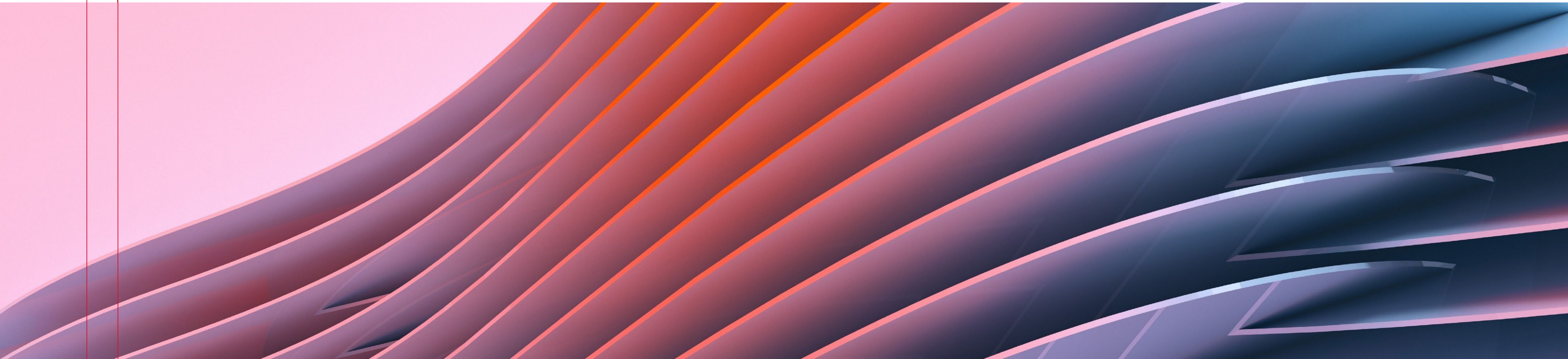
Kilpatrick

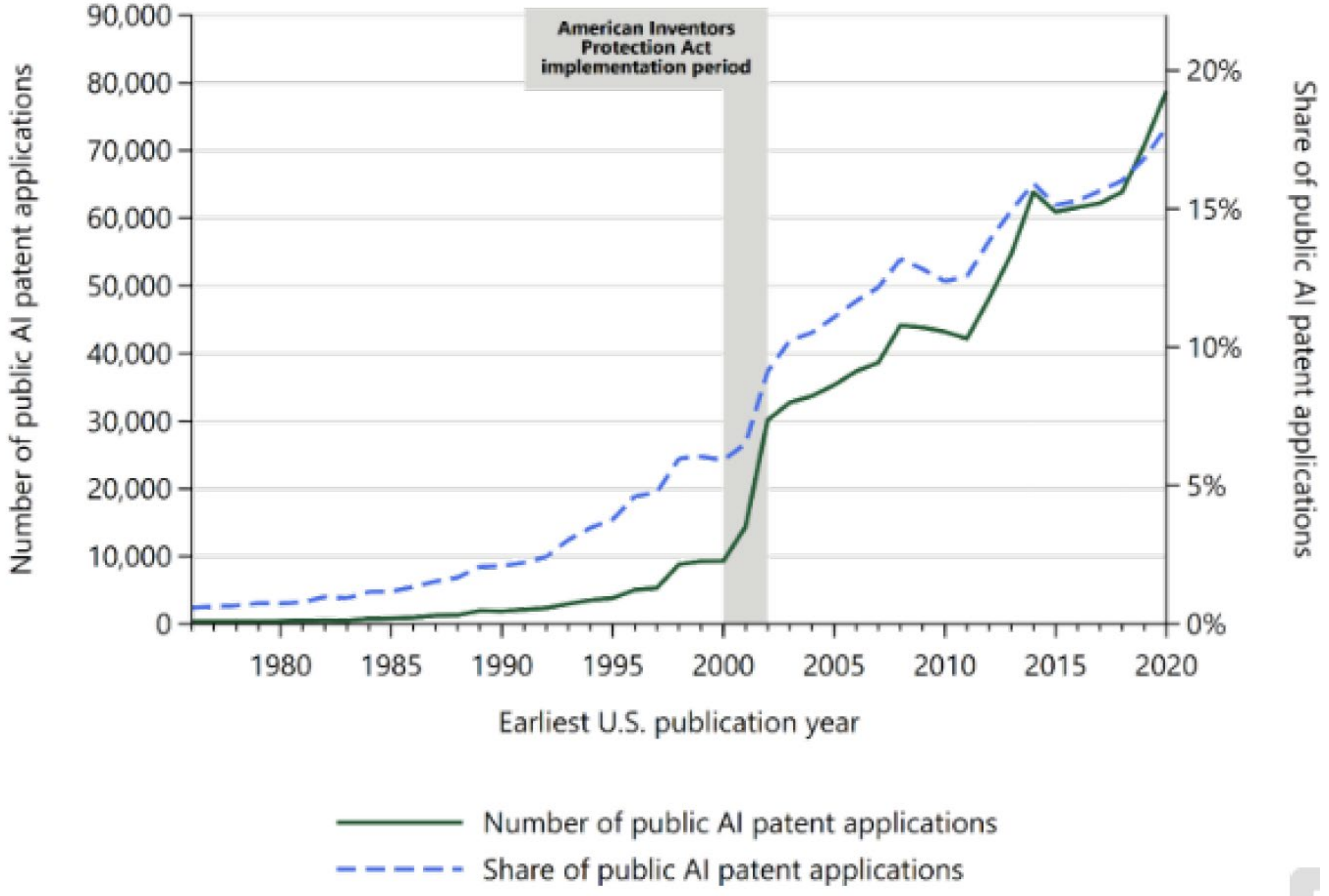
SEPTEMBER 18, 2024

PRESENTED BY DAVID RACZKOWSKI

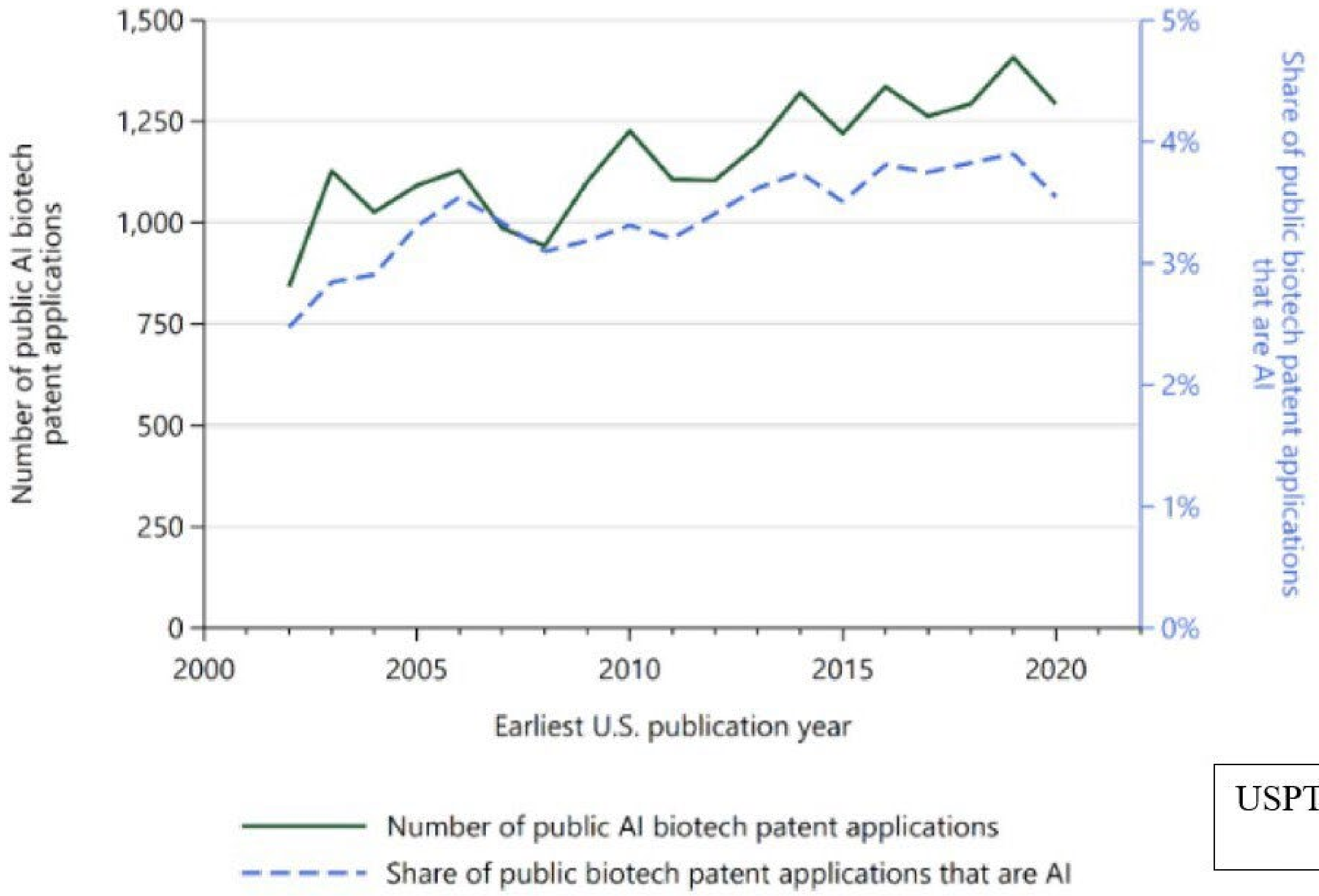
Machine learning models and uses in biology



Volume and share of public U.S. AI patent applications, 1976–2020

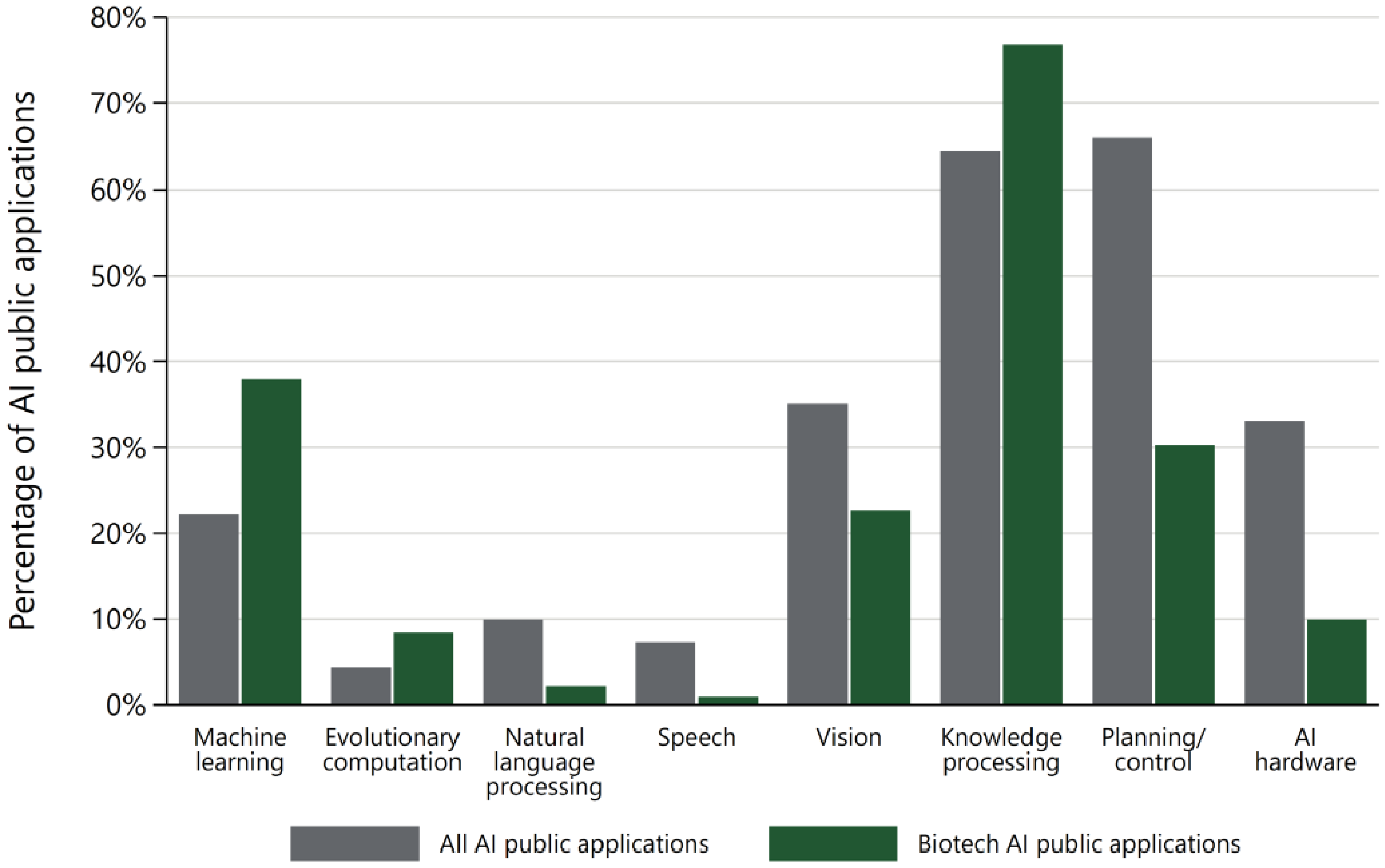


Volume and Share of Public U.S. Biotech AI Patent Applications, 2002-2020



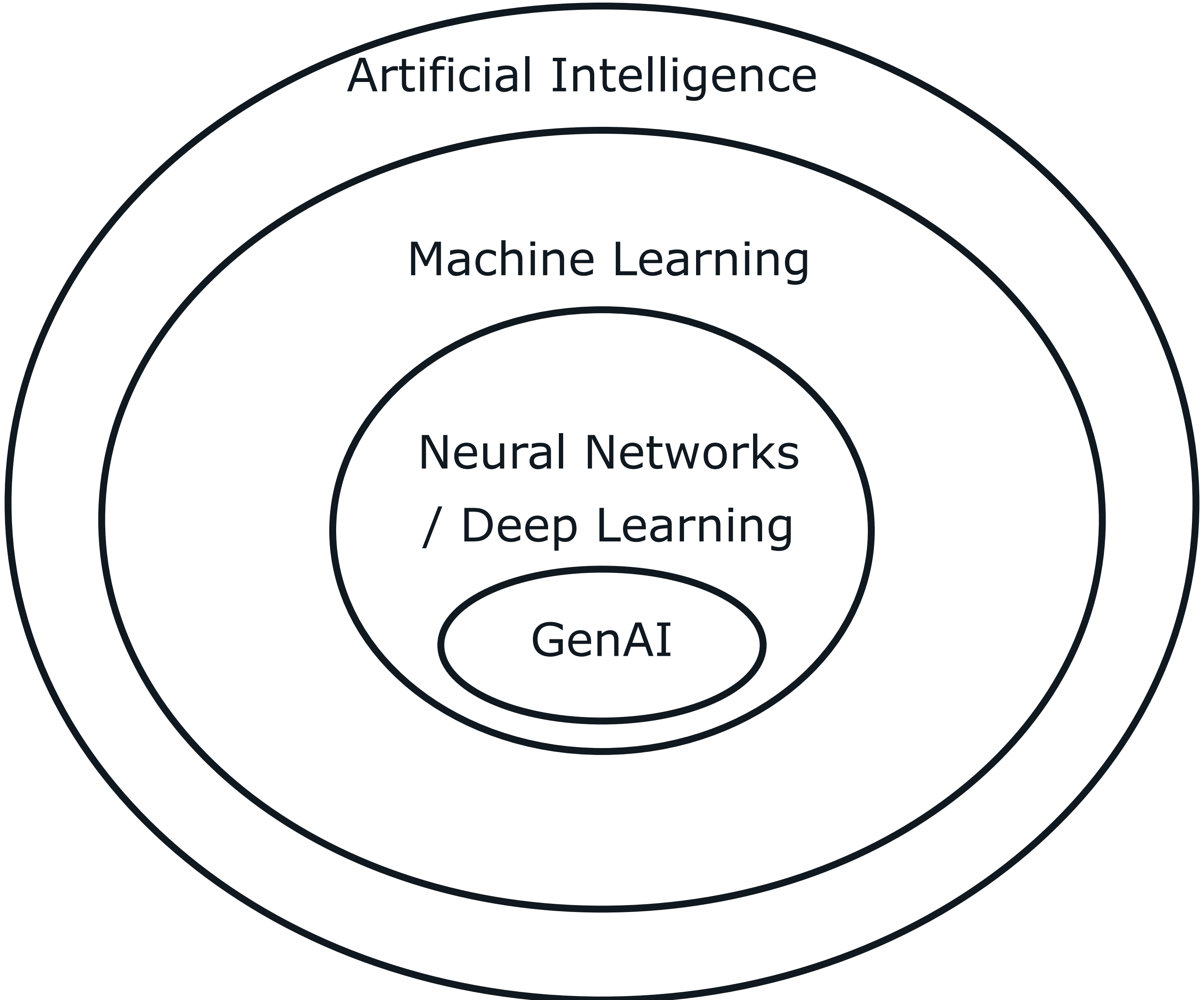
USPTO Data

Share of AI and Biotech AI Public Applications by AI Technology Component, 2020



ML technqiues

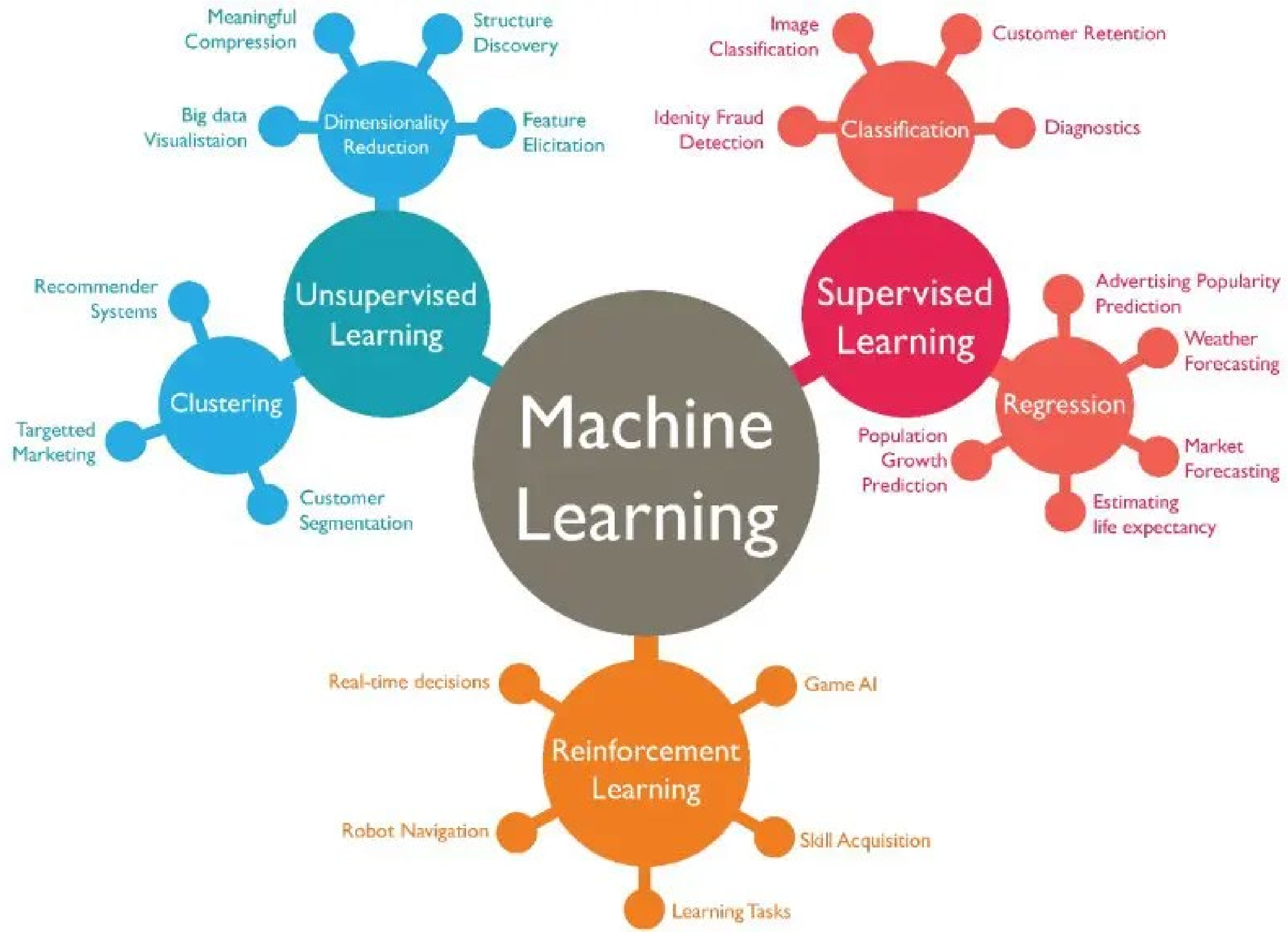
AI vs. Machine Learning vs. Neural Networks



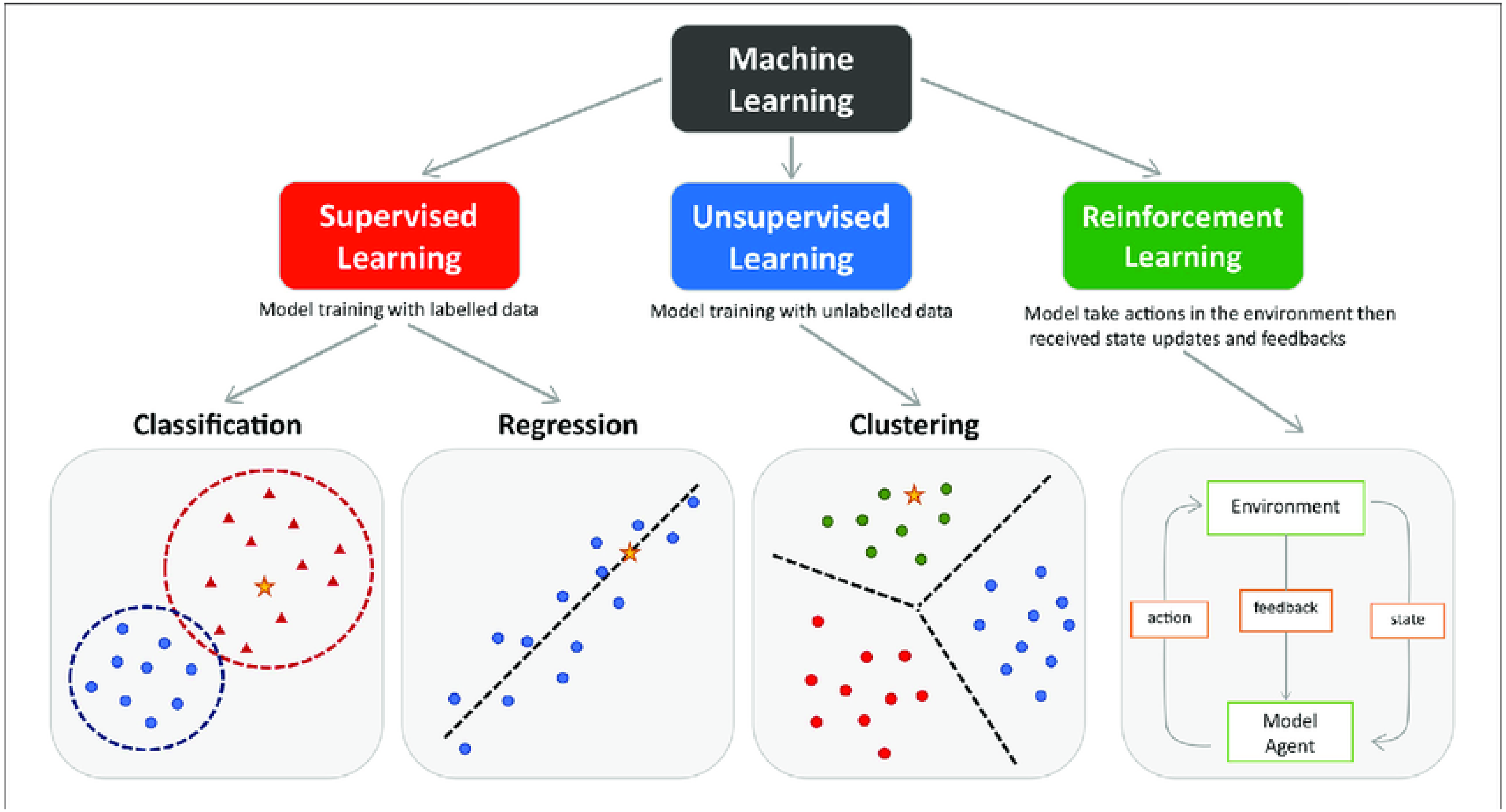
Types of Machine Learning

- **Machine Learning (ML):** A subset of AI that allows models to **learn and adapt** without following explicit instructions, by analyzing and drawing inferences from patterns in data.
- **Unsupervised learning**: Learns patterns exclusively from unlabeled data
  - Ex: cluster analysis (hierarchical, k-means, mixture models), dimensionality reduction (principal component analysis (PCA)), auto-encoder
- **Supervised Learning**: Input objects and a desired output value train a model using labeled data
  - Ex: neural network, support vector machines (SVM), decision trees (random forest), linear regression, logistic regression
- **Semi-supervised learning**: Combines supervised and unsupervised learning by using both labeled and unlabeled data to train artificial intelligence (AI) models for classification and regression tasks
- **Reinforcement Learning**: Combination with optimal control, concerned with how an intelligent agent takes actions in a dynamic environment to maximize cumulative reward.
  - Ex: Markov decision process (MDP), Dynamic programming, Value iteration, Policy iteration. Q-learning.

Types of Machine Learning

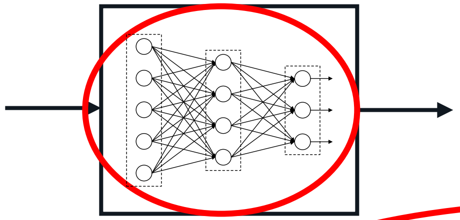


Types of Machine Learning

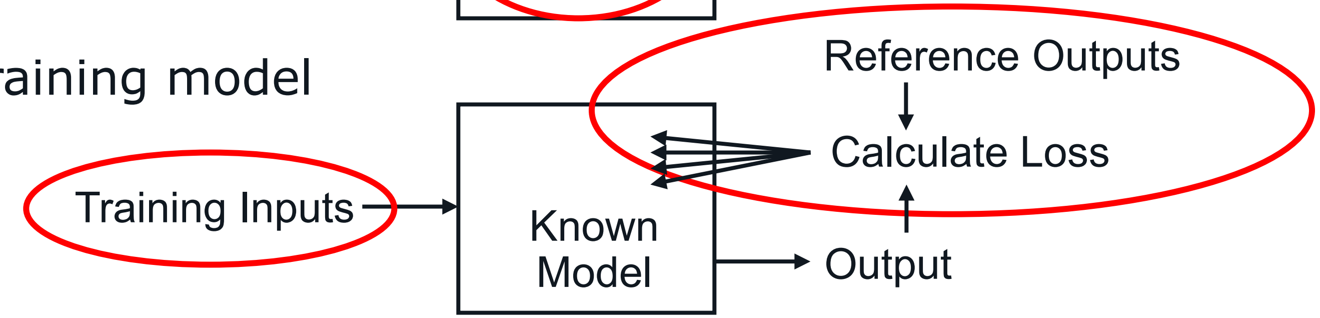


Types of Machine Learning Inventions

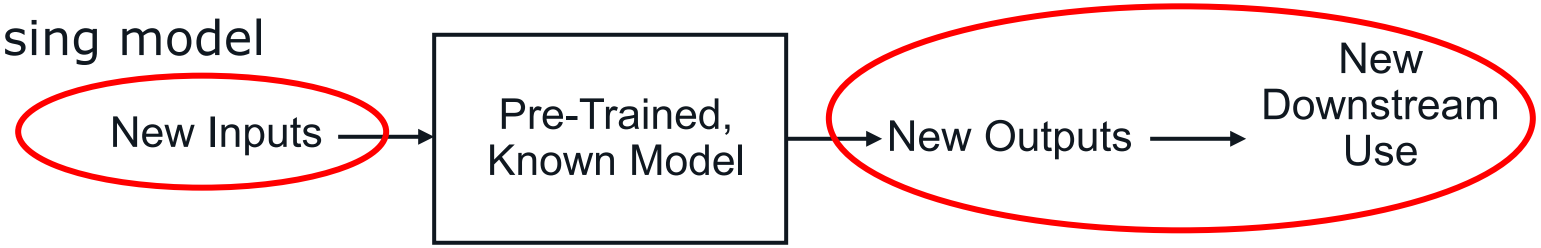
- Type 1 – New model



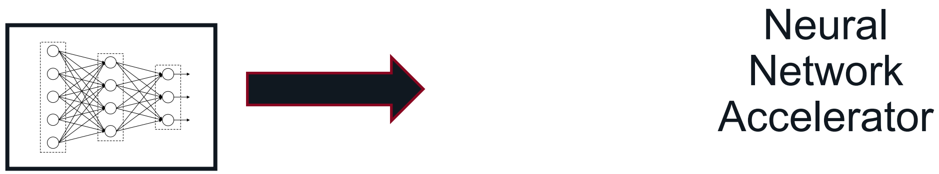
- Type 2 – New method of training model



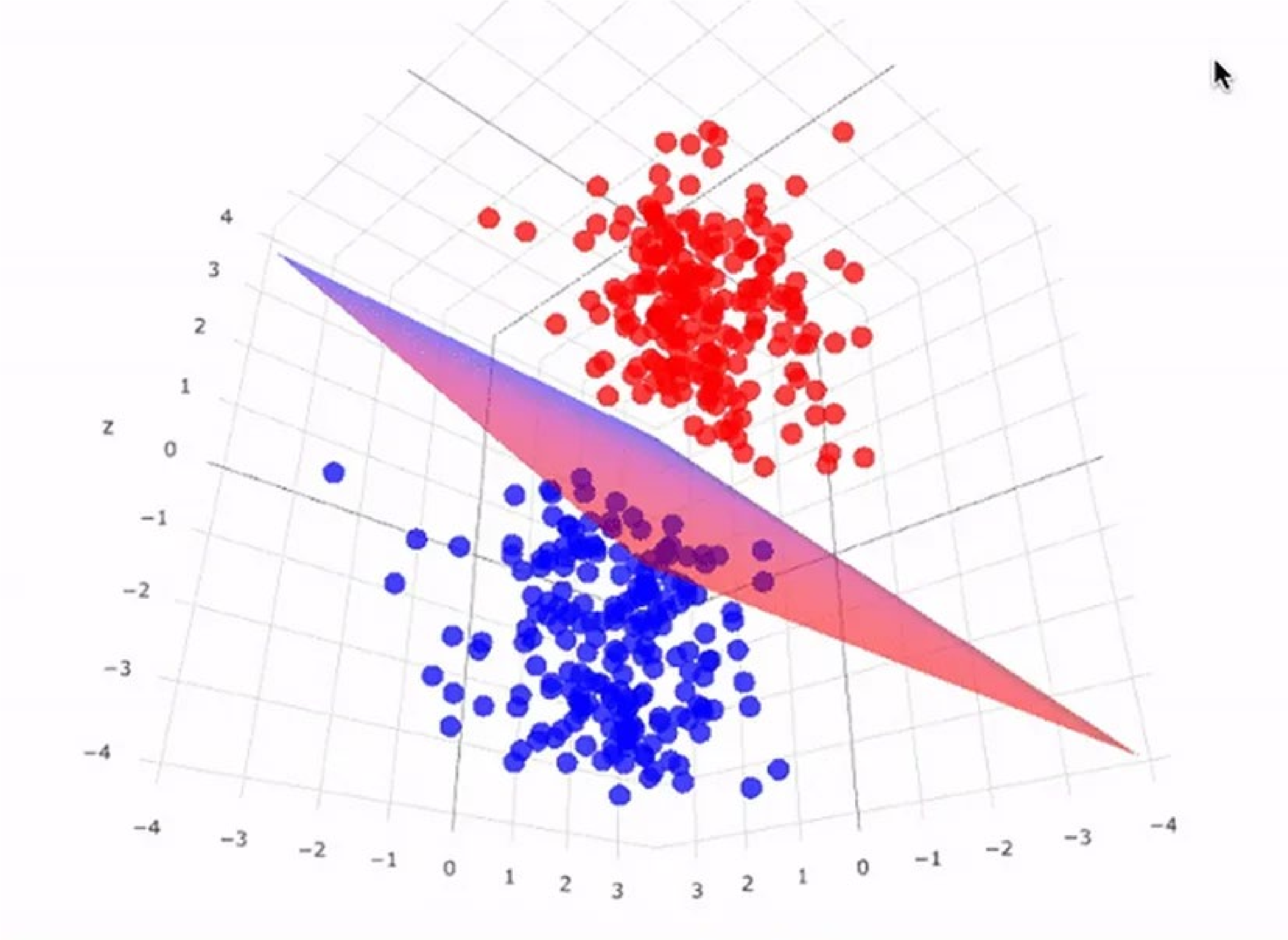
- Type 3 – New method of using model



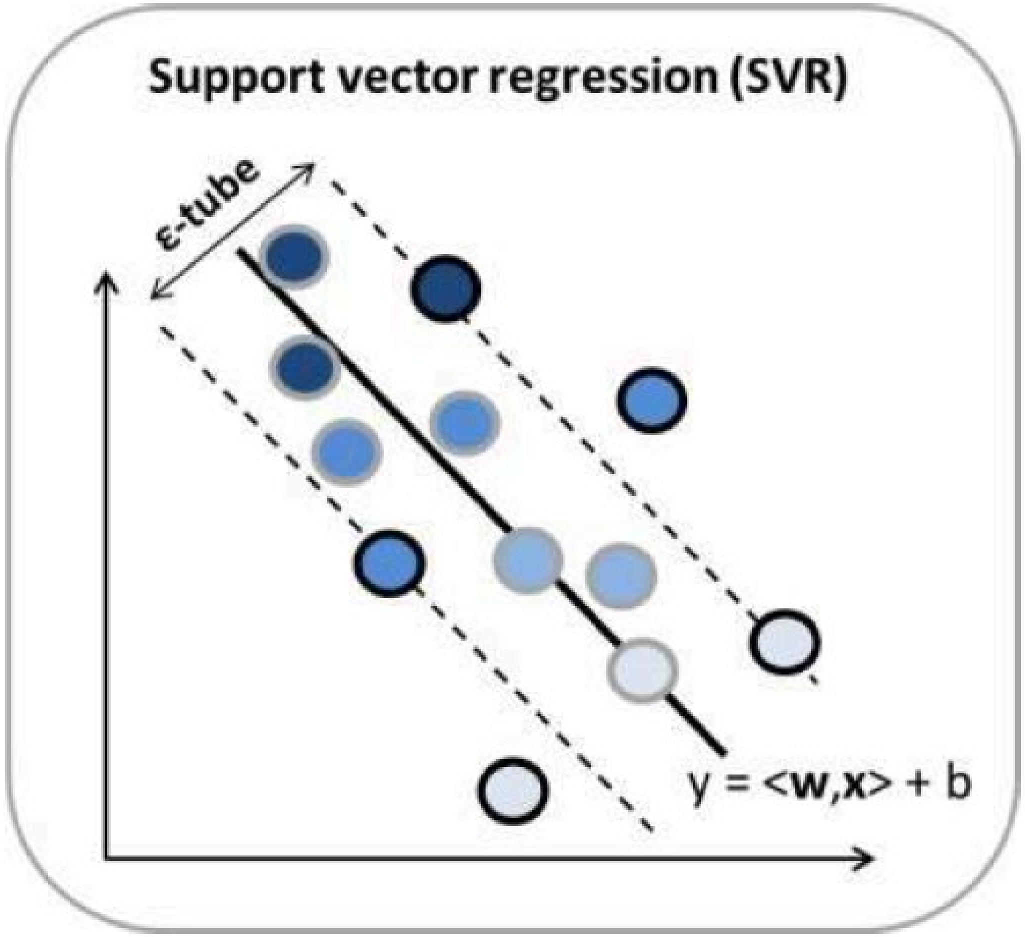
- Type 4 – New hardware for runtime/training



Support Vector Machines

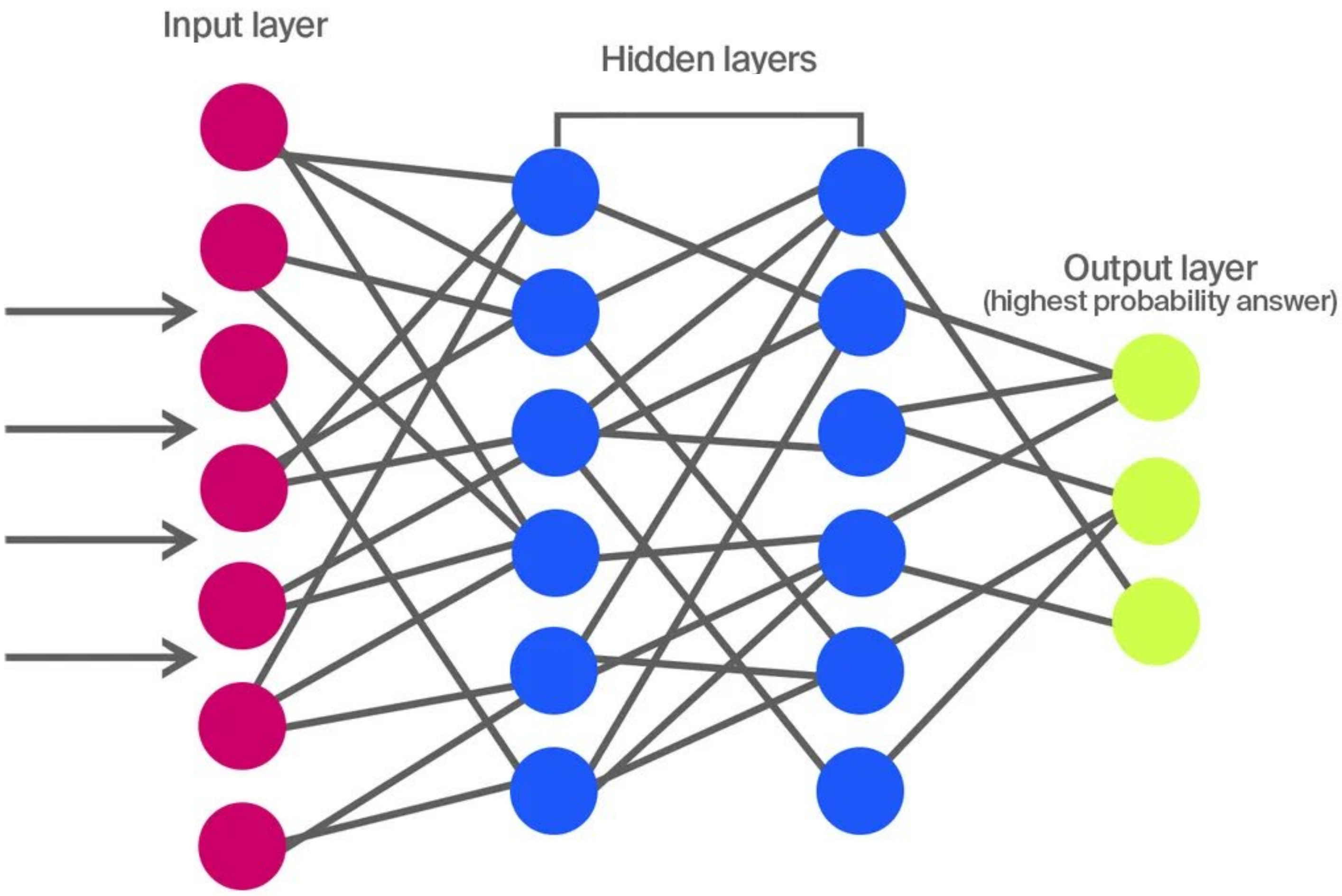


Support Vector Regression

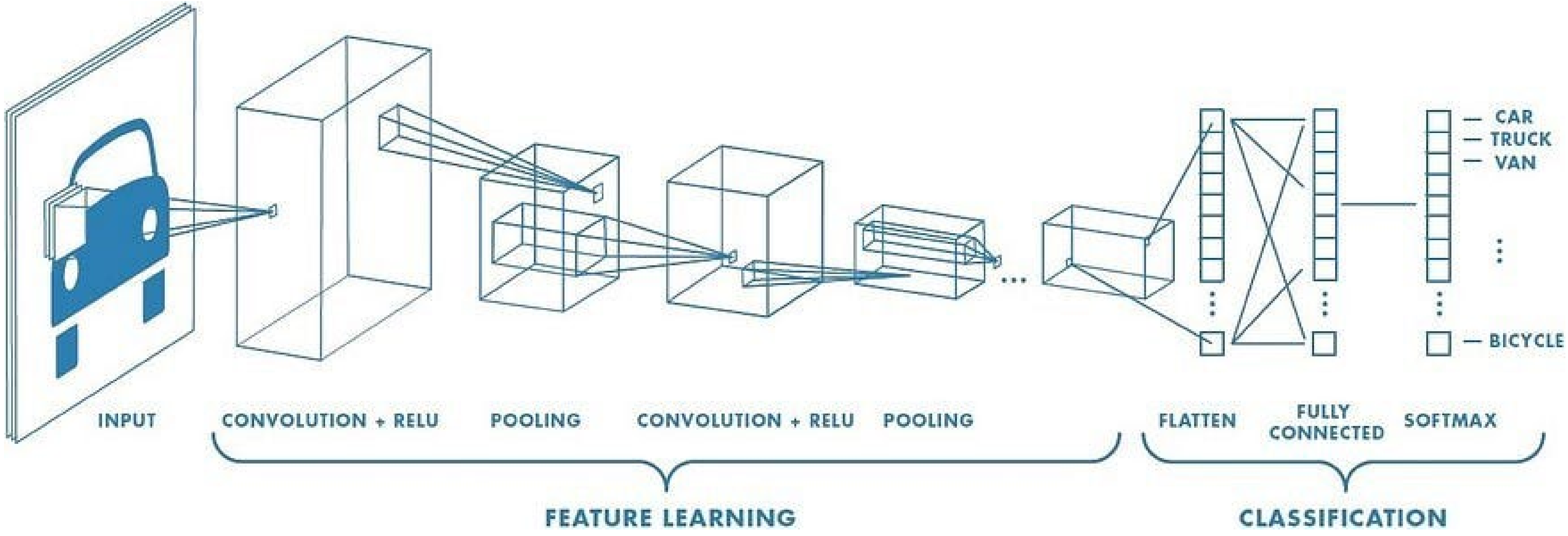


Neural Network

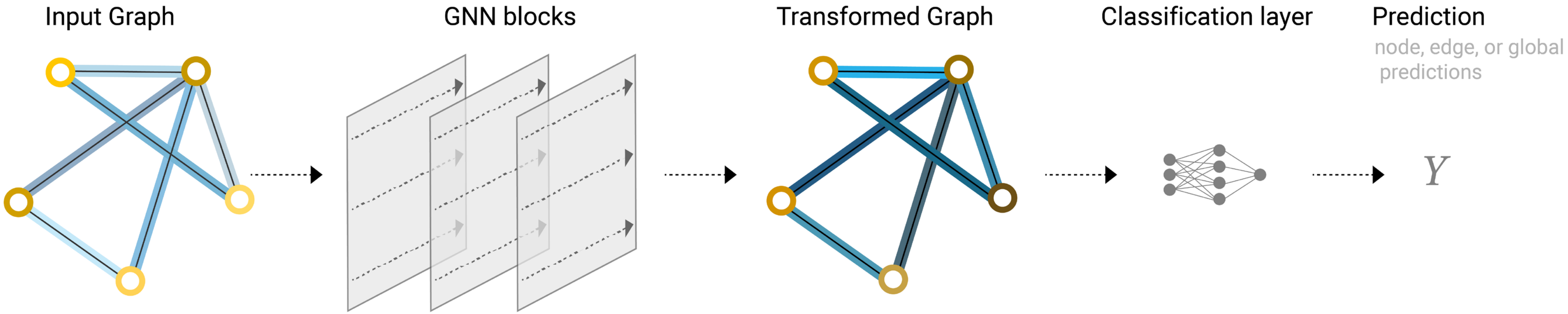
Measurements



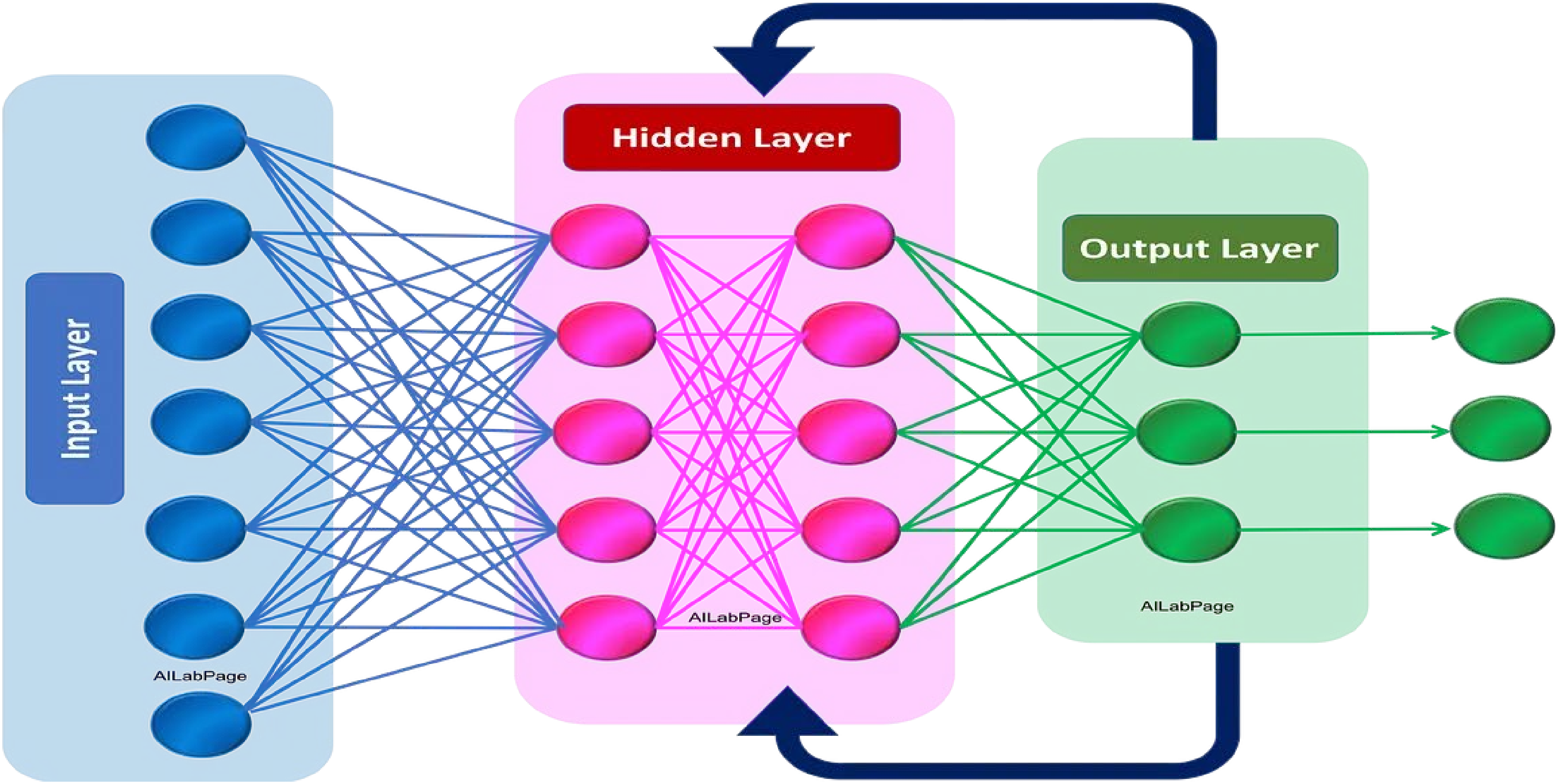
Convolutional Neural Network



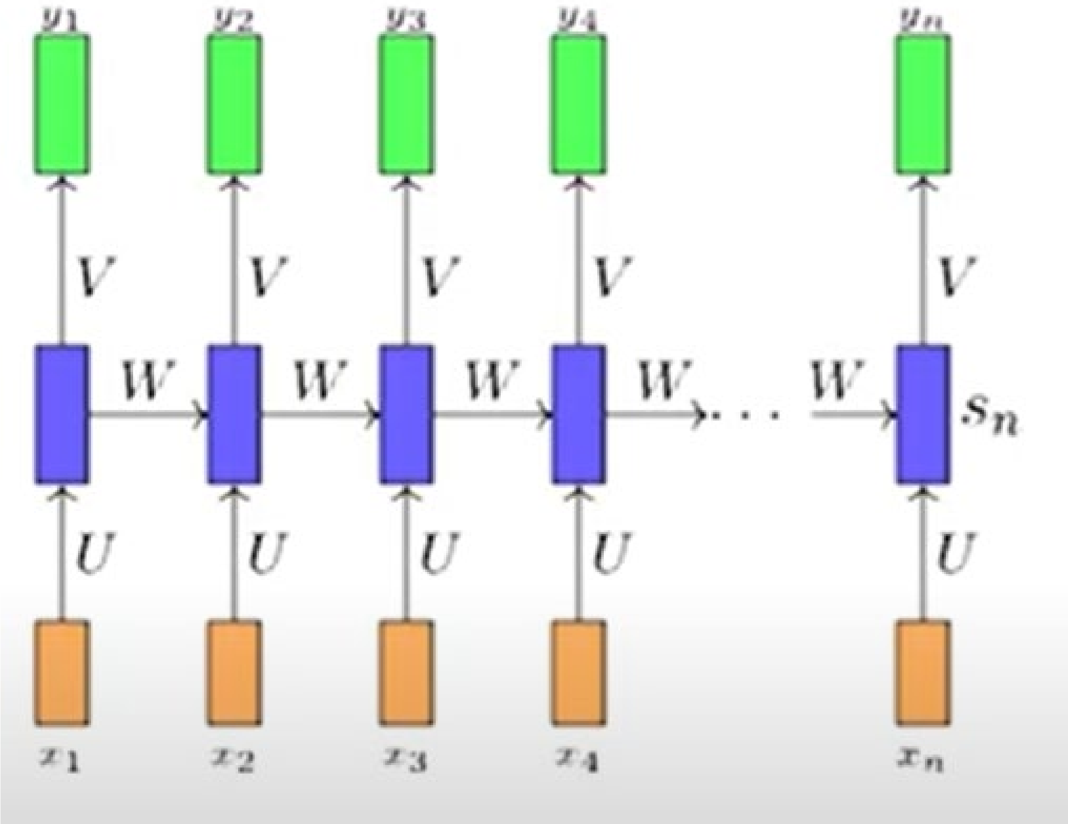
Graph Neural Network



Recurrent Neural Networks

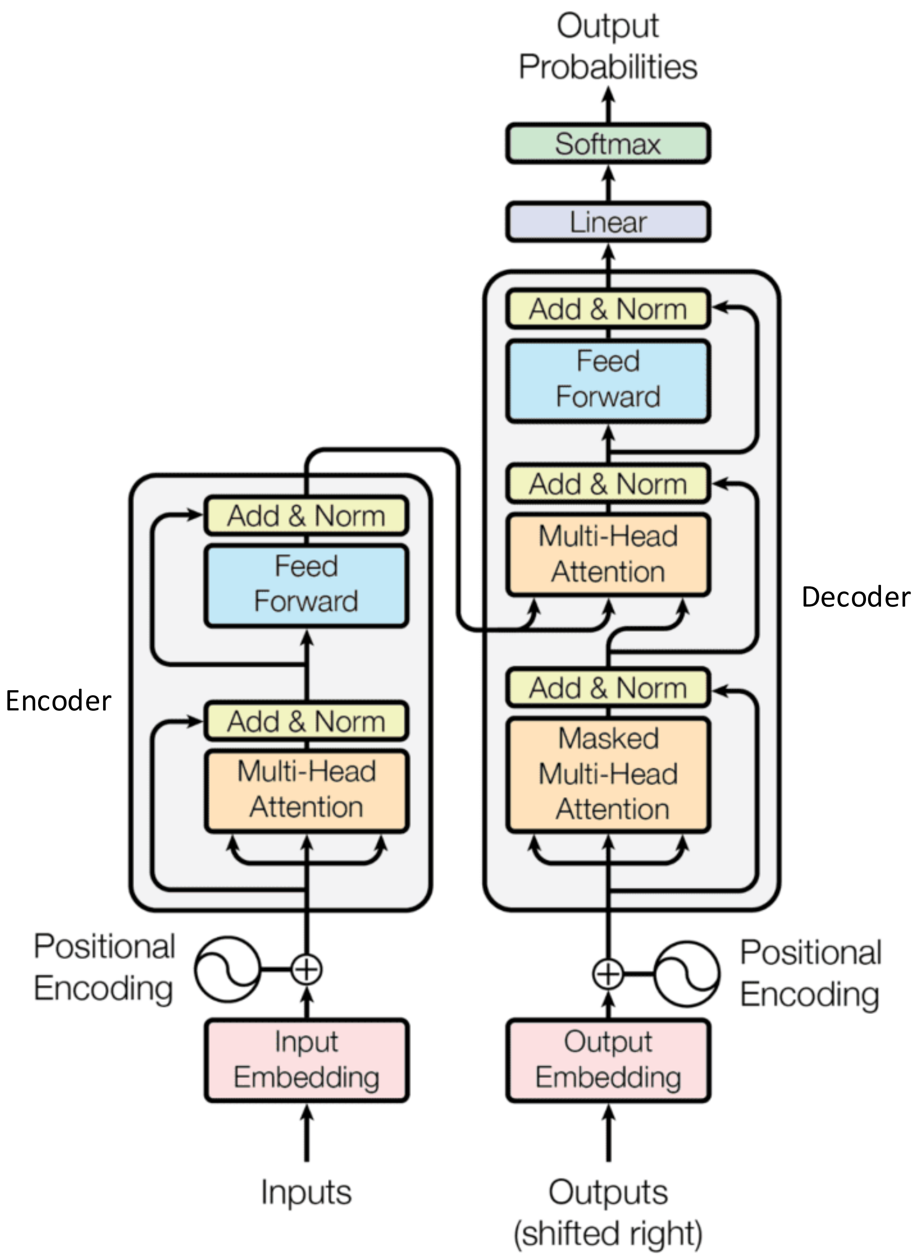


RNN



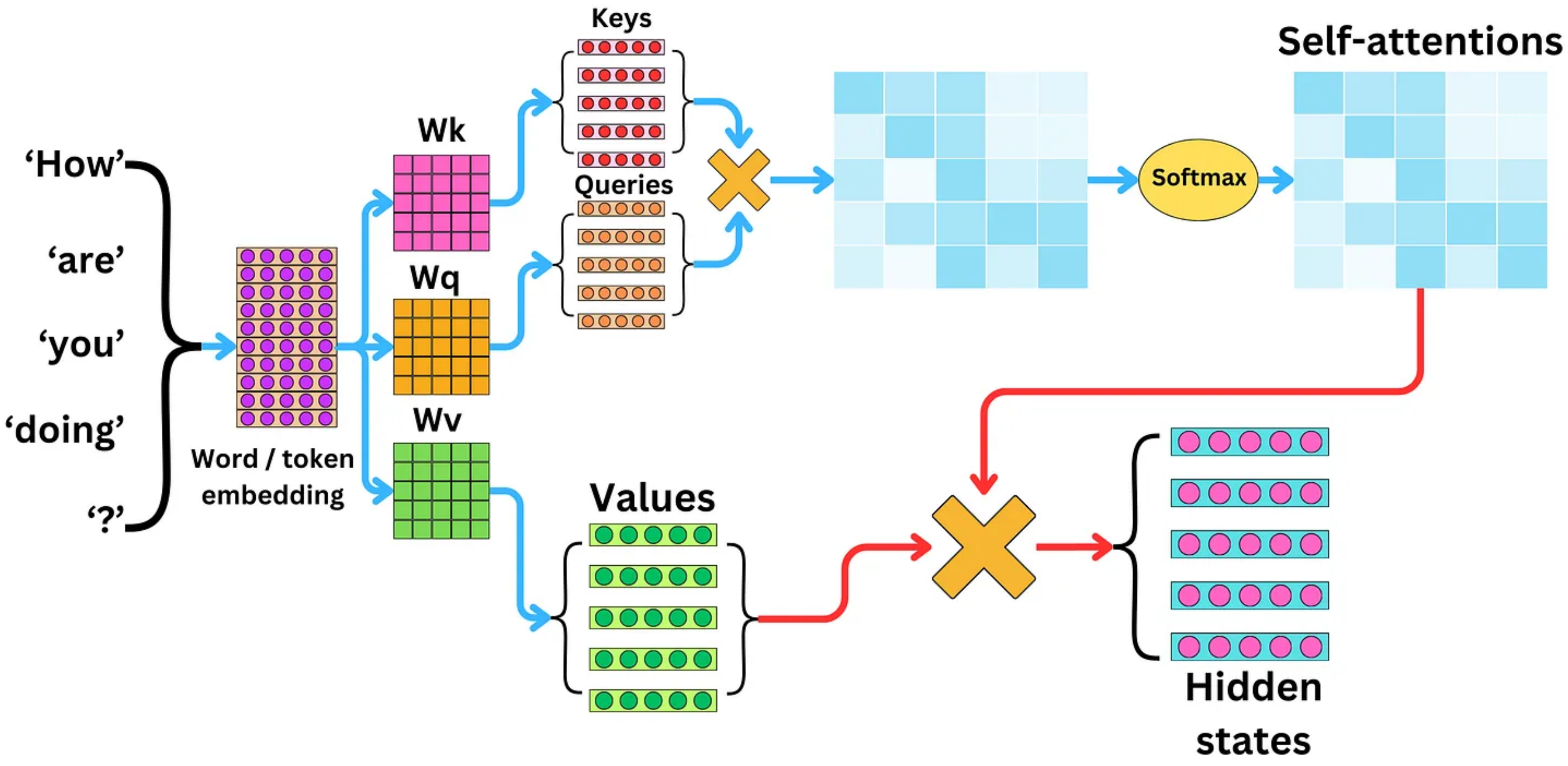
Transformer Model

- Encoder, on the left half of the Transformer architecture, is to map an input sequence to a sequence of continuous representations, which is then fed into a decoder.
- Decoder, on the right half of the architecture, receives the output of the encoder together with the decoder output at the previous time step to generate an output sequence.

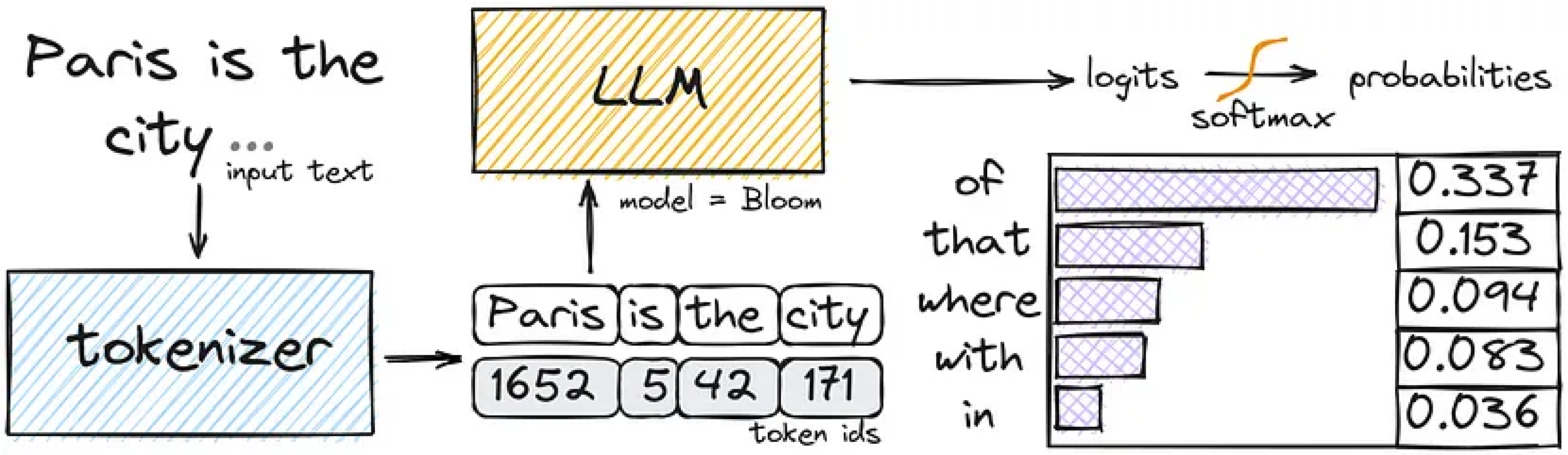


Transformer Models - Self-Attention head

- (relative importance of each component in a sequence relative to the other components in that sequence)



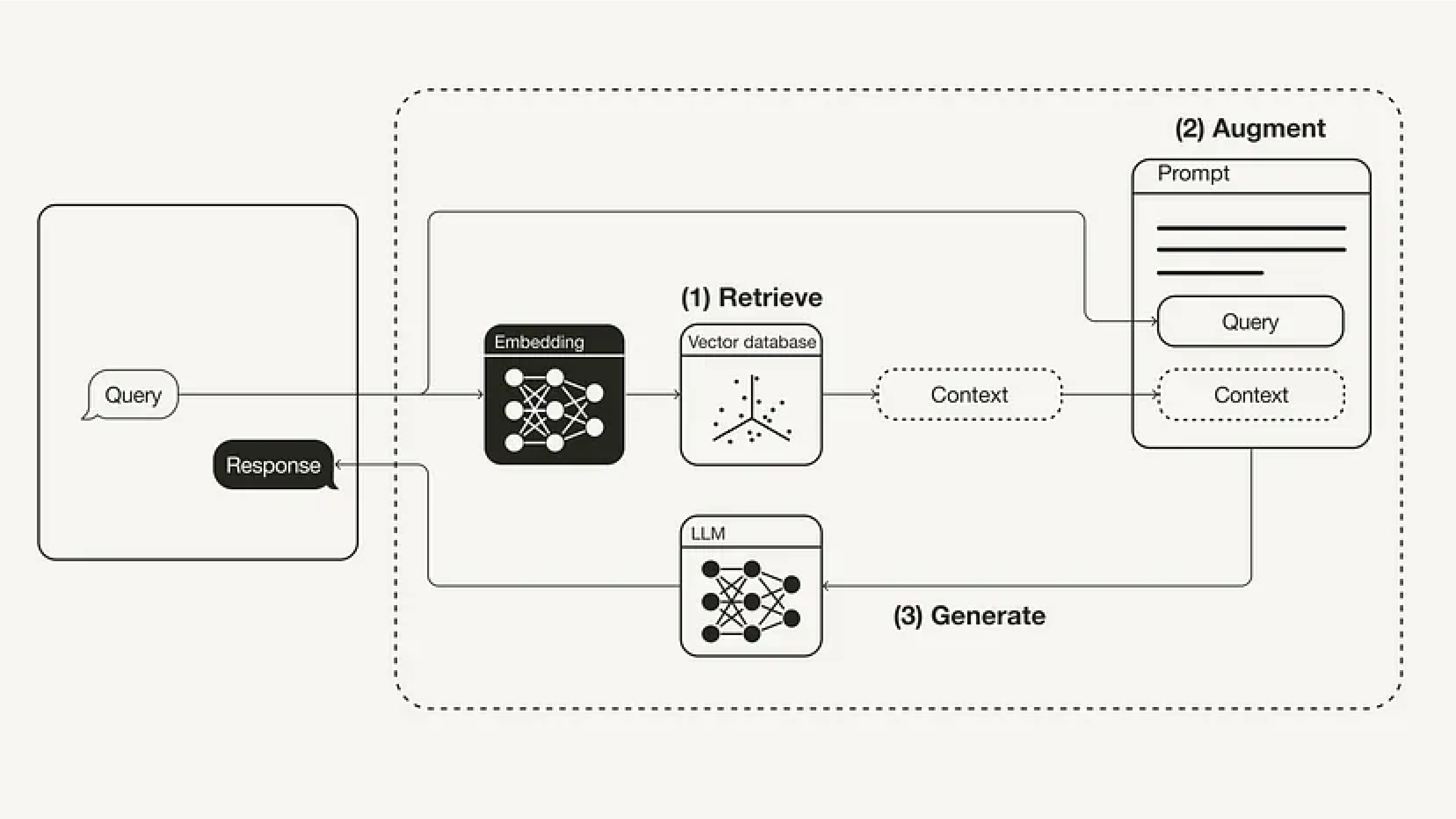
Generative AI



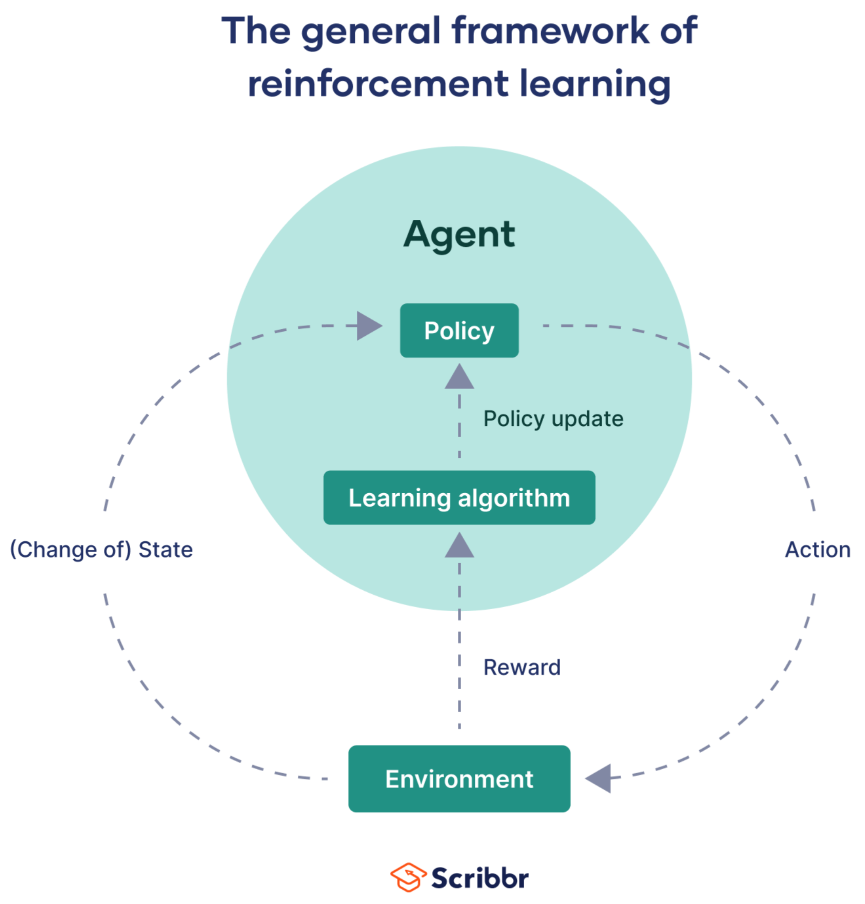
Improving Generative AI (LLMs)

- **Prompt engineering** - tailored inputs to get desired outputs from LLMs. Model's underlying structure stays same; prompts guide the model's response accurately. Example prompt engineering methods: few-shot prompting, analogical prompting, and chain of density.
- **Fine-tuning** - Fine-tuning involves adapting a pre-trained model by further training on new data to enhance its understanding and response capabilities regarding nuances not initially covered.
- **Retrieval-Augmented Generation (RAG)**
  - **Retrieve:** The user query is used to retrieve relevant context from an external knowledge source. For this, the user query is embedded with an embedding model into the same vector space as the additional context in the vector database. This allows to perform a similarity search, and the top k closest data objects from the vector database are returned.
  - **Augment:** The user query and the retrieved additional context are stuffed into a prompt template.
  - **Generate:** Finally, the retrieval-augmented prompt is fed to the LLM

Retrieval-Augmented Generation (RAG)



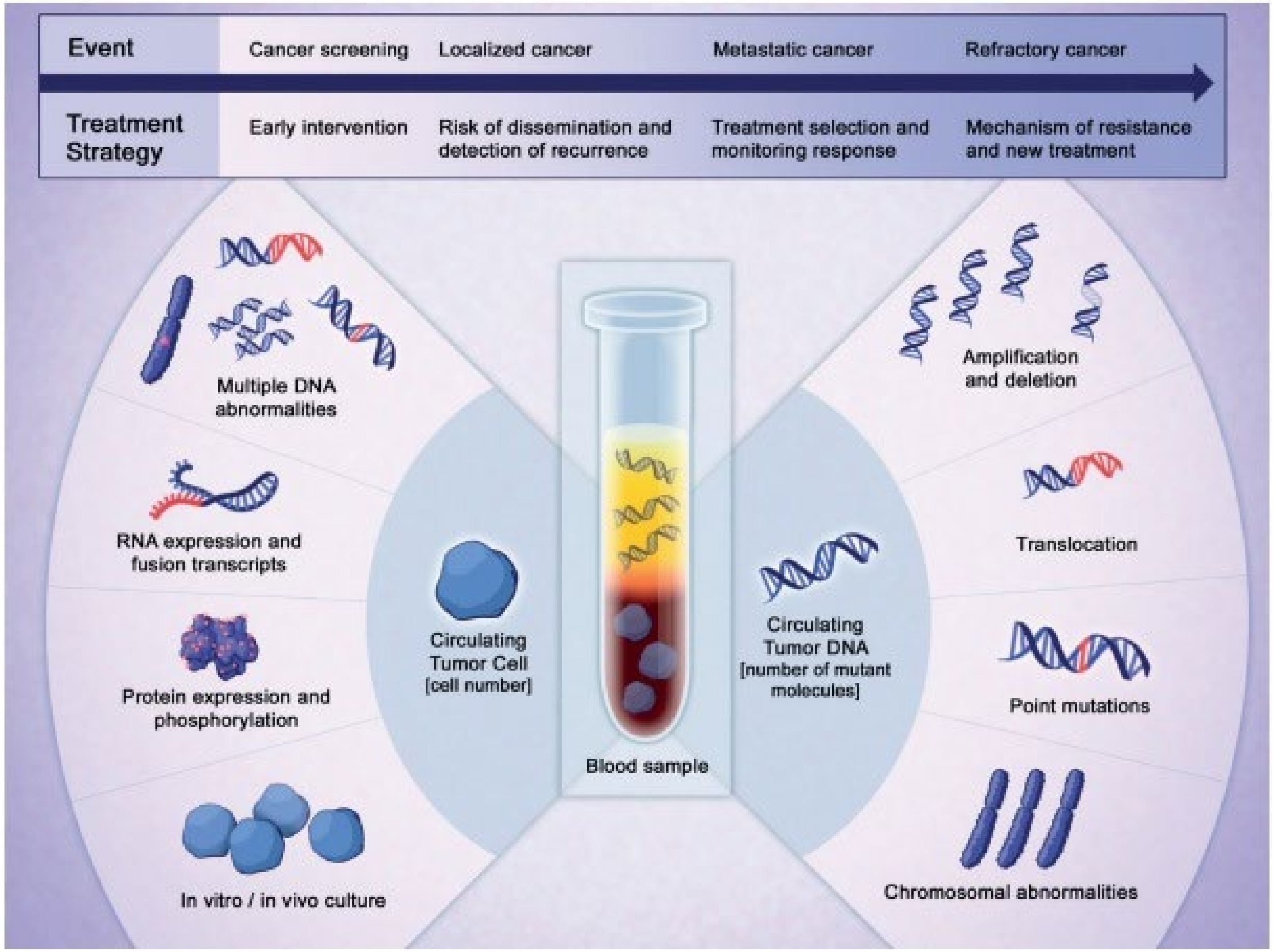
Reinforcement Learning



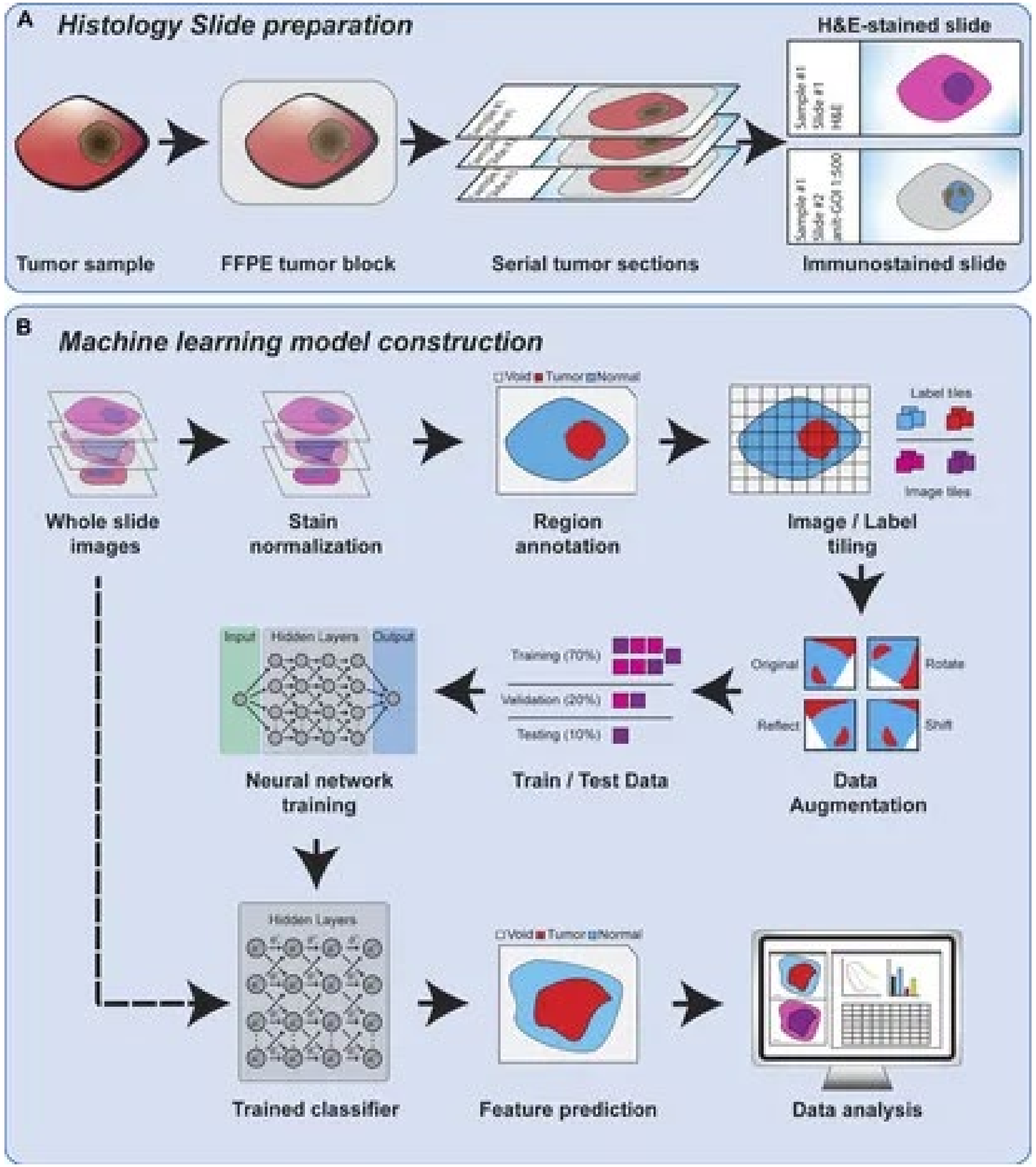
Diagnostics/Theranostics

MONTH 00, 2024

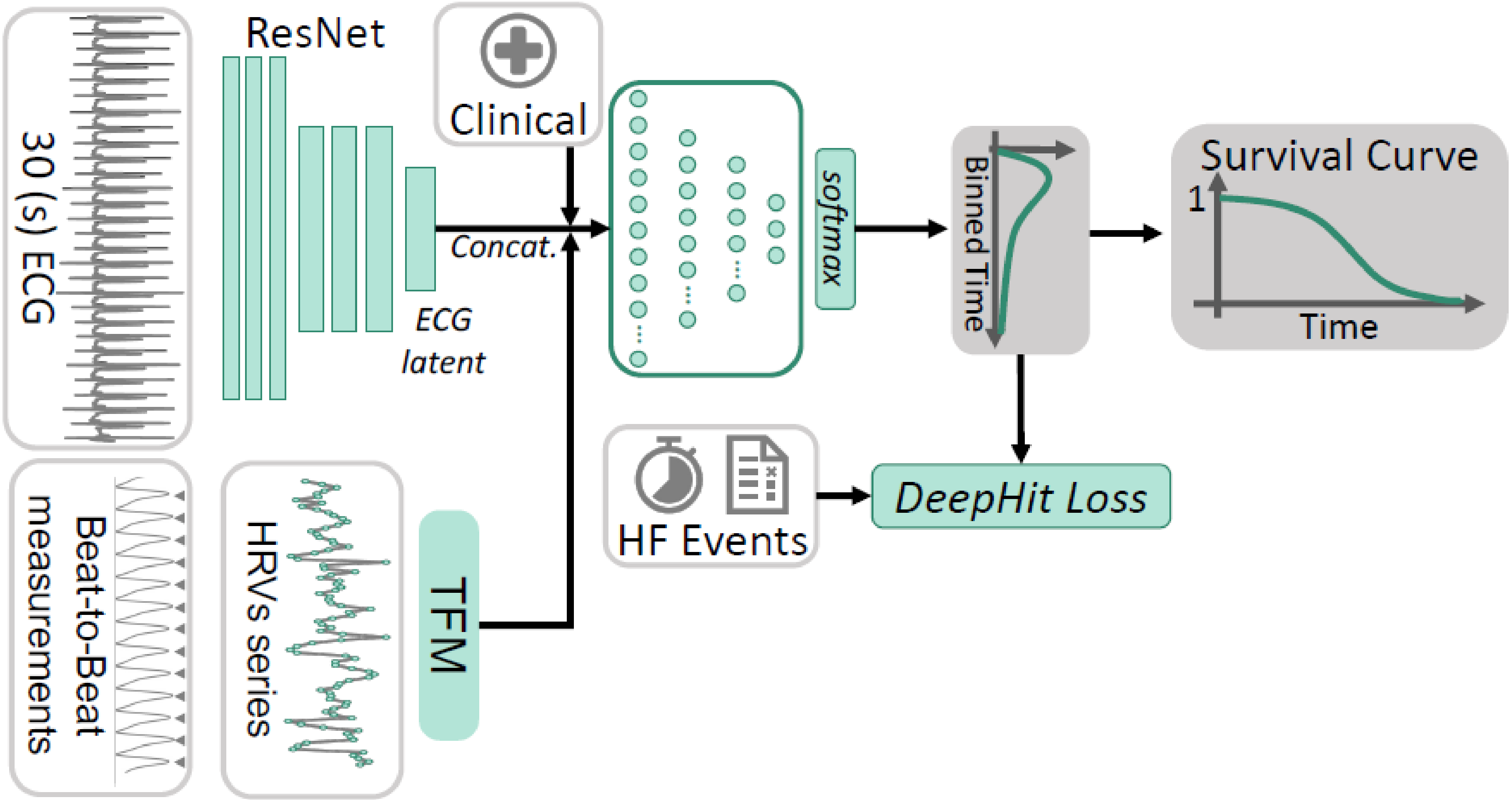
Molecular Assays: cancer, fetal, transplant



Histopathology Images



Heart measurements



Other Examples

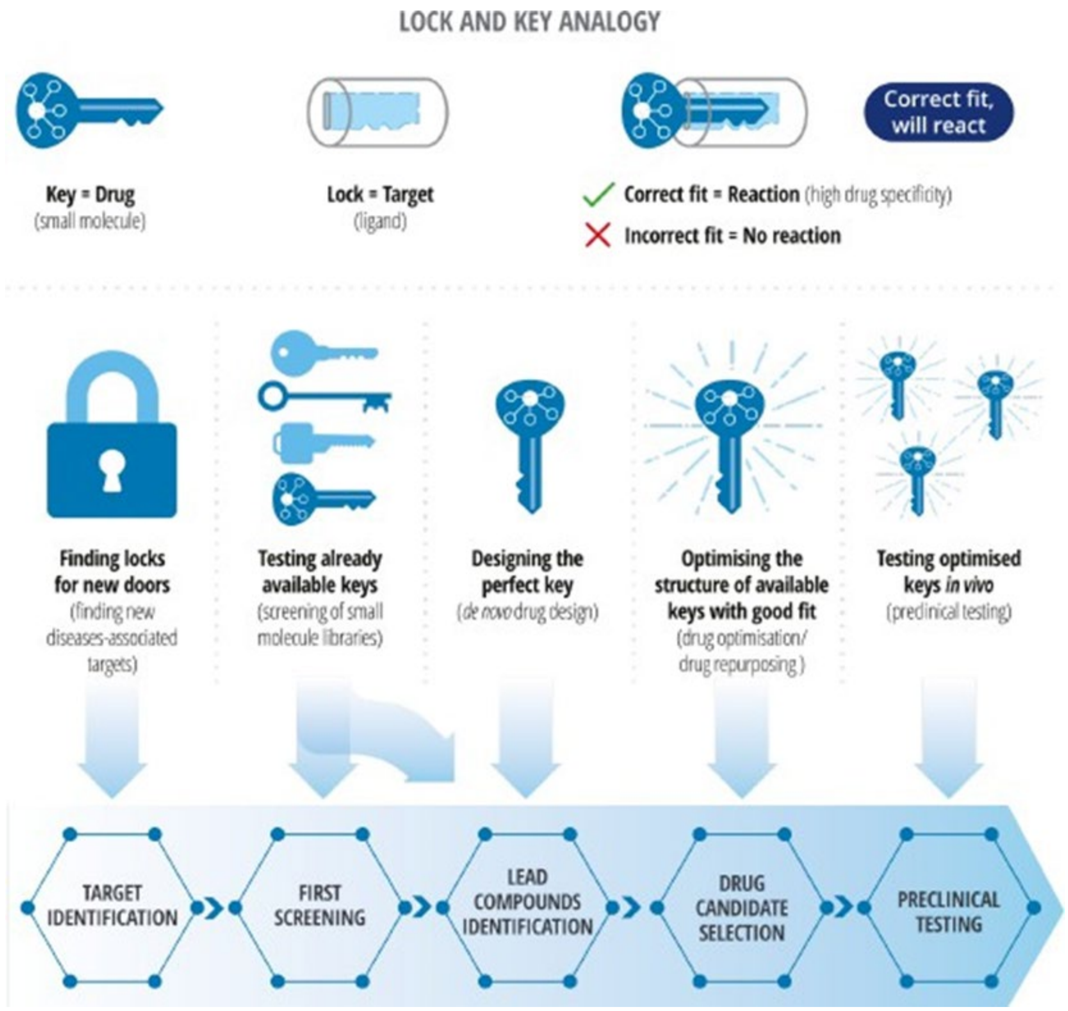
- Risk stratification for patient deterioration (e.g., death or ICU): routine vital signs, laboratory data, and patient demographics
- Plethysmogram (PLETH) and Photoplethysmogram (PPG)
  - sleep apnea, Atrial Fibrillation Prediction, Biometric identification, Blood glucose analysis, Blood pressure analysis
- Surgical Risk: identifying groups with statistical probability outcomes
- Trauma: Using CT head examinations, CNN detected intracranial hemorrhage and other acute brain findings, such as mass effect or skull fractures

Drug Discovery

MONTH 00, 2024

Kilpatrick

Lock and Key Analogy – Five Main Challenges for AI in Drug Discovery



AI Can Help Develop a Deeper Understanding of Targets and Identify New Molecules

Biology – Finding the Right Biological Targets

- Identify new genome-wide targets using AI mining of –omics data
- Map novel disease pathways, protein/drug and polypharmacological interactions based on AI analysis across networks (leveraging academic and experimental data)
- Identify new targets and leads from AI analysis of phenotypic results (e.g., images)
- Understand protein interactions, function, and drugability using protein structure prediction
- Identify novel binding sites and protein interactions through modeling of protein motion

Chemistry – Identifying and Designing Small Molecules for Preclinical Candidates

- Ability to explore a vast array of chemical space and conduct molecular dynamics to screen for small molecule hits
- AI analysis of molecule structure and experimental data to predict small-molecule structure/activity relationships
- Analysis of prospective protein target to generate lead-like small molecules
- AI mining of literature and internal data to predict optimal synthetic routes
- Preselects drug leads based on AI-predicted pharmacokinetic and pharmacodynamic properties
- Preemptively flags leads and candidates for predicted off-target effects

SOURCE: QURESHI ET AL. (2023)

AI-Based Software Tools for Drug Development Process

Reference	Description	Source code
AlphaFold2 [35]	Deep learning based model for 3D structure prediction of proteins from amino acid sequences	https://github.com/deepmind/alphafold/
DeepChem [80]	A deep learning library for drug discovery and computational chemistry	https://github.com/deepchem/deepchem
DeepBind [81]	A computational tool to analyze binding between the protein and DNA/RNA	https://github.com/MedChaabane/DeepBind-with-PyTorch
DeepBar [82]	A method for accurate and fast prediction of binding free energy	https://fastmbar.readthedocs.io/en/latest/
Deep-Screening [83]	Web-server based in deep learning for virtual screening of compounds	http://deepscreening.xielab.net/
DeepScreen [84]	High performance drug target interaction	https://github.com/cansyl/DEEPScreen
DeepConv-DTI [45]	A convolutional neural network based model for predicting drug-target interactions	https://github.com/GIST-CSBL/DeepConv-DTI
DeepPurpose [24]	A Deep learning library for drug-target interaction, drug-drug interaction, protein-protein interaction and protein function prediction	https://github.com/kexinhuang12345/DeepPurpose
DeepTox [85]	A deep learning model for toxicity prediction of chemical compounds	http://www.bioinf.jku.at/research/DeepTox/
AtomNet [86]	A deep convolutional neural network for bioactivity prediction	github
PathDSP [87]	A deep learning method for predicting drug sensitivity using cancer cell lines	https://github.com/TangYiChing/PathDSP
Graph level representation [88]	Learning graph representation for drug discovery	https://github.com/ZJULearning/graph_level_drug_discovery
Chemical VAE [89]	An auto-encoder based framework to generate new molecules	https://github.com/aspuru-guzik-group/chemical_vae/
DeepGraphMol [87]	A computational method for molecule generation with desired properties using graph neural networks and reinforcement learning	https://github.com/dbkgroup/prop_gen
TorchDrug [26]	A pytorch based flexible framework for drug discovery models	https://torchdrug.ai/

SOURCE: QURESHI ET AL. (2023)

Software for MD Simulation, Modeling, Docking, Visualization, and Analysis of Molecules

Reference	Description	Pros	Cons	Source code
AMBER [144]	A package for MD simulation	High Performance MD, Comprehensive trajectory analysis tools	License required for parallel CPU or GPU computation	https://ambermd.org/
ACEMD [145]	An accelerated platform for faster and longer biomolecular simulations	Super computer level performance	License required for ful functionality	https://www.acellera.com/
AutoDock Vina [146]	A program for molecular docking and screening	Receptor flexibility, blind docking	Difficult to dock small peptides	https://vina.scripps.edu/
DeePMD [147]	A deep learning package for MD simulation and energy representation	Optimized code, interfaced with Tensorflow	Model compression issues	https://github.com/deepmodeling/deepmd-kit/
RBio3D [148]	R package for the analysis of MD trajectories	Tools for protein-networks, conformations	-	http://thegrantlab.org/bio3d/
Pymol [149]	An interactive platform for visualization of molecules	Homology Modeling, Docking, Virtual Screening	License required for full features	https://pymol.org/2/
Rosetta Commons [124]	A tool for predicting the mutant structure	Protein modeling and folding	Preference for aromatics, Preference for hydrogen bonding	https://www.rosettacommons.org/

Manufacturing

MONTH 00, 2024

Kilpatrick

AI in manufacturing

- **Process Design and Scale-up:**
  - AI models such as machine learning—generated using process development data—could be leveraged to more quickly identify optimal processing parameters or scale-up processes,
- **Advanced Process Control (APC):**
  - AI methods can also be used to develop process controls that can predict the progression of a process by using AI in combination with real-time sensor data.
- **Process Monitoring and Fault Detection:**
  - AI methods can be used to monitor equipment and detect changes from normal performance that trigger maintenance activities, reducing process downtime
- **Trend Monitoring**
  - AI can be used to examine consumer complaints and deviation reports containing large volumes of text to identify cluster problem areas and prioritize areas for continual improvement.

Kilpatrick Confidential

Clinical Trials

MONTH 00, 2024

Kilpatrick

AI in Clinical Trials

Advanced data analytics and AI automation

TRIAL DESIGN

- Assess feasibility of protocol design for patient recruitment using RWD.
- Assess site performance (e.g. enrollment and dropout rates) with real-time monitoring.
- Analyse and interpret unstructured and structured data from previous trials and scientific literature.

TRIAL STARTUP

- Mine EHRs and publicly available content, including trial databases and social media, to help match patients with trials, by using NLP and ML.
- Create drafts of investigator and site contracts and confidentiality agreements by smart automation.

TRIAL CONDUCT

- Assess site performance (e.g. enrollment and dropout rates) with real-time monitoring.
- Analyse digital biomarkers on disease progression, and other quality-of-life indicators.
- Automate sharing of data across multiple systems.

STUDY CLOSEOUT

- Complete sections of the final clinical trial report for submission by using NLP
- Data cleaning by ML methods.

AI in Clinical Trials Cont…

AI-enhanced mobile applications, wearables, biosensors and connected devices

TRIAL STARTUP

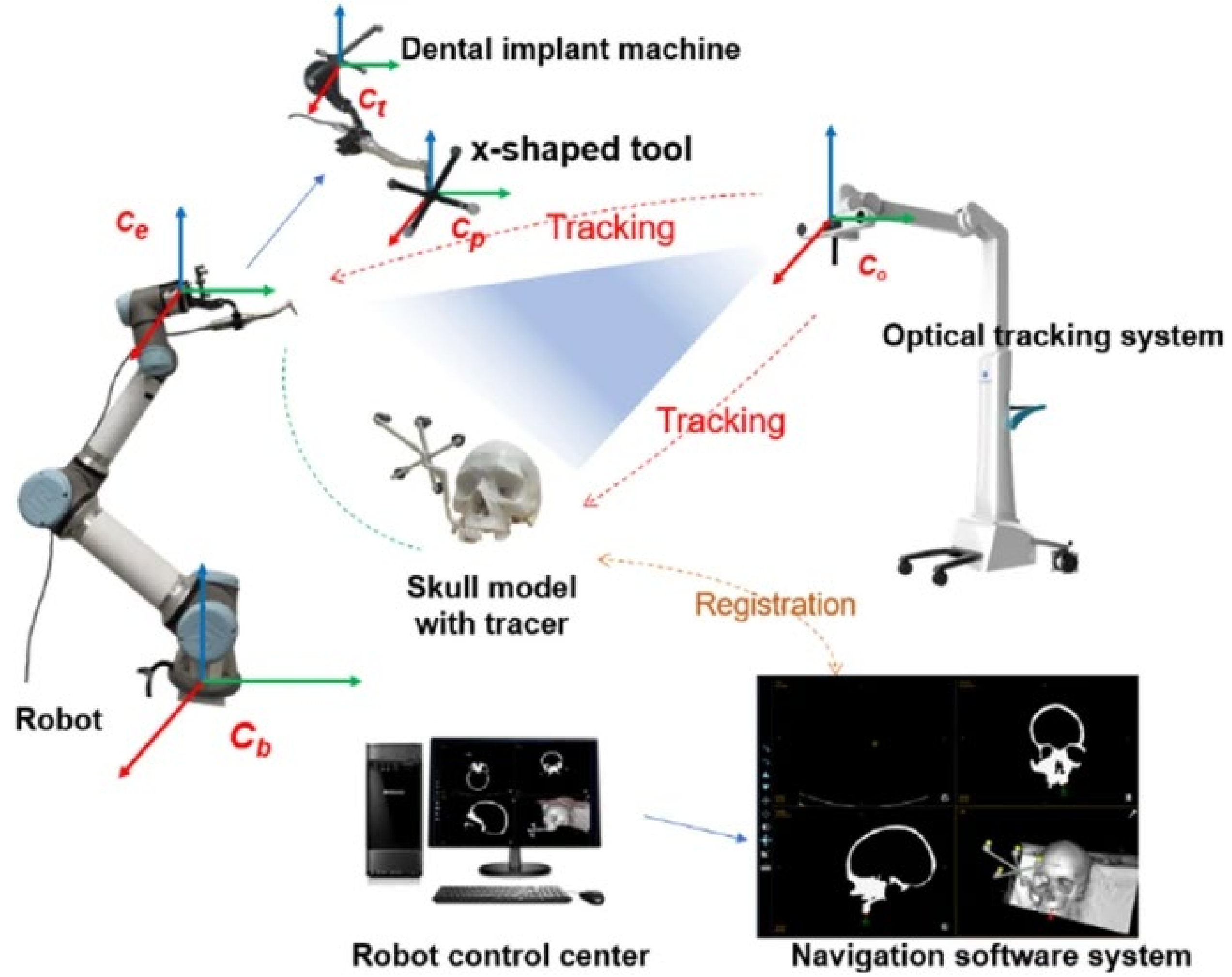
- Expedite recruitment and create a more representative study cohort through cloud-based applications.
- Simplify and accelerate the informed consent process using eConsent.

TRIAL CONDUCT

- Enhance adherence through smartphone alerts and reminders
- eTracking of medication using smart pillboxes, and tools for visual confirmation of treatment compliance.
- eTracking of missed clinic visits, and trigger non-adherence alerts.

Guided Surgery

Image-Guided Surgery



Registration for imaging

- Evaluating the risk of post-surgical complications and selecting individuals for operative procedures (Prognosis for outcomes for different treatment)
- Segmentation can be used to identify mass (tumor, to remove vascular calcifications)
- Placing an overlay on the surgeon's video screen during an operation to suggest where it is safer or less safe to operate
- Allow the gripper to react to the deformation of the gauze and progress of the cutting trajectory with a translation unit vector along an allowable set of directions. (B. Thananjeyan, et al. *2017 IEEE International Conference on Robotics and Automation (ICRA)*, Singapore, 2017, pp. 2371-2378,
- Anticipate next 15 to 30 seconds of operation and provide additional oversight during surgery
- Outcome of supervised autonomous procedures is superior to surgery performed by expert surgeons and RAS techniques in ex vivo porcine tissues and in living pigs (Shademan A, e al.,. Sci Transl Med. 2016 May 4;8(337):337ra64.

Speaker Contact Information

David Raczkowski, Ph.D.
Kilpatrick Townsend & Stockton LLP
Partner | San Francisco, CA
415 273-7555
draczkowski@ktslaw.com

Kilpatrick

Anchorage | Atlanta | Augusta | Beijing | Birmingham | Charlotte | Chicago | Dallas | Denver | Houston
Los Angeles | New York | Phoenix | Raleigh | San Diego | San Francisco | Seattle | Shanghai
Silicon Valley | Stockholm Tokyo | Walnut Creek | Washington D.C. | Winston-Salem