness, self-awareness, and a sense of identity.

Functionality:

- Understands its own internal states.
- Can make decisions based on self-reflection.
- Could potentially have desires, needs, and goals.

Status:
Purely theoretical and not currently achievable with existing technology.

Cognitive computing

Cognitive computing is an umbrella term for systems that mimic human problem-solving—combining perception (vision/speech), language, reasoning, and learning. Today it largely overlaps with AI assistants, translation, and search experiences. It emphasises human-centred design and assistive outcomes.

EVALUATION

AI evaluation tests are conceptual tools or thought experiments used to assess or critique the **intelligence, understanding, and capabilities** of artificial systems. Here's what they mean in the context of three well-known examples:

The Turing Test

Proposed by: Alan Turing (1950)

Purpose: To evaluate a machine's ability to exhibit intelligent behaviour indistinguishable from that of a human.

How it works:

- A human judge engages in a text-based conversation with both a human and a machine.
- If the judge cannot reliably tell which is which, the machine is said to have passed the test.

Significance:

- Focuses on **behavioural imitation**, not actual understanding or consciousness.

Criticized for being too focused on deception rather than true intelligence.

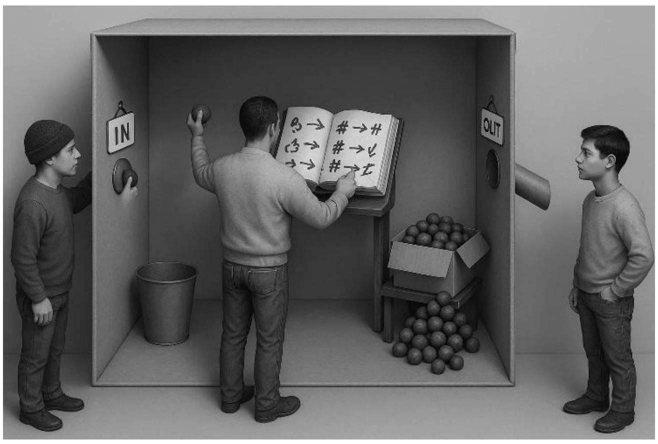
Chinese Room Argument

Proposed by: John Searle (1980)

Purpose: To challenge the idea that passing the Turing Test proves understanding or consciousness.

How it works:

- Imagine a person in a room who doesn't understand Chinese but follows a rulebook to manipulate Chinese symbols and produce appropriate responses. *The rulebook represents computer coding and data processing.*
- To outsiders, it appears this person understands Chinese, but internally, there is no comprehension.



Significance:

- Argues that **syntax (symbol manipulation)** is not the same as **semantics (understanding)**.
- Suggests that machines can simulate understanding without actually possessing it.

Wozniak Coffee Test

Proposed by: Steve Wozniak (co-founder of Apple)

Purpose: To test an AI's ability to function in the physical world using common sense.

How it works:

- The (robotic) AI is asked to enter any typical home and make a cup of coffee. *Where would the coffee be stored? What coffee making equipment is available? etc.*
- This involves recognizing objects, navigating space, and understanding human environments.
- Random situations instead of pre-prepared scenarios.

Significance:

- Emphasizes **embodied intelligence** and **real-world reasoning**.
- Highlights the gap between narrow AI and general intelligence.

Geospatial applicability:

For AI use-cases that have a spatial element, such as autonomous vehicles, the coffee test is an important concept to indicate that real world situations can be very random and are difficult to completely predict within controlled development scenarios.

The essence of the test is adaptability in unstructured environments.
Applying this concept to geospatial technology, we can draw parallels in terms of handling unpredictability and unknowns in spatial data:

1. Unpredictability in Satellite Imagery
Just as a robot faces unknown kitchen layouts, geospatial systems encounter:
- Unexpected objects (e.g., new construction, natural disasters, temporary structures).
- Unusual patterns (e.g., crop anomalies, unregistered roads).
- Unknown features (e.g., unidentified vessels, debris after storms).

A robust geospatial system should adaptively interpret and classify these anomalies without relying solely on pre-trained patterns.

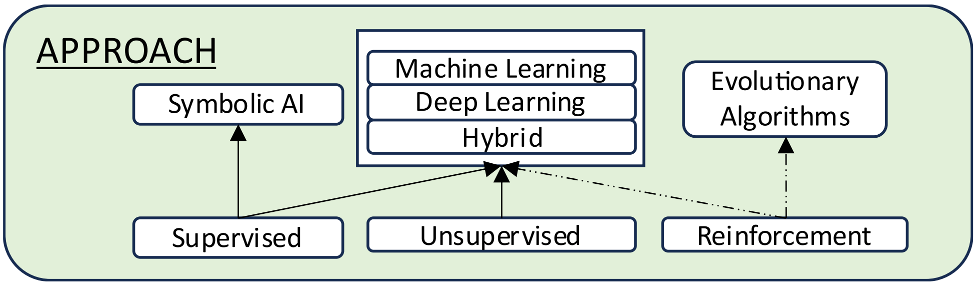
2. Dynamic Context Awareness
In the coffee test, the robot must infer context (where is the coffee, water, kettle?). Similarly, geospatial AI must:
- Infer contextual meaning of objects (e.g., is that a shadow or a new building?).
- Understand temporal changes (e.g., seasonal vegetation vs. deforestation).
- Handle multi-source data fusion (satellite, drone, IoT sensors).

3. Real-World Complexity
The real world is messy:
- Occlusions (cloud cover, shadows).
- Resolution variability (different sensors, angles).
- Data gaps (missing metadata, incomplete maps).

Applying the Wozniak principle means designing systems that don't fail when conditions deviate from the training set they should reason, hypothesize, and adapt.

4. Practical Applications
- Disaster Response: Detecting unexpected debris or temporary shelters after an earthquake.
- Urban Planning: Identifying informal settlements that don't match standard patterns.
- Environmental Monitoring: Spotting invasive species or illegal logging in unpredictable terrains.

5. Key Design Principles
- Adaptive AI models: Incorporate anomaly detection and unsupervised learning.
- Human-in-the-loop workflows: For ambiguous cases, enabling expert review.
- Continuous learning: Systems that update models as new patterns emerge.



AI APPROACH

AI training approaches refer to the **methods and paradigms** used to develop intelligent systems. These approaches vary in how they represent knowledge, learn from data, and adapt to environments. Here's a breakdown of the major types:

Symbolic AI

Symbolic AI refers to a type of artificial intelligence that uses clearly defined rules, logic, and symbols to represent knowledge and make decisions. Rather than learning from data like machine learning systems, symbolic AI relies on human-created models and relationships such as spatial rules, classifications, and hierarchies.

Symbolic AI can be useful when transparency and explainability are important, as the logic behind decisions can be easily traced and understood. Transparency and explainability are increasingly important topics for the use of AI, where some newer AI systems are 'black boxes' meaning that users cannot see what data and algorithms are being used to generate the outputs. In certain contexts such as legal or government scenarios it can be crucial that there is accountability around how and why certain decisions have been made, so this is a topic of much debate.

Approach: Uses **explicit rules and logic** (e.g. IF-THEN statements) to represent knowledge and make decisions.

Key Features:

- Based on human-readable symbols and rules.
- Requires manual programming of knowledge.
- Great for structured problems (e.g., expert systems).

Geospatial Examples:
In GIS, symbolic AI can be used to:

- Apply spatial rules (e.g., "rivers flow downhill" or "buildings must be within zoning boundaries").
- Automate reasoning based on geographic relationships (e.g., identifying land parcels that meet certain criteria).
- Support expert systems that use structured knowledge to guide planning or analysis.

Resources:

https://www.sciencedirect.com/science/article/pii/S1569843225000159

Machine Learning

Machine Learning is a type of AI where computers learn from data instead of being directly programmed. In GIS, machine learning can be used to automatically classify land cover from satellite images, detect changes in terrain, or predict traffic patterns based on historical data. The system learns by spotting patterns in examples and then applies that knowledge to new data. For example, if an image processing algorithm has never encountered a type of object before (maybe a lighthouse) it will store that object and group it with other similar objects (even if it doesn't know yet that those objects are called lighthouses).

Approach:
Enables systems to **learn patterns from data** without being explicitly programmed.

Key Features:

- Data-driven.
- Improves performance with more data.
- Includes supervised, unsupervised, and reinforcement learning.

Example:

- Spam email detection based on patterns in email content.

Geospatial Examples:
Classifying satellite imagery into feature categories like forest, urban, water, agriculture. If training data is trained to recognise the characteristics of these land uses (natural colours, geometries, multi-spectral signatures, coincident or logically non-coincident features etc) then it is possible to automatically update the trained models for those features so that subsequent automated classifications are more accurate.

In ArcGIS:
Machine learning has long been part of ArcGIS. It supports classification, prediction, clustering, and forecasting through algorithms like K-Nearest Neighbour, Support Vector Machine, decision tree ensembles, logistic regression, and MaxEnt. Tools such as Forest-Based and Boosted Classification and Regression handle tabular and spatial data, while clustering options include Multivariate Clustering, Density-Based Clustering, and Build Balanced Zones. ArcGIS also offers global regression models, time-series forecasting, and causal inference analysis to uncover relationships between variables. Unlike traditional ML, spatial machine learning incorporates location, shape, proximity, density into its algorithms, using methods like spatial autoregression and geographically weighted regression. Choosing the right algorithm depends on your data, problem, and goals.

- https://www.esri.com/content/dam/esrisites/en-us/esri-press/book-pages/sample-page/geoai-artificial-intelligence-in-gis.pdf
- https://www.esri.com/arcgis-blog/products/arcgis-pro/analytics/machine-learning-spatial-components-in-arcgis-pro
- https://resource.esriuk.com/blog/machine-learning-2025/?adusf=linkedin&aduc=esri+uk&adut=0179f03e-6b90-46ba-b049-375952c7938b
- https://www.esri.com/arcgis-blog/products/analytics/analytics/machine-learning-in-arcgis

Deep Learning

Deep Learning is a more advanced form of machine learning that uses layered networks (called neural networks) to understand complex data especially images and sounds. Deep Learning progressively uses multiple layers of information to extract higher-level features from the raw input. It uses more advanced methods and helps solve complex problems across large data volumes, with a focus on automated data extraction and pattern recognition. In GIS, deep learning is often used for tasks like identifying buildings, roads, or vegetation in aerial imagery with high accuracy. It's particularly good at handling large and detailed datasets, such as those from drones or high-resolution satellites.

Approach:
A subset of ML that uses **artificial neural networks** with many layers to model complex patterns.

Key Features:

- Excels at processing unstructured data (images, audio, text).
- Requires large datasets and computational power.
- Learns hierarchical representations.

Example:

- Image recognition in self-driving cars.

In ArcGIS:
Deep learning is widely available across the ArcGIS ecosystem, from ArcGIS Online for testing pretrained models to the ArcGIS API for Python for custom workflows. Pretrained GeoAI models save time by automating feature extraction from imagery, point clouds, and text tasks like digitizing building footprints or creating land-cover maps. These models, accessible via ArcGIS Living Atlas and other repositories, eliminate the need for extensive training and resources. Over 100 models exist today, supporting image feature detection, pixel and point cloud classification, and image redaction, making deep learning more accessible than ever.

- https://pro.arcgis.com/en/pro-app/latest/help/analysis/deep-learning/what-is-deep-learning-.htm
- https://www.arcgis.com/home/item.html?id=db4ccd9a286a471d8b937f79d88e96a3
- https://www.esri.com/content/dam/esrisites/en-us/esri-press/book-pages/sample-page/geoai-artificial-intelligence-in-gis.pdf
- https://www.esri.com/content/dam/esrisites/en-us/about/events/media/UC-2019/technical-workshops/tw-9271-1014.pdf
- https://www.esri.com/about/newsroom/arcwatch/where-deep-learning-meets-gis

Hybrid

Approach:
Combines **symbolic reasoning** with **machine learning** to leverage the strengths of both.

Key Features:

- Balances interpretability and adaptability.
- Useful in complex domains requiring both logic and learning.

Example:
An AI assistant that uses rules for scheduling and ML for understanding natural language.

Evolutionary Algorithms

Approach:
Inspired by **biological evolution**, these algorithms use mechanisms like mutation, crossover, and selection to evolve solutions.

Key Features:

- Good for optimization problems.
- Does not require gradient-based learning.
- Can explore large, complex search spaces.

Geospatial Example:
Evolutionary algorithms are being increasingly applied to address climate change challenges. These algorithms can help optimize complex systems, model climate phenomena, and design strategies for climate mitigation and adaptation.

Supervised

Approach:
Learns from **labelled data** (input-output pairs).

Key Features:

- Predicts outcomes based on past examples.
- Requires a large amount of labelled data.

Geospatial Example:

- Predicting house prices from features like size and location.

https://pro.arcgis.com/en/pro-app/latest/help/analysis/image-analyst/overview-of-image-classification.htm

Unsupervised

Approach:
Finds **patterns or structures** in **unlabelled data**.

Key Features:

- Useful for clustering, dimensionality reduction.
- No explicit output labels.

Geospatial Example:

- Customer segmentation in marketing.

Resource:
https://pro.arcgis.com/en/pro-app/latest/help/analysis/image-analyst/overview-of-image-classification.htm

Reinforcement

Approach:
Learns by **interacting with an environment** and receiving feedback in the form of rewards or penalties.

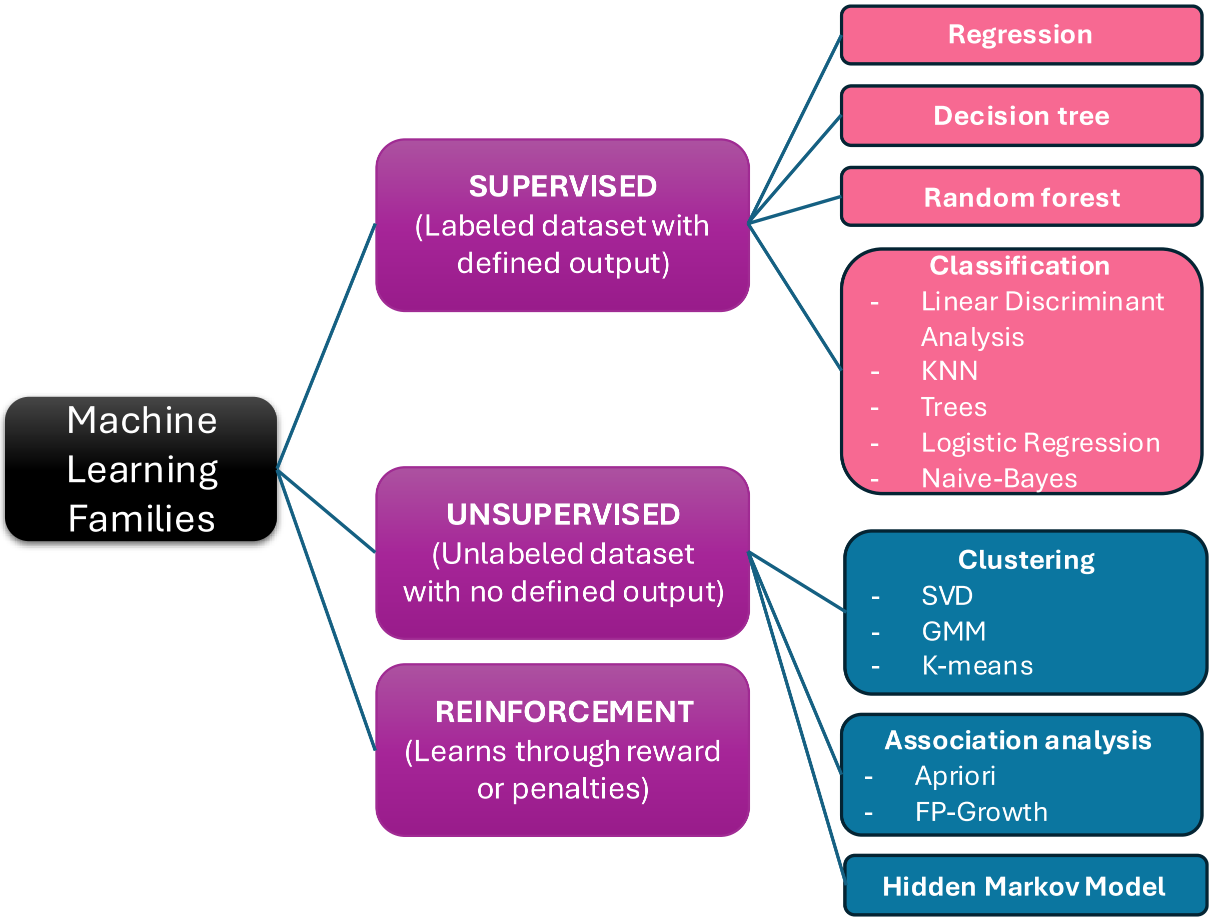
Key Features:

- Focuses on sequential decision-making.
- Balances exploration and exploitation.

Example:
Training a robot to walk or an AI agent to play video games.

Deep Dive: Classes of Algorithms Used in Machine Learning

The following section summarizes common machine learning algorithms, their descriptions, and whether they are used within in ArcGIS. This section may be relevant mostly to developers or others with an interest in the technical detail of how AI algorithms operate.



Algorithm	Category	Brief Description	ArcGIS Support	ArcGIS Reference
Regression	Supervised	Predicts continuous values from input features.	Yes Generalized Linear Regression; Forest-based Classification and Regression.	Regression Analysis: https://pro.arcgis.com/en/pro-app/3.4/tool-reference/spatial-statistics/regression-analysis-basics.htm Forest-based Classification and Regression: https://pro.arcgis.com/en/pro-app/latest/tool-reference/big-data-analytics/forest-based-classification-and-regression.htm
Decision Tree	Supervised	Splits data into branches based on feature values.	Yes Forest-based Classification and Regression.	Forest-based Classification and Regression: https://pro.arcgis.com/en/pro-app/latest/tool-reference/geoai/how-decision-tree-classification-and-regression-works.htm
Random Forest	Supervised	Ensemble of decision trees for robust predictions.	Yes Forest-based Classification and Regression.	Forest-based Classification and Regression: https://pro.arcgis.com/en/pro-app/latest/tool-reference/geoai/how-random-trees-classification-and-regression-works.htm
KNN (k-Nearest Neighbor)	Supervised (Classification)	Classifies based on closest labeled neighbors.	Yes Image Analyst classification workflows.	Image Classification Overview: https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/train-k-nearest-neighbor-classifier.htm
Logistic Regression	Supervised (Classification)	Predicts categorical outcomes using logistic model.	Yes Logistic Regression tool.	Logistic Regression: https://pro.arcgis.com/en/pro-app/3.4/tool-reference/spatial-statistics/how-glr-works.htm
Naïve Bayes	Supervised (Classification)	Probabilistic classifier using Bayes' theorem.	Yes Image Analyst classification workflows.	Image Classification Overview: https://www.esri.com/content/dam/esrisites/en-us/about/events/media/UC-2019/technical-workshops/tw-6165-494.pdf
SVM (Support Vector Machine)	Supervised (Classification)	Finds maximum-margin decision boundaries.	Yes SVM classification in Image Analyst.	Support Vector Machine: https://pro.arcgis.com/en/pro-app/latest/tool-reference/image-analyst/train-support-vector-machine-classifier.htm
SVD (Singular Value Decomposition)	Unsupervised	Matrix factorization for dimensionality reduction.	Not directly exposed; related functionality via PCA.	
Principal Component Analysis (PCA)	Unsupervised	Reduces dimensionality by transforming correlated variables.	Yes Principal Components tool.	Principal Components: https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/principal-components.htm
Covariance Matrix (in machine learning)		A matrix that shows how pairs of variables vary together across a dataset; it is widely used in statistical modelling, dimensionality reduction (e.g., PCA), and understanding relationships or correlations between features.		
K-means	Unsupervised (Clustering)	Groups data into clusters based on similarity.	Yes Spatially Constrained Multivariate Clustering.	K-means Clustering: https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/spatially-constrained-multivariate-clustering.htm

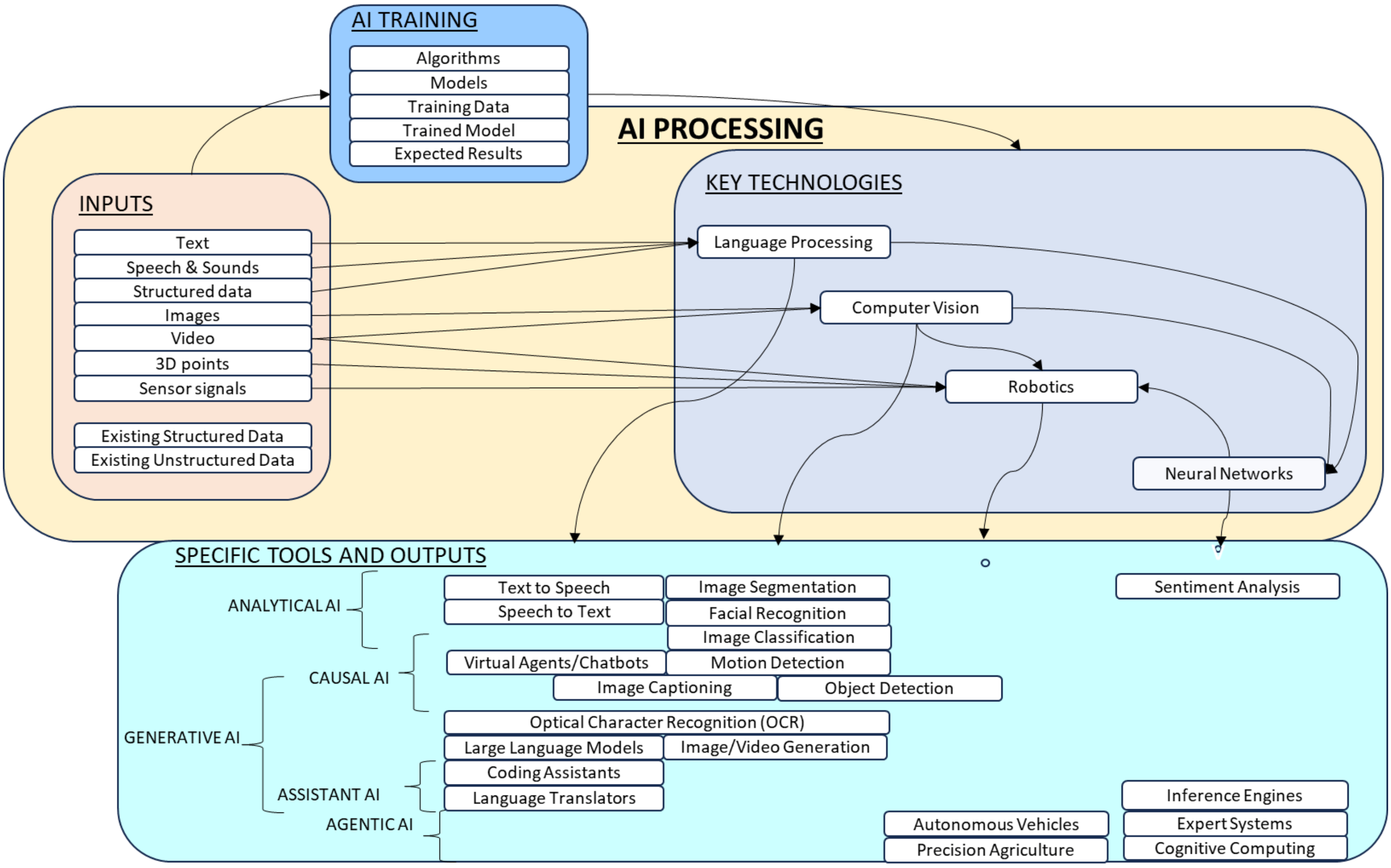
Algorithm	Category	Brief Description	ArcGIS Support	ArcGIS Reference
Affinity propagation		Affinity propagation is a clustering algorithm that identifies “exemplar” points by passing messages between data points until stable clusters emerge. Unlike k-means, it does not require the number of clusters to be specified ahead of time. It evaluates similarity between all pairs, selecting representative examples.		
Apriori / FP-Growth	Unsupervised (Association Analysis)	Finds frequent patterns or associations.	Not natively available in ArcGIS.	https://pro.arcgis.com/en/pro-app/3.3/help/data/parcel-editing/aboutmeasurementaccuracy.htm
Markov Model or Hidden Markov Model (HMM)	Probabilistic Sequence Model	Models sequential data with hidden states.	Not directly mentioned in ArcGIS websites.	https://github.com/simonscheider/mapmatching

Deep Dive: Detailed Terminology Used in Machine Learning

Defined Term	Brief Description	GIS Relevance	ArcGIS Reference
Reinforcement Learning (RL)	Learns actions via rewards and penalties.		Not specifically implemented in ArcGIS core tools.
Linear processing	Data being handled in a fixed, sequential order, where each step depends directly on the previous one.	Traditional geoprocessing workflows (e.g. buffering → intersect → summarise) execute steps linearly in a model or script.	https://pro.arcgis.com/en/pro-app/latest/help/analysis/geoprocessing/modelbuilder.htm
Tensor	A tensor is a multi-dimensional numerical array used to represent data in machine-learning models. Tensors generalise scalars, vectors, and matrices into higher dimensions, enabling efficient computation on GPUs. Deep-learning frameworks rely on tensor operations for model training and inference.	GIS raster imagery is often converted into tensors when used as input for ArcGIS deep-learning workflows.	https://pro.arcgis.com/en/pro-app/latest/help/analysis/image-analyst/deep-learning-models-in-arcgis.htm
Heuristic	A heuristic is a rule-of-thumb strategy (short-cut) or simplified decision process that helps an algorithm reach a good (though not necessarily optimal) solution efficiently. Heuristics are valuable when exhaustive search is computationally expensive.	ArcGIS Network Analyst uses heuristic techniques—derived from Dijkstra-based optimisations—to speed up route planning, such as hierarchical routing for nationwide networks.	https://support.esri.com/en-us/gis-dictionary/heuristic
Bayesian inference	Bayesian inference is a statistical reasoning framework that updates the probability of a hypothesis as new data becomes available. It is foundational in probabilistic machine learning and uncertainty quantification.	In ArcGIS geostatistics, Bayesian techniques (e.g., Empirical Bayesian Kriging) continually update spatial predictions based on observed sample points and modelled uncertainty.	https://www.esri.com/about/newsroom/app/uploads/2018/12/Machine-Learning-in-ArcGIS.pdf
Cost function	A cost function (or loss function) quantifies the error between a model's predictions and the true values. During training, AI models minimise this cost via optimisation algorithms, adjusting internal parameters to improve accuracy. Different cost functions suit different tasks—for example, cross-entropy for classification or mean-squared error for regression.	In a GIS least-cost path analysis, cost surfaces operate analogously: the system computes the lowest cumulative "cost" (travel difficulty, slope, friction) between locations—conceptually similar to minimising a model's error over an optimisation landscape.	https://pro.arcgis.com/en/pro-app/latest/help/analysis/raster-functions/cost-path-function.htm
One-hot encoding	One-hot encoding converts categorical variables into binary indicator fields—one column per category, with a 1 marking presence and 0 for absence. This prevents algorithms from assuming false numerical order in category labels.	In ArcGIS Pro, a user can encode land-use categories (e.g., residential, industrial, commercial) into separate binary fields to prepare data for	https://pro.arcgis.com/en/pro-app/latest/tool-reference/data-management/encode-field.htm

Defined Term	Brief Description	GIS Relevance	ArcGIS Reference
		regression-based suitability modelling.	
Covariance matrix	In AI and statistics, a covariance matrix is a square table showing the covariances between pairs of variables. Covariance measures how two variables change together, so the matrix summarises relationships across an entire dataset. It is fundamental in multivariate modelling, dimensionality reduction, and machine learning algorithms such as PCA. Covariance matrices help identify correlated features, stability issues, or redundancy in data, which can guide feature engineering.	ArcGIS uses covariance matrices when analysing multiple raster bands, such as in multiband imagery classification or geostatistical modelling.	https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/how-band-collection-statistics-works.htm
Chain-of-thought Prompting	Chain-of-thought prompting provides a model with examples of step-by-step reasoning, enabling large language models to break problems into intermediate steps and achieve more accurate multistep reasoning.		
Context window	A context window is the maximum amount of text (measured in tokens) an AI model can consider at once. It functions as the model's working memory and determines how much prior conversation or data the system can "remember" while generating responses.		
Computational Learning Theory (CLT)	Computational Learning Theory analyses the mathematical foundations of machine learning—studying learnability, sample complexity, algorithm efficiency, and error bounds. It provides theoretical guarantees about what concepts can be learned under given constraints.		
Latent space	A latent space is a compressed, lower-dimensional representation that captures the essential structure of complex data. Machine-learning models, especially deep and generative models, use latent spaces to encode hidden patterns and relationships not immediately visible in the raw data.		
Approximate nearest neighbour	Approximate-nearest-neighbour search retrieves points that are close to a query point without exhaustively scanning all data. It trades perfect accuracy for dramatic speed-ups, especially valuable in high-dimensional spaces. ANN methods power recommendation systems, vector search, and fast similarity queries.		

PART 4: HOW DOES AI WORK?



The concepts in the image above represent common AI domain terms that describe both how AI systems are built and the kinds of tools they produce. The upper sections outline the inputs, training, and processing technologies—such as algorithms, models, language processing, computer vision, robotics, and neural networks—which are the foundational steps involved in creating any AI capability. These elements explain where AI outputs come from and help demystify the process behind modern systems.

The lower section lists the specific tools and applications that emerge from these technologies, including speech-to-text, facial recognition, chatbots, image segmentation, large language models, and autonomous systems. These represent the practical, user-facing capabilities that organisations adopt. Overall, the terms in the diagram highlight the full lifecycle of AI—from raw data, through core processing technologies, to the final tools—helping geospatial professionals understand how AI works and what it can realistically achieve.

These terms (and others) are defined in the following section.

AI TRAINING

AI training is the process of teaching a computer system to recognise patterns and make decisions using example data. In GIS, this might involve showing the system labelled maps or images so it can learn to identify features like buildings or vegetation. Once trained, the system can analyse new data and produce useful results automatically.

Geospatial Example: • Training models in ArcGIS to recognise features in satellite imagery.

In the context of **AI training approaches**, the following terms describe the **core components and workflow** of how AI systems are developed and refined:

Algorithms

These are sets of instructions or rules that tell a computer how to solve a problem or perform a task. In GIS and AI, algorithms help process spatial data, find patterns, or make predictions like identifying land use from satellite images.

Algorithms are the **step-by-step procedures or mathematical formulas** that guide how an AI system learns from data.

Role in AI Training:

- Define how the model updates itself based on input data.

Examples include decision trees, gradient descent, and neural network backpropagation.

Labeled Data

Labelled data consists of **input-output pairs** where the correct answer (label) is provided for each example.

Role in AI Training:

- Essential for **supervised learning**.
- Helps the model learn the relationship between inputs and desired outputs.

Geospatial Example: • A satellite image of a building labelled as "building."

Models

A model is a computer program built using algorithms and data. It's designed to understand patterns and make decisions. A model is the **mathematical structure or architecture** that makes predictions or decisions based on input data.

Role in AI Training:

- The model is shaped and adjusted during training.

Examples include linear regression models, decision trees, and deep neural networks.

Geospatial Example:

- A model might predict flood risk based on terrain, rainfall, and historical data.

Multimodal Model

A multimodal model processes and relates information across more than one modality (e.g., text + imagery). In practice, *vision–language* models can describe, classify, or segment images based on natural-language prompts, and can also generate text grounded in pixels. These models support zero-shot and few-shot behaviours, reducing the need for task-specific training. They are increasingly used to query imagery semantically.

Geospatial Example:
ArcGIS now integrates **vision–language multimodal models** that allow a user to upload an aerial or satellite image and ask natural-language questions such as *"What do you see?"* or *"Segment the lake."* This uses paired image + text understanding to perform GIS tasks like segmentation and classification.
Resources:
https://www.esri.com/arcgis-blog/products/arcgis-pro/geoai/vision-language-models-geospatial-analysis

Training Data

This is the information used to teach an AI model how to do its job. It includes examples the model learns from. Training data is the information used to teach an AI model how to make decisions or predictions. Think of it like examples in a textbook: the AI studies these examples to learn patterns, relationships, and rules. The better and more representative the training data, the smarter and more accurate the AI becomes.

Imagine you want to teach a child to recognize different fruits. You show them **lots of pictures of apples, bananas, and oranges**, and tell them the name of each fruit. Over time, the child learns to identify fruits they've never seen before because they've learned patterns (colour, shape, size) from the examples. In AI terms, those pictures and labels are the **training data**.

Role in AI Training:

- The foundation for learning patterns and relationships.
- Can be labelled (for supervised learning) or unlabelled (for unsupervised learning).

Geospatial Example:
In GIS, training data might include labelled maps, sensor readings, or aerial photos that show known features like buildings or vegetation. AI training data in remote sensing is a dataset

used for land cover classification, where satellite imagery (such as from Sentinel-2 or Landsat 8) is paired with labelled data indicating land types like forest, urban, water, or agriculture. Each pixel in the imagery contains spectral information across multiple bands, and the corresponding labels serve as ground truth for supervised learning. This data enables machine learning modelslike Random Forests or Convolutional Neural Networksto learn patterns and classify new, unlabelled images. Public datasets like EuroSAT and BigEarthNet are commonly used for this purpose.

In **ArcGIS**, training data is often used in **deep learning workflows** for tasks like **object detection, image classification, or feature extraction** from satellite imagery. For example:

- If you want to detect **buildings in aerial imagery**, you need training data that includes:
  - **Images** (e.g., aerial photos)
  - **Labels** (e.g., polygons marking where buildings are)

ArcGIS tools like **"Export Training Data for Deep Learning"** prepare this data so it can be fed into AI models such as those in **ArcGIS Pro** or **ArcGIS Image Analyst**.

Reference Links:

- ArcGIS Pro: Export Training Data for Deep Learning
- ArcGIS Pro: Deep Learning in ArcGIS Pro

Trained Model

Definition:

A trained model is the **final product** after the model has been exposed to training data and optimized using an algorithm. Once a model has learned from training data, it becomes a trained model. This means it can now analyse new data and make predictions or classificationssuch as spotting damaged roads from drone footage.

Role in AI Training:

- Ready to make predictions or perform tasks on new, unseen data.
- Its performance depends on the quality of training data and the learning algorithm.

Reference Links:
https://www.esri.com/arcgis-blog/products/arcgis-pro/imagery/deep-learning-with-arcgis-pro-tips-tricks-part-2

Probabilistic

AI is considered **probabilistic** because it doesn't always give fixed or exact answersinstead, it makes decisions based on **likelihoods and patterns** learned from data.

Rather than following strict rules, AI models (especially those using machine learning) learn from examples and then estimate the probability that a certain outcome is correct. For instance, when an AI recognises an object in an image, it might say, "There's an 85% chance this is a building." It's not 100% certain, but it's making a best guess based on what it has seen before.

This probabilistic nature is especially important when dealing with uncertain, noisy, or incomplete data, or random situations (in AI randomness is often referred to using the term '**stochastic**').

Stochastic

Stochastic refers to processes that involve randomness, meaning the outcome can vary even with the same starting conditions.

This is common in real-world applications like:

- **Natural language processing** (understanding varied human speech or writing)
- **Image recognition** (handling blurry or complex visuals)
- **Geospatial analysis** (interpreting satellite imagery or predicting movement)

In short, AI doesn't "know" things in a rigid wayit **estimates**, **predicts**, and **adapts**, which makes it powerful but also means it can make mistakes or change its output depending on the context.

Entropy

In AI, entropy measures uncertainty or randomness in a distribution. High entropy indicates unpredictability, while low entropy suggests a more ordered or certain state. It is used in decision trees, information theory, and probabilistic models.

Geospatial Examples:
ArcGIS uses entropy to classify spatial relationships—for example, calculating local entropy to analyse how predictably one variable explains another.

Esri link:
https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/localbivariaterelationships.htm

Fuzzy logic

Fuzzy logic extends classical Boolean logic by allowing partial membership in a set rather than crisp boundaries. It captures uncertainty and gradual transitions, making it suitable for real-world phenomena that do not fit binary classification. AI systems use fuzzy logic for reasoning, decision-making, and modelling imprecise concepts. Fuzzy sets allow values to lie on a continuum from 0 to 1, representing degrees of truth.

Geospatial Examples:
ArcGIS provides fuzzy membership and fuzzy overlay tools to evaluate suitability surfaces where class boundaries (e.g., "steep", "near", "high risk") are gradual rather than absolute.

Esri link:
https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/applying-fuzzy-logic-to-overlay-rasters.htm

Temperature (in AI models)

A parameter that controls the randomness of text generation; **lower temperature** produces more predictable, deterministic output, while **higher temperature** increases creativity and variability by allowing the model to choose less probable words.

Hugging Face

Hugging Face is an open-source platform and community that provides tools, models, and libraries for natural language processing (NLP) and other AI tasks. It's widely used for sharing and deploying pre-trained models.

Geospatial examples:

- Using Hugging Face's language models to extract location names from environmental reports.
- Applying text summarisation models to condense lengthy planning documents into key spatial insights.
- Leveraging multilingual models to process geospatial data from international sources.

In-Context Learning (ICL)

In-context learning is the ability of large language models to adapt behaviour based on examples provided within a prompt, without retraining.

Geospatial example:
ArcGIS assistants and copilot-style tools use ICL to adapt responses based on user-supplied spatial context or examples during a session.

Ensemble Learning

Ensemble learning combines multiple models—such as decision trees, neural networks, or regressors—to improve predictive performance and reduce overfitting.

Geospatial Examples:
A GIS analyst may use AutoML or ArcGIS-supported machine-learning pipelines to build ensembles that predict landslide susceptibility, combining outputs of several spatial models for higher accuracy.

Esri link:
https://pro.arcgis.com/en/pro-app/latest/tool-reference/geoai/how-automl-works.htm

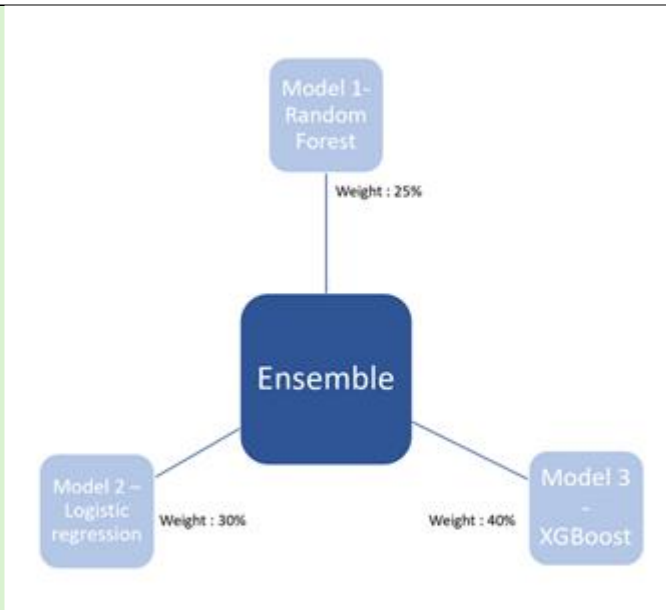


Figure 2- Source: https://pro.arcgis.com/en/pro-app/latest/tool-reference/geoai/how-automl-works.htm

Deep Dive: Detailed Terminology Used in Training

Defined Term	Brief Description	GIS Relevance	ArcGIS Reference
Feature engineering	Feature engineering is the craft of creating, transforming, and selecting variables that best represent informative patterns for machine-learning models. Effective feature engineering improves model accuracy, reduces complexity, and enhances interpretability.	Using ArcGIS Notebooks, an analyst might engineer new spatial features—such as distance-to-road, elevation gradients, or land-use diversity metrics—to improve a predictive model for retail-site selection or accident risk.	https://mediaspace.esri.com/media/t/1_c4bzw6hm
Data leakage	In AI, data leakage occurs when information from outside the training dataset improperly influences a model, causing overly optimistic performance during training but poor accuracy in real-world use. Leakage can happen through shared attributes, temporal overlap, or incorrect preprocessing. Preventing leakage is essential for trustworthy, generalisable models.	In GIS workflows, leakage may occur if a predictive model inadvertently uses future or confidential spatial data layers during training, leading to misleading map based predictions.	
Dendrogram	A dendrogram is a tree like diagram used in hierarchical clustering to illustrate how data points merge into clusters at different distance thresholds. It visually represents similarity relationships between items. In AI, dendrograms help understand cluster	ArcGIS uses dendrograms to show how spectral classes merge in multiband imagery during supervised classification.	https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/dendrogram.htm

	structure and determine optimal grouping.		
Residual network	A residual network is a deep-learning architecture that uses skip connections to allow gradients to flow more easily through very deep networks. This design mitigates vanishing-gradient issues, enabling models with dozens or hundreds of layers to train effectively. ResNets are widely used in image classification and feature extraction.	ArcGIS supports pretrained ResNet based models (ResNet 18, ResNet 34, etc.) for feature classification in imagery workflows, such as identifying building types or vegetation categories using deep learning toolsets	https://www.gichd.org/fileadmin/uploads/gichd/migration/fileadmin/GICHD-resources/info-documents/2024_GICHD_Innovation_Session_on_AI_for_Mine_Action/8._Machine_Learning_and_Neural_Networks.pdf
Cold start	The cold-start problem occurs when a system lacks sufficient data about new users or new items, hindering accurate predictions. It is common in recommendation systems and machine-learning pipelines reliant on historical patterns. Solutions include metadata enrichment, transfer learning, or synthetic data.		
Early stopping	Early stopping is a regularisation strategy that halts model training once validation performance stops improving. It prevents overfitting by stopping training at the optimal point before the model memorises noise. This technique reduces computation while improving generalisation.		
Corpus	A corpus is a large, structured collection of texts used to train, fine tune, or evaluate NLP models. Domain specific corpora (e.g., incident reports) improve performance on in scope tasks.		
Backpropagation	Backpropagation is a fundamental algorithm in machine learning and neural networks used to train models by efficiently adjusting their internal weights to reduce prediction error. It works by first computing the model's output and error, then propagating that error backwards through the network using calculus (specifically the chain rule) to calculate how much each weight contributed to the error. These gradients are then used by an optimisation method such as gradient descent to update the weights.	Backpropagation enables neural networks to learn from data, scales well to deep architectures, and underpins most modern AI applications, including image recognition, natural language processing, and speech recognition.	

Exploding gradient	An exploding gradient occurs when the values computed during backpropagation become excessively large, causing unstable or diverging training. It often affects deep neural networks, leading to numerical overflow, erratic weight updates, and failure to converge.		
Forward propagation	Forward propagation is the process by which input data flows through the layers of a neural network to produce an output. Each layer applies weights, biases, and activation functions, generating progressively richer representations until the final prediction is produced.		
Forward chaining	Forward chaining is a data-driven reasoning method used in rule-based AI systems. It starts with known facts and repeatedly applies inference rules to derive new information until a conclusion is reached. It is often used in expert systems that must update conclusions dynamically as new data arrives.		
Backward chaining	Backward chaining is a goal-driven inference method that starts with a hypothesis and works backward to determine which facts must be true to support that goal. It is used extensively in expert systems and logic programming environments.		

Evaluation & Reliability

Expected Results

These are the outcomes we hope to get from the trained model. Expected results are the **desired outputs or behaviors** the AI system should produce when given new data.

Geospatial Example:

- In GIS, expected results might include accurate land cover map updates when compared to a known accurate reference dataset.

Resource:
https://www.esri.com/arcgis-blog/products/arcgis-pro/imagery/performing-feature-extraction-classification-using-deep-learning-with-arcgis-pro

Role in AI Training:

- Used to evaluate the model's accuracy and effectiveness.
- Helps in tuning and improving the model.

Validation data

Validation data are held-out samples not used for training; they help tune hyperparameters and estimate model generalisation. Clear separation between train/validation/test reduces overfitting. For geospatial imagery, validation often uses distinct areas or dates.

Geospatial Examples:
ArcGIS Pro provides accuracy tools (e.g., confusion matrices and IoU for pixel classification) and guidance on training/validation chip sizes and strides for object detection.

Geospatial Resourcess:
• Training-data tips (tile/stride, splits): https://www.esri.com/arcgis-blog/products/arcgis-pro/geoai/tips-for-training-data-preparation-for-object-detection-models
• Compute accuracy tool (Pro 3.5): https://www.esri.com/arcgis-blog/products/arcgis-pro/geoai/whats-new-for-geoai-in-the-image-analyst-extension-of-arcgis-pro-3-5

Hallucination

In the context of AI, **hallucination** refers to an instance where an AI model **produces information that is false, fabricated, or not grounded in its training data**, yet presents it with confidence as though it were correct. Hallucinations occur when the model fills gaps, makes incorrect inferences, or generates plausible-sounding but inaccurate details due to limitations in its understanding, context handling, or the data it was trained on.

Retrieval-Augmented Generation (RAG)

A technique in which an AI model retrieves relevant information from an external knowledge source such as documents, databases, or vector stores and uses that information to generate more accurate, grounded responses, reducing the likelihood of hallucinations.

Confusion Matrix

Table showing classification accuracy across categories, for example: showing how many times a photo app correctly or incorrectly identified pets vs. people.

Geospatial Example:

- Validating crop classification results in precision agriculture.

Resource:
https://support.esri.com/en-us/gis-dictionary/confusion-matrix

Precision & Recall

Metrics for how accurately and completely features are identified.

- Precision: How many of the emails marked as spam were actually spam.
- Recall: How many of the actual spam emails were caught. Overfitting

Overfitting happens when an AI model learns the training data too well, including its noise and quirks, so it performs brilliantly on that data but poorly on new, unseen data. It's like memorising answers for a test instead of understanding the subject you'll fail when the questions change.

Geospatial examples:

- A land cover classification model trained on imagery from one city misclassifies rural areas elsewhere because it learned very specific urban patterns.
- A flood prediction model that works perfectly for one river basin but fails in another because it over-relied on local historical data.

Resource:
https://www.esri.com/arcgis-blog/products/arcgis-pro/geoai/tips-for-training-data-preparation-for-object-detection-models

Underfitting

Underfitting is the opposite problem the model is too simple and doesn't learn enough from the training data, so it performs badly both on the training set and new data. It's like skimming a textbook and missing the key ideas.

Geospatial examples:

- A model using only two features (e.g., elevation and rainfall) to predict landslides, ignoring other critical factors like soil type and vegetation.
- A basic clustering algorithm that lumps diverse land uses into one category because it can't capture complexity.

Model Drift

Model drift occurs when a model's accuracy declines over time because the real-world data it sees changes from what it was trained on. It's like using an old map for a city that's constantly building new roadsthe map becomes less useful.

Geospatial examples:

- An urban growth prediction model trained five years ago underestimates expansion because new transport corridors have changed development patterns.
- A crop health model becomes inaccurate as farming practices and crop varieties evolve.

Data Drift

Data drift means the characteristics of incoming data change over time compared to the data the model was trained on. Even if the model itself hasn't changed, its inputs no longer match expectations.

Geospatial examples:

- Seasonal changes in vegetation alter spectral signatures, confusing a land cover classifier.
- Sensor upgrades produce higher-resolution imagery, shifting pixel distributions and affecting model performance.

Explainable AI (XAI)

Explainable AI refers to techniques that make AI decisions understandable to humans. Instead of being a "black box," the model shows why it made a certain prediction. This builds trust and helps users verify results.

Geospatial examples:

- Showing which spectral bands influenced a classification of "forest" in satellite imagery.
- Explaining why a routing algorithm chose a particular evacuation path during a disaster scenario (e.g., prioritising elevation and road connectivity).

Resources:

- https://trust.arcgis.com/en/trusted-ai/trusted-ai.htm
- https://www.esri.com/arcgis-blog/products/arcgis-pro/geoai/whats-new-for-geoai-in-the-image-analyst-extension-of-arcgis-pro-3-5

Model Retraining

Model retraining is the process of updating an AI model with new data to restore or improve its accuracy. It's like refreshing your knowledge when new information becomes available.

Geospatial examples:

- Retraining a building detection model with recent aerial imagery to include new construction.
- Updating a flood risk model after major land use changes or new hydrological data.

Resource:

https://www.esri.com/arcgis-blog/products/arcgis-pro/geoai/whats-new-for-geoai-in-the-arcgis-pro-3-6-image-analyst-extension

Human-in-the-Loop (HITL)

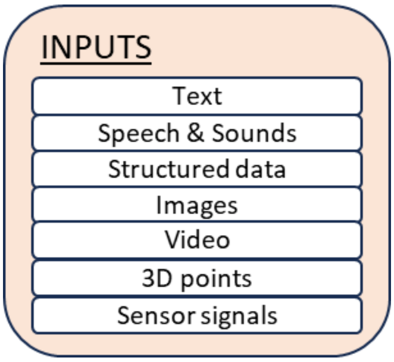
Human-in-the-loop systems combine automated AI with human judgement to validate, correct, or guide outcomes.

Geospatial resource:
Labelling imagery for deep learning and reviewing AI-detected features before final acceptance.
Geospatial AI, An Overview

Deep Dive: Detailed Terminology Used in Evaluation & Reliability

Defined Term	Brief Description	GIS Relevance	ArcGIS Reference
Intersection over Union	Measures the overlap between two shapes—typically a predicted bounding box and a ground truth bounding box. It is calculated as the area of intersection divided by the area of union. IoU is widely used to evaluate object detection and image segmentation models.	ArcGIS uses IoU when assessing deep learning based object detection accuracy for example, checking how well detected buildings, trees, or vehicles overlap with ground truth polygons.	https://pro.arcgis.com/en/pro-app/latest/tool-reference/image-analyst/compute-accuracy-for-object-detection.htm
Bounding Box (in computer vision)	A rectangular box that an AI model places around an object within an image to identify, classify, or track it; bounding boxes are the simplest form of spatial annotation used in object detection tasks.		
Log-likelihood	Log likelihood is the logarithm of the probability of observed data under a particular statistical model. It transforms products of small probabilities into sums, improving numerical stability and simplifying optimisation. Many statistical models, from logistic regression to hidden Markov models, are trained by maximising log likelihood.	ArcGIS uses likelihood based methods in tools such as Maximum Likelihood Classification (MLC), which implicitly involves log likelihood calculations when assigning raster cells to classes.	https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/maximum-likelihood-classification.htm
Rectified linear unit	A rectified linear unit is a neural network activation function defined as f(x)=max(0,x). It introduces	ReLU is used in deep learning models run within ArcGIS Pro (e.g., building footprint extraction),	Deep learning workflows in ArcGIS (which use ReLU-based models) are described here:

	non linearity whilst remaining computationally efficient and helps to avoid vanishing gradients in deep networks. ReLU has become the default activation in many modern deep learning architectures.	enabling convolutional neural networks to learn spatial features from aerial imagery more effectively.	https://www.esri.com/arcgis-blog/products/arcgis-enterprise/geoai/unlocking-the-power-of-geospatial-ai-with-arcgis https://developers.arcgis.com/python/latest/guide/how-wnet-cgan-works/
F1 Score	A harmonic mean of precision and recall. For example: a balanced score used to evaluate how well a medical diagnosis app identifies diseases.	Assessing wildfire detection performance in satellite imagery	https://www.esri.com/arcgis-blog/products/arcgis-pro/announcements/new-smarter-prediction-model-evaluation-in-arcgis-pro-3-6
Kappa Coefficient	A statistic measuring classification agreement beyond chance. For example: measuring how much two movie reviewers agree on ratings, beyond random chance.	Evaluating forest/non-forest classification accuracy.	https://support.esri.com/en-us/gis-dictionary/kappa-coefficient
Root Mean Square Error (RMSE)	Measure of average error in continuous predictions. For example: evaluating how far off a weather app's temperature predictions are from actual temperatures.	Assessing DEM accuracy improvements with AI.	https://support.esri.com/en-us/gis-dictionary/rmse https://content.esri.com/esri_content_doc/dbl/us/j10322_census_2020_accuracy_analysis_final.pdf
Mean Absolute Error (MAE)	Mean of absolute differences between predicted and actual values. For example: calculating the average error in predicted vs. actual delivery times for a food delivery app.	Evaluating predicted vs. observed rainfall maps.	



INPUTS

AI models can process a wide variety of **input types**, depending on the task and domain. Common input types and examples:

Text

Written language in natural or structured form.

Examples:

- Emails, chat messages, articles (e.g., for sentiment analysis or summarization)
- Corporate documents stored on file servers
- Organisation's chat messages through channels such as Microsoft Teams, or through discussion forums for user issues or defect resolution etc.
- Code snippets (e.g., for code completion)
- Product reviews (e.g., for classification)

Used in:
Natural Language Processing (NLP), chatbots, translation, search engines.

Geospatial Example:

- AI can interpret natural language queries like *"Find areas in Auckland with high vegetation near water bodies"* and convert them into spatial analyses using GIS data.

Speech & Sounds

Audio signals representing spoken language or other sounds.

Examples:

- Voice commands (e.g., "Play music" for virtual assistants)
- Call centre recordings (e.g., for emotion detection)
- Environmental sounds (e.g., detecting glass breaking or alarms)

Used in:
Speech recognition, audio classification, voice synthesis.

Geospatial Examples:

- AI can process spoken commands like *"Highlight flood-prone zones near the Waikato River"* and convert them into geospatial queries.
- AI can convert spoken updates into geospatial updates.

Resources:

- https://www.esri.com/arcgis-blog/products/arcgis-pro/geoai/vision-language-models-geospatial-analysis
  https://resource.esriuk.com/blog/a-voice-user-interface-for-the-arcgis-platform-using-rest-and-alexa/
- https://www.esri.com/arcgis-blog/products/developers/geoai/extend-the-capabilities-of-custom-apps-with-ai-skills
- https://www.esri.com/arcgis-blog/products/arcgis-pro/geoai/use-vision-language-models-to-optimize-object-classification

Structured data

Data organized in rows and columns, often from databases or spreadsheets. Can have significant differences from 'Text' (free text/unstructured text) or Speech above, because the context of the row and column structuring can add additional cues to comprehend the data organised in that structure.

Examples:

- Customer records (e.g., age, income, purchase history)
- Financial transactions (e.g., for fraud detection)
- Sensor logs in tabular format

Used in:
Predictive modelling, business intelligence, recommendation systems.

Geospatial Example:

- AI can analyse tabular datasets (e.g., census or traffic data) to map population density or congestion hotspots on a city map.

Unstructured data

Unstructured data refers to information that does not follow a predefined data model or organised format, making it difficult for traditional databases to store and process. Unlike structured data (rows and columns in a database), unstructured data is often text-heavy or media-rich and lacks consistent attributes. Examples include free-text documents, images, videos, audio recordings, and social media posts. AI and machine learning techniques are commonly used to interpret and extract meaning from unstructured data.

Geospatial Examples:

- **Social Media Posts with Location Tags**
  Text, images, and videos shared online that include geotags but lack structured attributes can provide valuable insights for GIS. These types of data are often used for crowd-sourced disaster mapping or tourism trend analysis, where location-based social content helps identify patterns and events in real time.

- **Full Motion Video (FMV)**
  Full Motion Video refers to video captured from UAVs or surveillance systems that includes embedded geospatial metadata. This type of data supports advanced GIS applications such as object detection and change detection in dynamic environments, enabling analysts to monitor and respond to evolving conditions.
- **PDF Reports and Scanned Maps**
  Documents containing spatial descriptions or embedded maps without machine-readable geometry are considered unstructured data. AI-driven techniques such as Optical Character Recognition (OCR) and Natural Language Processing (NLP) can extract place names and other geographic details from these sources, converting them into GIS-ready data for further analysis.

Resources:
https://www.esri.com/en-us/arcgis/products/locatext/overview

Images

Visual data in the form of 2D pixel arrays.

Examples:

- Photos of animals (e.g., for species classification)
- Medical scans (e.g., X-rays for diagnosis)
- Handwritten digits (e.g., for digit recognition)

Used in:
Computer vision, facial recognition, car number plate recognition, medical imaging.

Geospatial Example:
- AI can interpret satellite or drone imagery to detect land-use changes and classify regions into urban, agricultural, or forested areas.

Video

Sequences of images (frames) over time, often with audio.

Examples:

- Camera Car video and LiDAR data collection.
- Camera Car parking enforcement vehicles.
- Undersea Pipeline camera tracing footage.
- Surveillance footage (e.g., for activity recognition)
- Sports replays (e.g., for highlight detection)
- Driver monitoring (e.g., for drowsiness detection)

Used in:
Action recognition, video summarization, autonomous vehicles.

Geospatial Examples:

- AI can analyse time-lapse videos to monitor shoreline erosion or urban expansion over time.
- AI can analyse video of a transport routes to identify where maintenance is required, based on the location of images at a certain timestamp from a moving video.

Resources:

- https://pro.arcgis.com/en/pro-app/latest/help/analysis/image-analyst/introduction-to-full-motion-video-in-arcgis-pro.htm
- https://www.esri.com/en-us/lg/industry/public-works/stories/county-innovates-using-geoai-to-inventory-ada-curb-ramps-saving-significant-time-money

3D points

Spatial data representing objects or environments in 3D space.

Examples:

- LIDAR scans (e.g., for self-driving cars)
- 3D models of buildings (e.g., for architecture or AR)
- Robotics navigation maps

Used in:
3D object detection, robotics, augmented reality.

Geospatial Example:

- AI can process LiDAR point clouds to generate elevation models and identify flood-risk zones in low-lying areas.

Sensor signals

Time-series data from physical sensors.

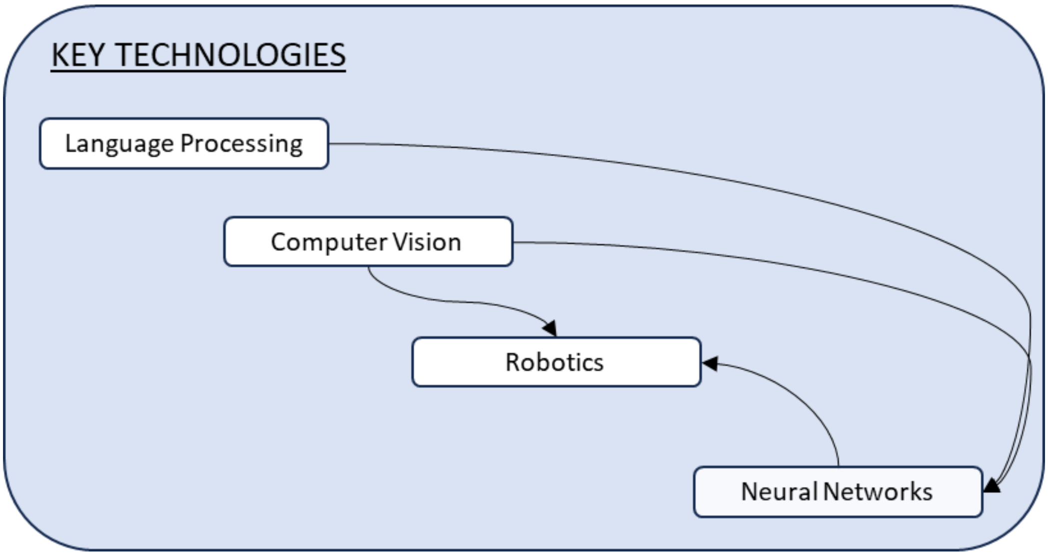
Examples:

- Accelerometer data (e.g., for fitness tracking)
- Temperature and humidity readings (e.g., for climate control)
- River gauges for detecting flood levels.
- Motion detectors

Used in:
IoT applications, health monitoring, industrial automation.

Geospatial Example:

- AI can use IoT sensor data (e.g., temperature, humidity, soil moisture) to predict drought conditions and visualize them spatially.



KEY TECHNOLOGIES

Key technologies of AI have four major domains: Language Processing, Computer Vision, Robotics and Neural Networks. These domains are explained in the following section.

Natural Language Processing (NLP)

NLP enables machines to understand, interpret, generate, and respond to human language.

Key Concepts:

- **Tokenization & Parsing**: Breaking text into words or phrases.
- **Named Entity Recognition (NER):** Identifying names, places, dates.
- **Entity Annotation**: Entity annotation (often via Named Entity Recognition) identifies and labels real world entities (people, places, organisations, etc.) in text. It turns unstructured text into structured records for search, linking, and mapping.
- **Machine Translation**: Translating between languages. Machine translation (MT) automatically converts text from one language to another using models trained on parallel corpora (multiple versions of the same text in different languages). Neural MT provides fluency and domain adaptation.
- **Language Models**: Predicting and generating human-like text (e.g., GPT, BERT).
- **NLU (Natural Language Understanding):** Natural Language Understanding enables AI systems to interpret user input, extract meaning, classify text, detect entities, and convert written language into structured information for downstream reasoning.
- **Natural Language Generation (NLG):** Natural Language Generation is the AI process of producing human-readable text from structured or unstructured data. NLG systems convert data patterns into coherent narratives, summaries, or explanations.
- **Prompt Engineering:** Prompt engineering is the practice of crafting precise, structured input prompts to optimize the performance of language models—improving accuracy, relevance, and reliability of AI outputs across tasks.

- **Meta Prompt / System Prompt** : A *system prompt* (often called a *meta prompt*) is the foundational instruction given to a generative AI model that sets its role, tone, constraints, and high-level goals. It steers how the model interprets all subsequent user prompts and can enforce organisational guardrails. Good system prompts are concise, unambiguous, and aligned with responsible-AI policies. They are distinct from user prompts, which ask for a specific output in the moment.

Geospatial Examples:
In ArcGIS Hub Assistant, site editors can define the assistant's "personality" and guidance via a *system prompt* so the assistant answers public questions about open data in a consistent style.

Geospatial Resources: https://doc.arcgis.com/en/hub/get-started/hub-assistant.htm

Esri demonstrates using NER to extract addresses and place names from reports, then geocoding them for mapping.

Geospatial Resources: https://mediaspace.esri.com/media/t/1_edd2mk6y

Examples:

- Chatbots and virtual assistants (e.g., Siri, ChatGPT)
- Language translation (e.g., Google Translate)

Geospatial Examples:

- **Place Name Disambiguation**: NLP helps AI figure out which "Paris" you mean France or Texas when processing text for geocoding.
- **Smart Map Search**: Converts queries like "nearest cycleway to Queen Street" into spatial searches for navigation or GIS apps.

Geospatial Resources:

- ArcGIS Pro 3.0 introduced a GeoAI Text Analysis toolset that brings natural language processing capabilities into GIS, allowing users to work with unstructured text by performing tasks such as entity extraction, classification, transformation, and other NLP workflows; it highlights how text often contains geographic information and shows how these new tools can extract, organise, and analyse such text to support real-world GIS problems. https://www.esri.com/arcgis-blog/products/arcgis-pro/analytics/text-analysis-in-arcgis-pro-3-0

- The Process Text Using AI Model tool in ArcGIS Pro enables users to apply custom or pretrained deep-learning language models to text from feature classes, tables, or files in order to perform tasks such as text transformation, entity recognition, classification, translation, summarisation, and text generation. https://pro.arcgis.com/en/pro-app/latest/tool-reference/geoai/process-text-using-ai-model.htm

- Writing highly specific prompts for ArcGIS AI Assistants—for example: *"Create a table summarizing land-use categories, sorted by area, and highlight high-density residential polygons."* https://www.esri.com/arcgis-blog/products/arcgis/mapping/prompt-writing-for-ai-assistants

- ArcGIS Instant Apps and Survey123 include Translation assistants that use Azure AI Translator to generate multilingual app or survey text, streamlining public facing GIS experiences. Instant Apps Translation Assistant: https://www.esri.com/arcgis-blog/products/instant-apps/mapping/instant-apps-translation-assistant-beta
Survey123 Auto Translate: https://www.esri.com/arcgis-blog/products/survey123/field-mobility/translate-your-surveys-with-the-click-of-a-button

Computer Vision

Computer vision enables machines to interpret and make decisions based on visual data.

Key Technologies:

- *Image Classification*: Identifying objects in images.
- *Object Detection*: Locating multiple objects in a scene.
- *Facial Recognition*: Identifying or verifying individuals.
- *Image Segmentation*: Dividing an image into meaningful parts.
- *Optical Character Recognition (OCR)* **:** OCR is a technology that allows computers to read text from images or scanned documents, like turning a photo of a street sign or an old map into editable and searchable text.

Geospatial Concepts:
Computer vision plays a critical role in GIS, especially in the context of remote sensing, because it enables automated interpretation and extraction of meaningful information from imagery and sensor data. The defined terms below elaborate on the concepts of computer vision and have particular relevance to the GIS related domain of the processing of remotely sensed imagery.

Spectral Signature Recognition
Identifying materials by their spectral reflectance patterns. E.g. Distinguishing crop types in multispectral data.

Semantic Segmentation
Pixel-level classification with meaningful categories. E.g. distinguishing between roads, rivers, and forests.

Instance Segmentation
Separating individual objects of the same class. E.g. differentiating each tree in a plantation.

Change Detection
Identifying differences between georeferenced images over time. E.g. monitoring coastal erosion.

Feature Extraction

Automated identification of spatial features from imagery. E.g. digitizing road networks from satellite data.

Visual grounding / Vision–language models

Visual grounding links natural language descriptions to specific objects or regions within an image. It requires models to understand both spatial relationships and linguistic context to localise the referenced feature accurately.

Resources:

- https://www.esri.com/arcgis-blog/products/arcgis-pro/geoai/revolutionizing-image-segmentation-with-sam-segment-anything-model
- https://www.esri.com/arcgis-blog/products/arcgis/geoai/dev-summit-2024-extending-the-segment-anything-model-sam
- https://www.esri.com/arcgis-blog/products/arcgis-pro/geoai/text-sam-extracting-gis-features-using-text-prompts
- https://www.esri.com/arcgis-blog/products/arcgis-pro/imagery/performing-feature-extraction-classification-using-deep-learning-with-arcgis-pro
- https://www.esri.com/training/catalog/661ed9e56eb80b0e91ead1b4/extracting-features-with-deep-learning-using-arcgis-online/
- https://www.esri.com/arcgis-blog/products/arcgis-living-atlas/imagery/learn-to-use-ai-to-extract-information-from-world-imagery
- https://www.esri.com/arcgis-blog/products/arcgis-pro/geoai/vision-language-models-geospatial-analysis

Robotics

Robotics combines AI with mechanical systems to perform tasks in the physical world.

Key Technologies:

- *Path Planning*: Navigating through environments.
- *Computer Vision*: For object recognition and manipulation.
- *Sensor Fusion*: Integrating data from multiple sensors (e.g., cameras, LIDAR).
- *Reinforcement Learning*: Teaching robots through trial and error.

Geospatial Example:

- Autonomous Drones for Mapping: AI-powered drones capture aerial imagery and generate high-resolution maps for GIS.
- Surveying Robots: Ground robots equipped with sensors collect geospatial data in hard-to-reach areas like mines or disaster zones.
- Precision Agriculture: Robots use GPS and AI to navigate fields, apply fertilizers, and monitor crop health.
- Infrastructure Inspection: Robots inspect bridges, pipelines, or power lines using computer vision and geospatial positioning.

Resources:

https://community.esri.com/t5/education-blog/robotics-and-arcgis-online-for-the-classroom/ba-p/1218331

Neural Networks

Neural networks are computational models inspired by the human brain, used to recognize patterns and make predictions. They are the foundation of deep learning, a subset of machine learning, and are particularly powerful for tasks involving complex patterns, non-linear relationships, and large datasets.

A typical neural network consists of:

1. **Layers**

- *Input Layer*: Takes in raw data (e.g., satellite image pixels, coordinates, sensor readings).
- *Hidden Layers*: Perform transformations and extract features.
- *Output Layer*: Produces predictions (e.g., land cover class, object location).

2. '**Neurons'.** Think of a neuron as a tiny decision-maker that looks at some information, does a bit of math, and then decides what to pass on. The process works something like this:

- **Receives Inputs** The neuron gets numbers from other neurons or from the data itself (like pixel values from an image or coordinates from a map).
- **Applies** *Weights* Each input is multiplied by a weight this tells the neuron how important that input is. Bigger weight = more influence.
- **Adds a** *Bias* In the domain of AI processing a bias is just a number added to help the neuron make better decisions like a little nudge to shift its thinking. This should not be confused with "AI bias" which in other contexts can mean a systematic tendency for an algorithm to produce skewed or unfair outputs due to biased training data, model structure, or evaluation. Bias can lead to unequal treatment of groups or misinterpretation of spatial patterns. The first definition of bias outlined here can generate instances of the second, but does have a distinct meaning.
- *Activation Function* This is a rule that decides whether the neuron should "fire" (pass its result on) or not. It helps the network handle complex patterns.

A simple analogy: Imagine you're deciding whether to go for a walk:

- **Inputs**: Weather, time of day, how tired you are.
- **Weights**: You care a lot about the weather, a little about the time, and not much about being tired.
- **Bias**: You generally like walking, so you're slightly biased toward going.
- **Activation Function**: You decide "Yes" if the total score is high enough.

That's basically what a neuron does it takes inputs, weighs them, adjusts with a bias, and makes a decision. Now imagine that vast numbers (in some cases billions) of neurons all making tiny decisions in a short amount of time. Consider GPT-4 which is capable of processing billions of parameters about which words are related to each other, which order they are in, which are the most important words in a sentence (downplay "the" but give more attention to "GPT-4"), and what words were in the previous sentences and the following sentences. In simple terms these processes are the core of what AI does.

Geospatial Example:

- Analysing a pixel in an image and also processing other pixels that are nearby for similarities or differences or patterns, and doing this with millions of pixels in a short timespan.

In GIS, biased training data (e.g., imagery mainly from urban regions) can cause an AI model to misclassify rural land cover or under-detect features in under-represented geographies.

- **Esri Link (AI & fairness context):** https://www.esri.com/en-us/geospatial-artificial-intelligence/overview

3. **Weights and Biases.** These are the parameters the network learns during training. They determine how strongly inputs influence outputs.

Key Technologies:

- *Feedforward Neural Networks*: Basic structure for prediction tasks.
- *Convolutional Neural Networks (CNNs)*: Specialized for image data.
- *Recurrent Neural Networks (RNNs)*: Designed for sequential data like text or time series.
- *Transformers*: Advanced architecture for NLP and vision tasks. Transformers are neural network models introduced in the paper *"Attention Is All You Need"* (Vaswani et al., 2017). They are designed to handle sequential data (like text) but without relying on traditional recurrent or convolutional structures.
- *Attention head:* a component within a transformer model that learns to focus on specific parts of an input sequence. Multiple heads capture different relationships simultaneously, enabling models to understand context, dependencies, and semantics. Attention heads form the foundation of modern large language and vision-language models. They support parallel processing of contextual information for improved predictions.
- *Perceptron*: the simplest form of a neural network unit: it computes a weighted sum of inputs, applies an activation function, and produces a binary or continuous output. It forms the foundational building block of larger networks, though by itself it can only learn linearly separable patterns.

Key Concepts related to Transformers

Tokens

In AI language models, **tokens** are the basic units of text that the model reads, processes, and generates. A token can be a **whole word**, a **part of a word**, punctuation, a number, or even whitespace, depending on how the model's tokenizer is designed.

Modern AI models break text into tokens because this allows them to handle language more efficiently and consistently across different languages, writing styles, and word lengths.

How tokens work

1. **Tokenisation**
   Before a model can understand or generate text, the input text is converted into a sequence of tokens.
   - For example, the sentence:
     "The geospatial system works well."
     may become something like:
     `["The", "geo", "spatial", "system", "works", "well", "."]`
     (Actual tokens may differ depending on the *tokeniser*.)
2. **Numerical Representation**
   Each token is mapped to a numerical ID from the model's vocabulary.
   The model does not "see" words—only these numerical representations.
3. **Model Processing**
   During both understanding and generation, the model predicts tokens one at a time, which gradually form full sentences.

Why tokens matter

- **Context window size**
  A model can only handle a certain number of tokens at once (its "context window").
  Larger windows allow the model to consider more text when forming responses.
- **Cost and speed**
  AI usage is often priced per token.
  More tokens = higher cost and longer processing time.
- **Accuracy and meaning**
  The way words are split into tokens affects how well the model understands nuance, rare words, or technical language (such as GIS terminology).

Examples of token behaviour

- **Common words** ("cat", "system", "analysis") are often single tokens.
- **Rare or technical words** ("hydrogeomorphology", "GeoAI") might be broken into multiple tokens.
- **Punctuation** is usually treated as separate tokens.
- **Languages without spaces** (e.g., Chinese, Japanese) rely heavily on sub word tokenisation.

Self-Attention Mechanism

Instead of processing words one by one (like RNNs), transformers use *attention* to weigh the importance of each word in a sentence relative to others.

This allows the model to capture long-range dependencies efficiently.

Here's a simple example of how **Self-Attention** works in a Transformer:

Sentence:

The cat sat on the mat.

Suppose we want to compute the attention for the word **“cat”**.

Step 1: Represent Words as Vectors

Each word is converted into an embedding (a numeric vector). For simplicity, assume:

The → [0.1, 0.2]

cat → [0.3, 0.7]

sat → [0.4, 0.5]

on → [0.6, 0.9]

the → [0.1, 0.2]

mat → [0.8, 0.3]

Step 2: Compute Query, Key, and Value

For each word, the model creates:

- **Query (Q)**: What am I looking for?
- **Key (K)**: What do I offer?
- **Value (V)**: My actual content.

These are linear transformations of the embeddings.

Step 3: Calculate Attention Scores

For **“cat”**, we take its Query and compare it with every Key (including its own) using a dot product:

Score(cat, The) = Q_cat · K_The

Score(cat, cat) = Q_cat · K_cat

Score(cat, sat) = Q_cat · K_sat

…

This gives raw scores like:

[0.2, 0.9, 0.4, 0.1, 0.2, 0.3]

Step 4: Apply Softmax

Convert scores into probabilities:

[0.05, 0.40, 0.20, 0.10, 0.05, 0.20]

This means **“cat”** pays most attention to itself (0.40) and some to “sat” and “mat”.

Step 5: Weighted Sum of Values

Multiply each Value vector by its attention weight and sum them:

Output for "cat" = 0.05*V_The + 0.40*V_cat + 0.20*V_sat + ...

This creates a new representation of “cat” that incorporates context from the whole sentence.

Encoder-Decoder Architecture

Encoder: Processes input text and creates contextual representations.

Decoder: Generates output (e.g., translated text, summaries) using those representations.

Positional Encoding

Since transformers don’t process sequences in order, they add positional information to tokens so the model understands word order.

Multi-Head Attention

Multiple attention layers run in parallel, capturing different types of relationships between words.

Why Are Transformers Important in NLP?

Scalability: They handle large datasets and parallelize well.

Performance: State-of-the-art results in translation, summarization, question answering, and more.

Foundation for Large Language Models (LLMs): GPT, BERT, T5, and similar models are all based on transformer architecture.

Applications

Machine Translation (e.g., Google Translate)

Text Summarization

Sentiment Analysis

Chatbots and Conversational AI

Code Generation and Document Understanding

Examples:

Image recognition (CNNs)

Language generation (Transformers like GPT)

Time-series forecasting (RNNs, LSTMs)

The models and algorithms behind facial recognition in smartphones that unlock the device by identifying the user's face.

Tokens in Image Processing

In the context of **AI image interpretation**, *tokens* are relevant because modern vision models no longer process images purely as grids of raw pixels; instead, they often **convert patches (or 'pixel blocks' of an image into "visual tokens"**, which the model then analyses in a similar way to how language models process text tokens.

1. Pixels are grouped into patches

Instead of looking at every pixel individually, many AI vision models (especially Vision Transformers, or ViTs) divide an image into **small, fixed-size patches**—for example, 16×16 pixels.

Each patch becomes a **visual token**, meaning a compact mathematical representation of that part of the image.

This allows the model to process an image as a *sequence* rather than a grid.

2. Tokens help the model understand patterns

Visual tokens capture the essential features of the pixel patch, such as:

- texture
- edges
- colour patterns
- shapes
- spatial context

Because each token represents a meaningful chunk of the image, the model can recognise higher-level structures the same way language models recognise words or subwords.

3. Tokens enable long-range relationships

Pixel grids make it hard to understand relationships between distant parts of an image (e.g., a car on the left and a road sign on the right).

Visual tokens allow the model to compare information **across the whole image at once**, enabling stronger reasoning about:

- object relationships
- scene layout
- context and depth cues
- anomalies in geospatial imagery

This is especially relevant in GIS, where patterns may span large areas.

4. Efficiency and scalability

Processing every pixel directly is computationally expensive.
By converting pixels into tokens:

- the model can scale to very large images
- computation becomes more efficient
- training becomes feasible for high-resolution or multi-band imagery

This matters in remote sensing workflows, where images are frequently huge.

5. Tokens act as the building blocks of image understanding

Just as text models build meaning from tokens, image models build meaning from **visual tokens**. The model predicts labels, bounding boxes, or segment masks based on interactions between tokens, not raw pixels.

Geospatial Example:

- ArcGIS leverages neural networks for advanced 3D GIS tasks, such as generating realistic urban environments. In the case study below, Esri's Urban Scene Generator (USG) uses a generative neural network trained on large raster and vector datasets from multiple cities to create practical building distributions aligned with urban planning concepts like the "15-minute city".
- The process involves selecting an appropriate neural network architecture, tuning hyperparameters, and normalizing diverse spatial features (e.g., building coverage, terrain models). Training these models requires massive compute resources, so Esri uses GPU virtualization with NVIDIA AI Enterprise to optimize performance.
- This setup allows splitting GPUs across virtual machines for parallel experiments and later consolidating resources for full-scale production training. Neural networks in this context

enable ArcGIS to automate complex spatial modelling, producing realistic 3D city layouts for planning and analysis.

Case Study:
https://www.esri.com/arcgis-blog/products/3d-gis/3d-gis/virtualizing-3d-training-models-with-nvidia-ai-enterprise

Resources:

https://www.esri.com/arcgis-blog/products/arcgis-pro/geoai/use-remote-sensing-foundation-models-with-arcgis

https://proceedings.esri.com/library/userconf/devsummit19/papers/DevSummitPS_140.pdf

https://pro.arcgis.com/en/pro-app/latest/arcpy/image-analyst/pixelblock-class.htm

Advanced AI Architectures

Convolutional Neural Network (CNN)

A Convolutional Neural Network is a type of deep learning model primarily used for processing grid-like data, such as images. It employs convolutional layers to automatically learn spatial hierarchies of features, making it highly effective for image recognition and classification tasks.

Geospatial Example:

- CNNs can classify satellite imagery into land cover types, such as distinguishing between urban areas, forests, and water bodies.

Recurrent Neural Network (RNN)

A Recurrent Neural Network is designed to handle sequential data by maintaining a memory of previous inputs through recurrent connections. This makes it suitable for tasks involving time series or language modelling.

Geospatial Example:

- RNNs can analyse temporal patterns in climate data to predict future temperature or rainfall trends for specific regions.

Long Short-Term Memory (LSTM)

LSTM networks are a specialised form of RNN that address the problem of long-term dependencies by using memory cells and gating mechanisms. They excel at learning patterns over extended sequences without losing context.

Geospatial Example:

- LSTMs can forecast traffic congestion by learning from historical traffic flow data across a city's road network.

Autoencoder

An Autoencoder is a neural network used for unsupervised learning, aiming to compress data into a lower-dimensional representation and then reconstruct it. This is useful for feature extraction and noise reduction.

Geospatial Example:

- Autoencoders can reduce the dimensionality of hyperspectral imagery, making it easier to process and analyse large datasets.

Variational Autoencoder (VAE)

A Variational Autoencoder extends the concept of autoencoders by introducing probabilistic elements, enabling the generation of new data samples similar to the original dataset. It is widely used in generative modelling.

Geospatial Example:

- VAEs can create synthetic satellite images to augment training datasets for land cover classification.

Transformer Model

The Transformer model revolutionised sequence processing by using self-attention mechanisms instead of recurrence, allowing for parallel computation and better handling of long-range dependencies.

Geospatial Example:

- Transformers can process large-scale geospatial text data, such as extracting location-based insights from social media posts.

Attention Mechanism

Attention mechanisms enable models to focus on the most relevant parts of the input when making predictions, improving performance in tasks with complex dependencies.

Geospatial Example:

- Attention can help identify key regions in satellite imagery for disaster assessment, such as focusing on flood-affected areas.

Graph Neural Network (GNN)

Graph Neural Networks are designed to work with graph-structured data, learning relationships between nodes and edges. They are ideal for modelling networks and spatial connectivity.

Geospatial Example:

- GNNs can analyse road networks to optimise routing and predict traffic bottlenecks based on connectivity patterns.

DATA STRUCTURES

Graph Databases

- **Definition**:
  A type of database that stores data as:
  - **Nodes**: Represent entities or things like people or places.
  - **Edges**: Represent relationships between entities.
- **Key Difference from Traditional Databases**:
  - Traditional relational databases use tables and foreign key relationships.
  - Graph databases are optimized for handling **complex connections** more naturally and efficiently.

Important Note on Terminology
- Terms like **nodes**, **edges**, and **vectors**:
  - Have **different meanings in AI vs GIS contexts**.
  - This can cause confusion for GIS professionals.
  - Always check **context** to understand the correct meaning.

Graph Databases in AI

- Used to represent **networks of information**, such as:
  - **Social networks** (connections between people).
  - **Recommendation systems** (relationships between users and products).
  - **Knowledge graphs** (structured relationships between concepts).

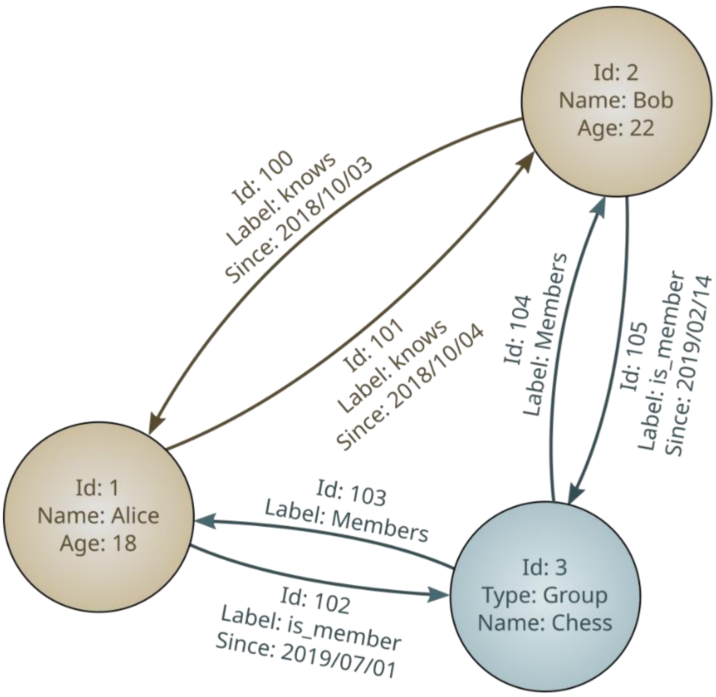


Figure 3- https://en.wikipedia.org/wiki/Graph_database

- Benefits:
    - Helps AI systems **reason**, **search**, and **learn** by understanding connections.

Geospatial Example:

ArcGIS Knowledge uses **knowledge graphs** for GIS applications. https://www.esri.com/en-us/arcgis/products/arcgis-knowledge/overview

Resources:

- https://www.esri.com/arcgis-blog/products/arcgis-enterprise/data-management/what-is-a-knowledge-graph
- https://www.esri.com/about/newsroom/arcnews/arcgis-knowledge-graphs-are-coming-to-the-web

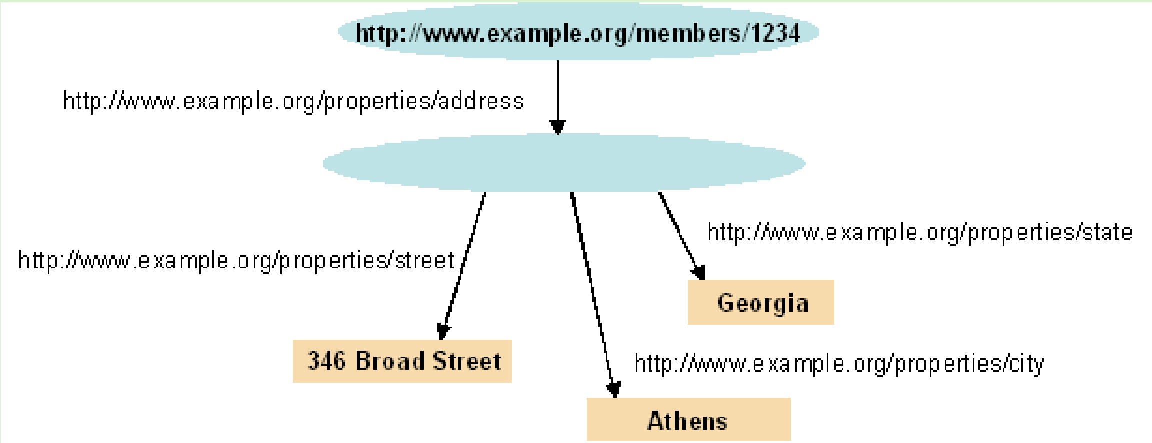


Figure 4- Resource Description Framework (RDF) Source: https://en.wikipedia.org/wiki/Graph_database

Vector Databases

- **Definition**:
  A specialized database designed to:
    - Store **vectors** (lists of numbers representing data in machine-readable form).
    - Perform **fast similarity searches** between vectors.

Vectors (in the context of AI)

- Vectors often come from **embeddings** created by AI models:
    - Convert text, images, or audio into numerical form.
    - Enable machines to understand and compare data based on meaning, not just keywords.

Vector Embeddings

What Are Embeddings?

Think of embeddings as **numeric fingerprints for words, sentences, or even images**.

Computers don't understand text directly they understand numbers.

An **embedding** is a way to turn something like the word *"cat"* into a list of numbers (a vector) that captures its meaning.

Words with similar meanings (like *"cat"* and *"kitten"*) will have embeddings that are close together in this numeric space.

Example:

"cat" → [0.3, 0.7, -0.2, ...]

"dog" → [0.4, 0.6, -0.1, ...]
These vectors are positioned so that similar concepts are near each other.

Semantic Embeddings

What Are Semantic Embeddings?

Semantic embeddings go a step further:

They don't just encode the word they encode **meaning in context**.

For example, the word *"bank"* in "river bank" vs. "money bank" will have different semantic embeddings because the context changes the meaning.

They are used for tasks like **search**, **recommendations**, and **question answering**, because they capture the *intent* or *concept* behind text, not just the exact words.

Why Are They Important?

Search: Instead of matching exact words, semantic embeddings let you find things that mean the same thing.

Chatbots & AI: They help models understand what you're asking, even if you phrase it differently.

Clustering & Classification: Group similar ideas together automatically.

Analogy:
Imagine a huge map where every word or sentence is a dot. Embeddings are the coordinates of those dots. Semantic embeddings make sure dots with similar meanings are close together even if the words are different.

- Help computers understand **meaning**, not just spelling or appearance.

- Examples:
    - “Avenue” and “Boulevard” → vectors very close (similar meaning).
    - “Avenue” and “Cycleway” → slightly further apart.
    - “Avenue” and “River” → much further apart.

Similarity Search

Similarity search is a way for AI to find items that are alike by comparing their vector representations in a database.

- Process:
    - AI converts a query (e.g., a sentence) into a vector.
    - Searches for other vectors that are **close** in numerical space.
- Advantages:
    - Faster and more flexible than traditional keyword matching.
    - Supports **semantic understanding** rather than exact word matching.

Vector Databases in GIS

- Applications:
    - Match **place names** or **street addresses**.
    - Interpret **map labels**.
    - Link geographic features with related descriptions.
- Especially useful for:
    - **Messy or historical data** where exact matches are difficult.

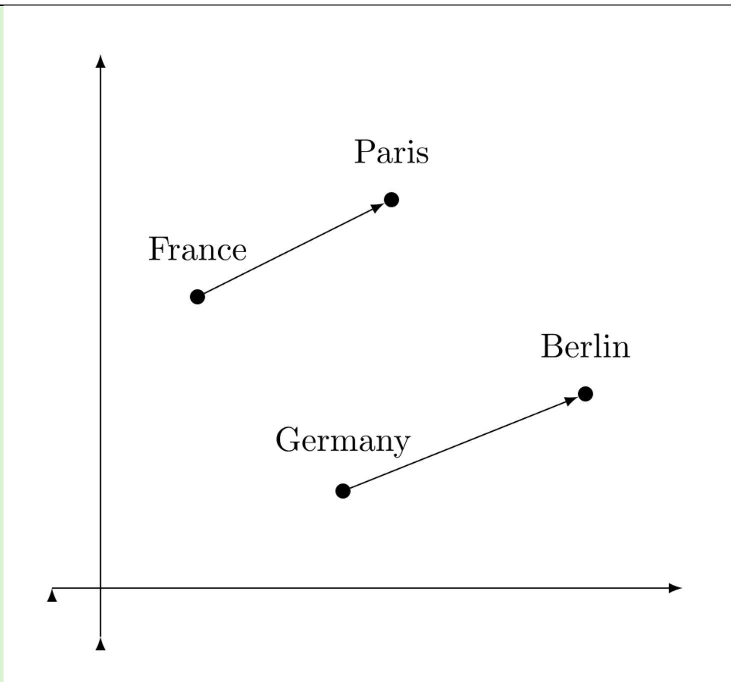


Figure 5 Source: https://en.wikipedia.org/wiki/Word_embedding

Support Vector Machine (SVM)

A Support Vector Machine is a type of machine learning algorithm used for classification and regression. It works by finding the best boundary (or "hyperplane") that separates different classes of data. Think of it as drawing a line that divides two groups of points as clearly as possible.

Geospatial examples:

- Classifying land cover types (forest, water, urban) from satellite imagery using spectral signatures.
- Identifying crop types in agricultural fields based on multispectral data.
- Detecting urban vs. rural areas in aerial photographs.

Geospatial Resources:
https://community.esri.com/t5/esri-technical-support-blog/classifying-images-in-arcgis-for-desktop-10-4/ba-p/901086

Synthetic Data

Synthetic data are artificially generated records (text, images, sensor streams) that statistically resemble real data without exposing sensitive details. In ML they augment scarce classes and improve robustness. Care is needed to avoid introducing bias or unrealistic artefacts.

Geospatial Examples:
When training object-detection models in ArcGIS Pro, teams often augment training data (tile size/stride/rotations) to simulate variety—akin to synthetic generation—to boost performance for small or rare features.

Geospatial Resources: https://www.esri.com/arcgis-blog/products/arcgis-pro/geoai/tips-for-training-data-preparation-for-object-detection-models

Indigenous Data

Indigenous Data refers to data about Indigenous Peoples, lands, and knowledge, and *Indigenous Data Sovereignty* asserts their rights to govern that data. Ethical GIS practice respects sovereignty, consent, and cultural protocols. Mapping should avoid erasure and promote visibility.

Geospatial Examples:
Esri highlights Indigenous data sovereignty and place-name restoration, plus resource hubs supporting Native Nations' geospatial governance and open-data efforts.

Geospatial Resourcess:
• ArcNews article on Indigenous place-names & data sovereignty: https://www.esri.com/about/newsroom/arcnews/putting-indigenous-place-names-and-languages-back-on-maps
• Esri Native Nations resources: https://www.esri.com/en-us/industries/tribal/overview

Data Processing & Integration

Model Context Protocol (MCP)

A standardised **communication protocol** designed to help AI models especially large language models interact with external tools, services, or data sources. It acts like a “universal connector”, enabling models to:

- Read files, query databases, fetch live data
- Execute functions and leverage APIs
- Handle and share context across tools and workflows

Developed by **Anthropic** and later open-sourced, MCP is gaining traction from providers like OpenAI, Microsoft Azure, and Google DeepMind.

Geospatial Resources:

- https://www.linkedin.com/posts/grant-s-carroll_esri-ai-geoai-activity-7364023419547353089-HRjQ/
- https://modelcontextprotocol.io/docs/getting-started/intro
- https://community.esri.com/t5/geodev-germany-blog/unser-r%C3%BCckblick-auf-den-esri-european-dev-amp-tech/ba-p/1668160
- https://www.linkedin.com/posts/martenhogeweg_olympics-mcp-esri-share-7431037870704078848-hV_B?utm_source=social_share_send&utm_medium=android_app&rcm=ACoAAAHZ_ooB6eHwedGMDc5KKh9coNpHrO4MKbg&utm_campaign=copy_link

Agent to Agent (A2A)

The **A2A protocol (Agent-to-Agent or Agent2Agent)** is an **open, standardised communication protocol** that enables autonomous AI agents, built by different vendors, using different frameworks, and running on different platforms,to **discover, authenticate, exchange information, and collaborate** with one another. Originally introduced by Google in April 2025, it is now an open-source project under the Linux Foundation. Its purpose is to provide a universal “language” for multi-agent systems, allowing agents to coordinate tasks, share state updates, and interoperate securely across organisational and technological boundaries. The protocol uses familiar web technologies such as **HTTP/HTTPS, JSON-RPC 2.0, and Server-Sent Events**, supports enterprise-grade security (OAuth 2.0, API keys, mTLS), and is designed to solve integration challenges that occur when agents must work together in complex enterprise workflows.

Resource:

https://developers.googleblog.com/en/a2a-a-new-era-of-agent-interoperability/

Geospatial Example:

A specialised AI Agent has been developed that is able to run specific analysis tasks for district plan analysis. This agent advertises its capabilities so that other agents can find it and if required use it within their workflows. This agent can then be placed within a large multi-agent system and when a request comes in from a parent agent for district plan analysis, the two agents can "talk" to each other to run the analysis.

-

Big Data

Big data refers to extremely large and diverse datasets whose **volume, velocity, and variety** exceed the capabilities of traditional data-processing systems. AI systems often rely on big data to identify subtle patterns, train complex models, and produce robust predictions. These datasets usually require distributed computing frameworks and advanced analytical pipelines to process efficiently.

Geospatial Examples:
ArcGIS GeoAnalytics tools to sift through billions of GPS traces to detect traffic hotspots or urban mobility patterns.
Esri link:
https://esriaustralia.com.au/resources/data/big-data-and-gis

Data Fusion

Data fusion is the process of integrating data from multiple sources, such as sensors, databases, or imagery, to create a unified dataset that is more accurate and informative than any single source alone. This technique often involves aligning data with different resolutions, formats, or timeframes to improve decision-making and analysis.

Geospatial Example:

- In GIS, data fusion can combine satellite imagery with ground-based sensor data to enhance land cover classification accuracy.

Resources:

- https://www.esri.com/~/media/files/pdfs/library/brochures/pdfs/data-fusion-centers.pdf
- https://www.esri.com/arcgis-blog/products/arcgis-pro/analytics/spatial-data-science-using-arcgis-notebooks-blog-2-data-fusion-wrangling-and-preparation

Image Fusion

Image fusion involves merging two or more images from different sensors, spectral bands, or resolutions into a single image that retains the most relevant spatial and spectral information from each source. This improves visual interpretation and analytical performance in applications such as remote sensing.

Geospatial Example:

- In GIS, image fusion can combine thermal and optical imagery to provide better insights for environmental monitoring.

Pan-Sharpening

Pan-sharpening is a technique that integrates high-resolution panchromatic images with lower-resolution multispectral images to produce a single image with both high spatial detail and rich colour information. This process is widely used to improve the clarity and usability of satellite imagery.

Geospatial Example:

- In GIS, pan-sharpening is commonly applied to enhance urban mapping and infrastructure analysis.

Resources:

- https://support.esri.com/en-us/gis-dictionary/pan-sharpening
- https://pro.arcgis.com/en/pro-app/latest/help/analysis/raster-functions/fundamentals-of-pan-sharpening-pro.htm
- https://support.esri.com/en-us/knowledge-base/how-to-increase-the-spatial-resolution-of-multispectral-000031853

Atmospheric Correction

Atmospheric correction refers to the removal of atmospheric effects, such as haze, scattering, and absorption, from remotely sensed imagery to retrieve accurate surface reflectance values. This step is essential for ensuring reliable quantitative analysis of satellite data.

Geospatial Example:

- In GIS, atmospheric correction is crucial for calculating vegetation indices when assessing crop health.

Cloud Masking

Cloud masking is the process of detecting and excluding cloud-covered pixels from satellite imagery to prevent errors in analysis. Clouds can obscure land features and distort spectral readings, making this step vital for accurate remote sensing.

Geospatial Example:

- In GIS, cloud masking is necessary for producing reliable land cover maps from optical satellite data.

Resources:

- https://www.esri.com/arcgis-blog/products/arcgis-pro/imagery/clean-up-your-landsat-imagery-removing-cloud-and-cloud-shadow
- https://www.esri.com/arcgis-blog/products/arcgis-online/imagery/create-cloud-free-imagery-composite-from-landsat-in-living-atlas

Noise Reduction

Noise reduction involves applying techniques to minimise random variations or unwanted signals in data, improving image quality and analytical accuracy. Noise can originate from sensor limitations, transmission errors, or environmental factors.

Geospatial Example:
- In GIS, noise reduction is often used on radar imagery to enhance terrain mapping in complex landscapes.

Data Augmentation

Data augmentation is the process of artificially expanding a dataset by applying transformations such as rotation, scaling, flipping, or adding synthetic variations. This technique is commonly used in machine learning to improve model robustness and performance.

Geospatial Example:
- In GIS, data augmentation can help train AI models for land classification when only limited satellite imagery is available.

Data Imputation

Data imputation refers to filling in missing or incomplete data using statistical methods, interpolation, or machine learning techniques to maintain dataset integrity and usability. This is particularly important for large datasets where gaps can affect analysis.

Geospatial Example:
- In GIS, data imputation can estimate missing elevation values in a digital terrain model for hydrological studies.

Spatiotemporal Modelling

Spatiotemporal modelling is the analysis and prediction of phenomena that vary across both space and time, often using advanced statistical or AI-based approaches. It enables understanding of dynamic patterns and trends in geographic data.

Geospatial Example:
- In GIS, spatiotemporal modelling can forecast urban growth patterns based on historical land use and population data.

AI Infrastructure & Deployment

Cloud AI

Cloud AI uses cloud computing resources to train, deploy, and manage AI models, offering scalability and cost efficiency without relying on local hardware.

Geospatial Examples:

- Running a cloud-based AI service to detect deforestation patterns from global satellite imagery.
- Using cloud-hosted machine learning APIs to classify terrain types in large geospatial datasets.

Edge AI

Edge AI refers to running AI models on local devices (e.g., sensors, drones, mobile devices) rather than in the cloud, enabling real-time processing with minimal latency.

Geospatial Examples:

- Deploying AI on drones to identify crop health during aerial surveys without needing internet connectivity.
- Using edge AI on handheld GPS devices to classify soil types in the field instantly.

GPU Acceleration

GPU acceleration uses Graphics Processing Units (GPU) to speed up AI computations, particularly for deep learning tasks that involve large datasets and complex models.

Geospatial Examples:

- Accelerating training of convolutional neural networks for satellite image segmentation.
- Using GPUs to process LiDAR point clouds for terrain modelling and feature extraction.

Distributed Training

Distributed training splits AI model training across multiple machines or GPUs to handle large datasets and reduce training time.

Geospatial Examples:

- Training a global climate prediction model using distributed computing across multiple servers.
- Running distributed training for urban feature detection using petabytes of aerial imagery.

Federated Learning

Federated learning allows multiple devices or organisations to collaboratively train AI models without sharing raw data, preserving privacy.

Geospatial Examples:

- Collaborating across municipalities to train a flood prediction model without sharing sensitive local data.
- Using federated learning to improve traffic prediction models across different cities without centralising GPS data.

Model Deployment

Model deployment is the process of integrating a trained AI model into a production environment where it can be accessed and used by applications or end-users.

Geospatial Examples:

- Deploying a trained AI model into a GIS platform to automatically classify land use in uploaded imagery.
- Integrating a predictive model into a web-based mapping application for real-time hazard alerts.

Inference

Inference is the stage where a trained AI model makes predictions or classifications on new, unseen data.

Geospatial Examples:

- Running inference on satellite images to detect illegal mining sites.
- Using inference to predict flood-prone areas based on current weather and terrain data.

Causal inference

Causal inference seeks to determine whether one factor *causes* an outcome, rather than merely being correlated with it.

Geospatial Resource:

- https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/how-causal-inference-analysis-works.htm

Data Pipelines

In AI, **data pipelines** are automated processes that ingest, clean, transform, store and deliver data so that machine-learning models can be reliably trained and used in production. They typically include stages such as data ingestion from sources like databases or sensors, storage in cloud data lakes, processing with tools such as Spark or Dataflow, and preparation of features for training and inference. These pipelines ensure that data is consistent, high-quality and available for real-time or batch AI workloads—including applications like computer-vision workflows, retrieval-augmented generation, IoT analytics and geospatial processing—ultimately acting as the production line that makes AI systems usable and dependable.

Geospatial resource:
https://www.esri.com/en-us/arcgis/products/arcgis-data-pipelines/overview

Pipeline Orchestration

Pipeline orchestration involves automating and managing the sequence of tasks in an AI workflow, such as data preprocessing, model training, and deployment.

Geospatial Examples:

- Orchestrating a pipeline that ingests satellite imagery, cleans data, trains a model, and updates a GIS dashboard.
- Automating workflows for continuous monitoring of urban sprawl using AI and geospatial data.

Emerging & Research-Frontier Terms

Foundation Model

A foundation model is a very large, pre-trained AI model that has learned from massive amounts of general data. It can then be adapted (fine-tuned) for specific tasks without starting from scratch. Think of it as a “base layer” of knowledge that can be customised.

Geospatial examples:

- Fine-tuning a vision foundation model to classify land cover types from satellite imagery.
- Adapting a language foundation model to extract location names from environmental reports.

Resource:
https://www.esri.com/en-us/arcgis/deep-learning-models

Emergent Behaviour

Emergent behaviour occurs when AI systems unexpectedly exhibit new capabilities or patterns not explicitly programmed, often arising from scale, training complexity, or multi-agent interactions.

Synthetic Data

Synthetic data is artificially created data that mimics real-world data. It’s used to train AI models when real data is scarce, sensitive, or expensive to collect.

Geospatial examples:

- Generating extra labelled crop images to improve agricultural classification models.
- Creating simulated flood maps for training disaster response algorithms.

Resource:
https://community.esri.com/t5/education-blog/how-to-get-started-with-geoai-in-arcgis-a/ba-p/1675534

GAN (Generative Adversarial Network)

A GAN is a type of AI model that learns to generate realistic new data by pitting two networks against each other one creates fake data, the other tries to spot fakes. Over time, the generator gets very good at producing convincing outputs.

Geospatial examples:

- Creating realistic synthetic satellite imagery for areas with poor coverage.
- Generating cloud-free versions of images for remote sensing analysis.

Zero-Shot Learning

Zero-shot learning means an AI model can handle tasks or categories it has never seen before, without extra training. It uses general knowledge to make educated guesses.

Geospatial examples:

- Detecting a new crop type in satellite imagery without retraining the model.
- Identifying an unfamiliar building type in aerial photos using contextual clues.

Resources:

- https://www.esri.com/arcgis-blog/products/arcgis-pro/geoai/vision-language-models-geospatial-analysis
- https://www.esri.com/arcgis-blog/products/arcgis-pro/geoai/revolutionizing-image-segmentation-with-sam-segment-anything-model

Few-Shot Learning

Few-shot learning is when an AI model learns a new task from only a handful of examples, instead of thousands. It's useful when labelled data is hard to get.

Geospatial examples:

- Classifying wetlands with just a few labelled samples.
- Learning to recognise rare features like volcanic vents from limited imagery.

Meta-Learning

Meta-learning is "learning how to learn." It trains models to quickly adapt to new tasks by understanding patterns across many previous tasks.

Geospatial examples:

- A model that quickly adjusts to classify different land cover types in new regions.
- Adapting to new sensor data formats without extensive retraining.

Active Learning

Active learning is when an AI model asks humans for help on uncertain cases, so it learns faster and better. It focuses effort where it matters most.

Geospatial examples:

- A land cover classifier flags ambiguous pixels for expert review.
- A building detection model requests validation for unclear roof shapes.

Self-Supervised Learning

Self-supervised learning trains models on unlabelled data by creating its own "pseudo-labels" or tasks. It's a way to learn from huge datasets without manual labelling.

Geospatial examples:

- Pretraining on millions of unlabelled satellite images before fine-tuning for classification.
- Learning spatial patterns from raw LiDAR point clouds without labels.

Resource:
https://www.esri.com/arcgis-blog/products/arcgis-pro/imagery/whats-new-for-geoai-in-the-image-analyst-extension-of-arcgis-pro-3-4

NeRF (Neural Radiance Fields)

NeRF is an AI technique that reconstructs 3D scenes from 2D images by modelling how light interacts with surfaces. It creates realistic 3D views from photos.

Geospatial examples:

- Building 3D city models from aerial imagery for urban planning.
- Reconstructing terrain from drone photos for flood modelling.

Physics-Informed AI

Physics-informed AI combines physical laws with AI models to improve accuracy and realism. It ensures predictions make sense in the real world.

Geospatial examples:

- Embedding hydrological principles in flood prediction models.
- Using soil physics in drought forecasting algorithms.

Spatial Knowledge Graph

A spatial knowledge graph links geographic entities (places, features) and their relationships in a graph structure, making it easier for AI to reason about spatial connections.

Geospatial examples:

- Modelling road networks and their connectivity for traffic optimisation.
- Linking buildings, utilities, and land parcels for smart city planning.

Anthropomorphism

Anthropomorphism within the context of AI, is the attribution of human emotions, motivations, or cognitive traits to AI systems. This may influence user perceptions, trust levels, or expectations about AI capability.

Geospatial examples:

- May be used in assistants, particularly avatar assistants, or for training avatars.

Capsule network

A capsule network is a neural architecture that groups neurons into “capsules” which encode not only the presence of features but also their spatial relationships. It aims to overcome CNN limitations by preserving hierarchical structure. Capsules communicate via dynamic routing, allowing more robust recognition of rotated or distorted objects. These models show promise in structured visual understanding.

AI Risks and Security Considerations

The following examples indicate some of the challenges, risks and other considerations when dealing with AI.

Cybersecurity

Cybersecurity protects systems and data against unauthorised access, disruption, and misuse through controls such as encryption, identity, monitoring, and patching. For AI, it also covers model/feature security, prompt injection risks, and supply-chain transparency.

Geospatial Examples: The ArcGIS Trust Center documents security hardening, FedRAMP/ISO posture, and Trusted-AI materials to help GIS teams deploy secure, compliant AI features. **Geospatial Resources:** https://trust.arcgis.com/en/

Public AI/ML system

A *public* AI/ML system typically refers to cloud-hosted (multi-tenant) services managed by a vendor, offering rapid updates andr elasticity. Data residency, privacy, and opt-in controls are governed by the provider's trust programme.

Geospatial Examples: ArcGIS Online's AI assistants (beta/preview) can be enabled by an organisation admin for members to use with Trusted-AI transparency cards and privacy guardrails. **Geospatial Resourcess:** • Configure AI assistants (ArcGIS Online): https://doc.arcgis.com/en/arcgis-online/administer/configure-assistants.htm • Trusted AI overview & transparency cards: https://trust.arcgis.com/en/trusted-ai/trusted-ai.htm

Private AI/ML System

A *private* AI/ML system runs within your organisation's controlled environment (on-premises or single-tenant cloud), prioritising data control and custom integration over vendor-managed elasticity. It suits sensitive data and stricter compliance.

Geospatial Examples: Many organisations deploy ArcGIS Enterprise (including Image/GeoAI capabilities) in their own infrastructure and apply Trust Center security hardening and governance to manage AI-powered workflows. **Geospatial Resources:** https://links.esri.com/ArcGISTrustCenter (ArcGIS Trust Center—security, privacy, compliance)

AI Privacy Impact Assessment

An AI Privacy Impact Assessment (AI-PIA) evaluates how an AI feature collects, processes, stores, and shares data, and the privacy risks that may arise. It documents mitigations (data minimisation, opt-in, retention, human-in-the-loop) and references lawful bases and standards. In ArcGIS, this is supported by *Trusted AI* artefacts—such as AI Transparency Cards—detailing prompts, storage, subprocessors, and limitations. Organisations often include AI-PIAs in governance before enabling assistants.

Geospatial Examples:
Before turning on AI assistants in ArcGIS Online, administrators review transparency cards and Trust Center guidance to confirm privacy and security posture.

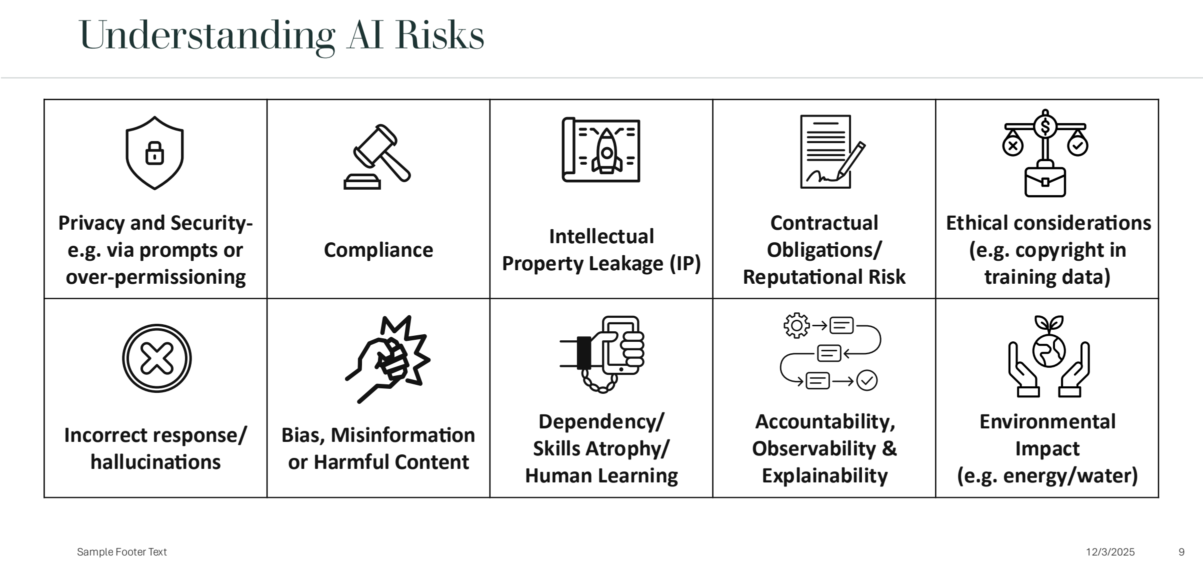
Geospatial Resourcess:
• Transparency cards list: https://trust.arcgis.com/en/trusted-ai/ai-transparency-cards-list.htm
• Transparency card structure (what's disclosed): https://trust.arcgis.com/en/trusted-ai/ai-transparency-cards.htm
https://content.esri.com/resources/enterprisegis/advancing_trusted_ai_in_arcgis.pdf

Deepfakes

AI-generated or AI-manipulated synthetic media—typically images, audio, or video—that convincingly depict people saying or doing things they never actually did, often created using deep-learning models such as GANs (Generative Adversarial Networks).

Misinformation (in the context of AI)

False or misleading information that an AI system may unintentionally generate, amplify, or spread, typically without malicious intent but still capable of causing confusion or harm if not identified or corrected.



The framework in the image above indicates some of the main areas of risks from using AI, which are detailed in the sections below.

Privacy and Security

- *Prompt leakage*: A consultant pastes an entire client RFP (with pricing and proprietary methodologies) into a public LLM to draft a response; the data becomes part of model training data or stored prompts.
- *Shadow tools*: Staff use unapproved “AI data cleaners” that silently upload CSVs containing PII and sensitive geospatial layers to a third-party server.
- *Government access*: Sensitive datasets are processed in a cloud region subject to foreign legal compulsion (e.g., extraterritorial discovery) without sufficient client consent.
- *Data sovereignty violations*: Training pipelines replicate domestic datasets to offshore regions for finetuning, breaching contractual residency obligations or community expectations.
- *Over-permissioned agents*: An AI integration has tenant-wide read access and inadvertently surfaces HR records or confidential board minutes in a user’s chat due to weak filtering.

Compliance

- *Copyright exposure*: The AI generates a map design that replicates a licensed cartographic style (or copyrighted basemap assets) without attribution or a valid license, leading to infringement claims.
- *Regulatory overreach*: Using AI to assess eligibility (e.g., public services or grants) without mandated consent, audit logs, or human review—potentially violating Privacy Act/GDPR principles.
- *Sector rules breach*: Generating health-related insights with an AI tool lacking appropriate controls; outputs are stored in general-purpose collaboration spaces.
- *Untracked datasets*: Personal data piped into AI experimentation notebooks without records of processing activities.

Intellectual Property Leakage (IP)

- *Cross-client contamination*: Internal bid strategy documents are used to “teach” an LLM assistant, and then similar language or ideas appear in a different client’s deliverable.
- *Model training misuse*: Vendor finetunes on proprietary GIS scripts, making bespoke automation logic identifiable via probing prompts.

Contractual Obligations

- *Unapproved subprocessors*: A team embeds a third-party AI plugin in deliverables without notifying the client—breaching contract terms on subcontracting/data processors.
- *Residency and retention violations*: Project data processed and cached outside the agreed geography; outputs kept indefinitely in LLM logs contrary to retention clauses.
- *Undisclosed AI use*: AI-authored sections in a technical report are not flagged, breaching transparency commitments in the MSA/SOW.

Reputational Risk from Overuse or Inappropriate Use of AI

- *Client does not expect AI to be used*: For some activities, such as recommendations reports, many clients expect that they will receive a report tailored to their specific situation, and that the report will have been considered and written by an expert human. If they then discover that an LLM contributed to the report (especially if errors are discovered) then this can create a very negative perception of the work.
- *AI-flavored deliverables*: Client receives a generic, shallow strategy that reads like a template—perceived as low-effort, harming credibility.
- *Cultural insensitivity*: AI produces recommendations about indigenous datasets without contextual nuance or consultation—seen as extractive or disrespectful.
- *Automated comms misfires*: AI-scheduled outreach spams stakeholders with poorly targeted updates, damaging trust and engagement.

Ethical Considerations

(e.g., copyrighted material in training data, or indigenous data sovereignty)

- *Opaque provenance*: Using models trained on copyrighted content (books, imagery) without clear licensing leads to ethical and legal concerns when outputs mimic source material.
- *Indigenous data misuse*: Training a model on datasets involving Māori taonga or culturally sensitive locations without co-governance, consent, or benefit-sharing; outputs commercialize knowledge improperly.
- *Exploitative analytics*: AI-driven resource allocation avoiding community consultation, reinforcing inequities or ignoring cultural obligations.

Incorrect Response / Hallucinations

- *Fabricated citations*: AI generates references or datasets that don't exist in a planning proposal, which are then included in a draft.

- *Invented geospatial features*: AI labels regions with non-existent hazards or misinterprets coordinate systems, leading to faulty maps.
- *Confident nonsense*: AI asserts a regulation applies when it doesn't (or is outdated), misguiding compliance decisions.

Bias, Misinformation, or Harmful Content

- *Spatial bias*: AI recommendations overweight urban data, under-serving rural or indigenous communities due to training data coverage skew.
- *Stereotyping in outputs*: AI-generated community profiles use biased language or assumptions, harming relationships and accuracy.
- *Amplified misinformation*: AI summarizes social media trends about infrastructure risks that are unverified, influencing planning falsely.
- *Attacks on a person*: AI image and video generators can be used to take a person's likeness and create images or videos depicting that person in faked situations or activities that are harmful. AI can be used to direct trolling bots to attack a person via social media.

Dependency on AI / Skills Atrophy

- *Over-reliance*: Analysts stop practicing key GIS fundamentals (coordinate transformations, topology checks), trusting AI fixes that occasionally skew datasets.
- *Shallow problem-solving*: Teams default to AI for ideation, reducing creative and critical thinking in solution design.
- *Tool-first mentality*: Project planning prioritizes "what AI can do" rather than rigorous requirements analysis or stakeholder discovery.

Human Learning (including for junior staff)

- *Shortcut learning*: Juniors ask AI for every step of ETL, bypassing understanding of data quality, projection decisions, and QA practices.
- *Lost tacit knowledge*: Mentors rely on AI review tools, reducing hands-on code walks, cartography critique, or field validation.
- *Portfolio gaps*: Junior staff lack evidence of independent reasoning—work is AI-assisted with minimal personal analysis.

Accountability from Using AI

- *Responsibility diffusion*: A poor recommendation is blamed on the model; no single owner for validation or sign-off.

- *No escalation pathways*: An AI error affecting a client isn't triaged; incident response is ad hoc, causing delays and compounding harm.
- *Ambiguous governance*: Decisions to deploy new prompts/agents occur without change control, bypassing risk assessment and approvals.

Observability of AI Models

- *Invisible drift*: Over time, retrieval sources change and accuracy drops, but there's no telemetry or alerting.
- *Unlogged access*: Prompts and responses aren't recorded; investigators cannot reconstruct how sensitive data surfaced.
- *Data lineage gaps*: Outputs cannot be traced back to source versions—making audits or corrections nearly impossible.

Explainability of AI Models and Outputs

- *Black-box decisions*: Risk scoring across regions lacks feature importance or rationale; clients cannot challenge outcomes.
- *No uncertainty bands*: Generated maps and forecasts present point estimates without confidence levels; misleads decision-makers.
- *Opaque trade-offs*: Complex model ensembles produce recommendations without counterfactuals; stakeholders lack clarity.

Environmental Impact (e.g., energy/water)

- *Compute sprawl*: Teams run multiple large finetunes and daily batch inference without governance; cloud costs and emissions balloon.
- *Water usage blindness*: Workloads tied to data centers with high WUE; environmental commitments are undermined.
- *Inefficient pipelines*: Redundant embeddings and unoptimized RAG indexing cause excessive energy consumption.

INDEX

Researchers for this eBook

Abby Heazlewood is a Communications Intern at Eagle Technology based in Auckland, New Zealand. Abby is currently studying Politics & International Studies and Media & Communication at the University of Melbourne. Abby had not used AI prior to creating this document and was interested to learn more about these types of tools.

Nathan Heazlewood is a Principal Consultant and Business Consulting Practice Leader at Eagle Technology. Nathan has more than 30 years experience working in the GIS domain. In his current role Nathan assists clients with developing geospatial strategies. Many organisations are currently struggling to formulate their strategies for AI and GIS, and many GIS Professionals are wondering how to adapt their career plans to the rise of AI, and these are topics that Nathan is able to assist with.

Robert Mi is an Auckland-based geospatial professional specialising in GIS, remote sensing, and data analytics, with a focus on applying emerging technologies like drones to real-world problems. He combines strong technical skills with consulting experience, contributing to projects that improve spatial decision-making and delivering innovative solutions within the geospatial industry.

Vipul Bhatia is a Senior GIS Consultant specialising in ArcGIS Enterprise, with expertise in solution architecture, system design, and enterprise implementations. He provides technical leadership across projects, supported by strong certifications and experience delivering GIS solutions for organisations across New Zealand.

AI Usage Statement

This document was developed with the assistance of artificial intelligence (AI) tools. These tools were used to support drafting, editing, and refining content. All outputs generated by AI have been reviewed, validated, and, where necessary, modified by the author to ensure accuracy, relevance, and appropriateness.