Priority challenges in the application of AI to ecology management

Jeff Boisvert
Civil Engineering, U of Alberta
jbb@ualberta.ca

Mike Flannigan, Thompson Rivers University,
University of Alberta
mflannigan@tru.ca

Randy Goebel, Computing Science, U of Alberta
rgoebel@ualberta.ca
Piyush Jain, Natural Resources Canada
jain@ualberta.ca

Nature, 16 August 2022

EDITORIAL | 16 August 2022

We must get a grip on forest science – before it's too late

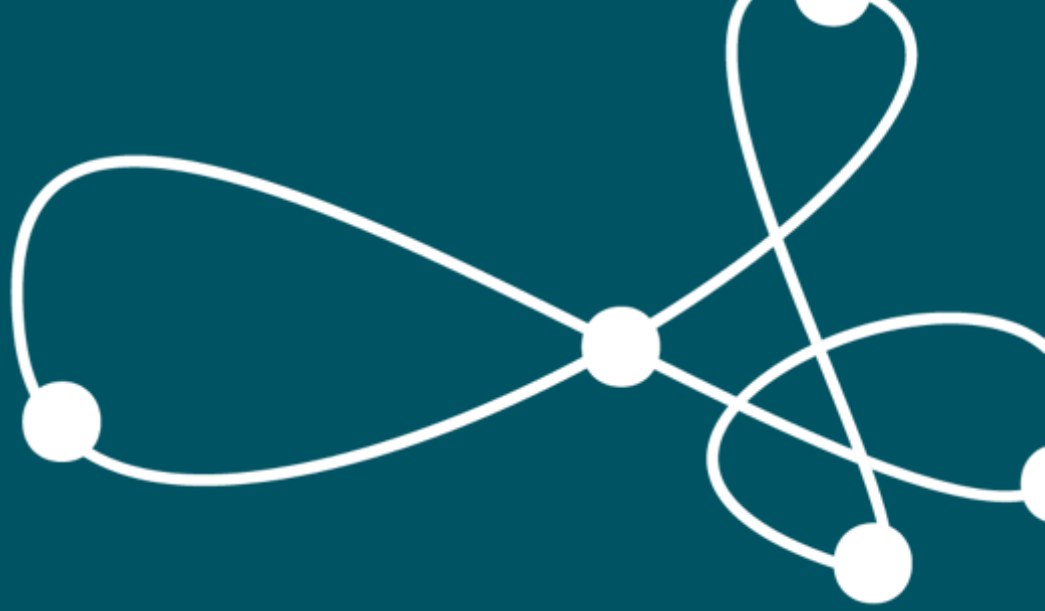
Trees are one of our biggest carbon hopes. Supporting the scientists studying them should be a much higher priority.

- Central is the development of dynamic global vegetation models (DGVMs).

https://www.nature.com/articles/d41586-022-02182-0?WT.ec_id=NATURE-20220818&utm_source=nature_etoc&utm_medium=email&utm_campaign=20220818&sap-outbound-id=ED733BD4EDEE67C46E39D80E82206E65F6358D08

Outline

- Jasinski's Rule: 3 components necessary for successful application of AI
- Data-driven management of natural resources
  - Infrastructure
  - biology
- Examples from the University of Alberta
  - predictive models for forest fire
  - biology of mountain pine beetle
  - environmental properties of soil
- Canadian Federal initiatives
  - Natural Resources Canada
  - Forest Products Innovation
  - Agricultural Climate Solutions
- Summary

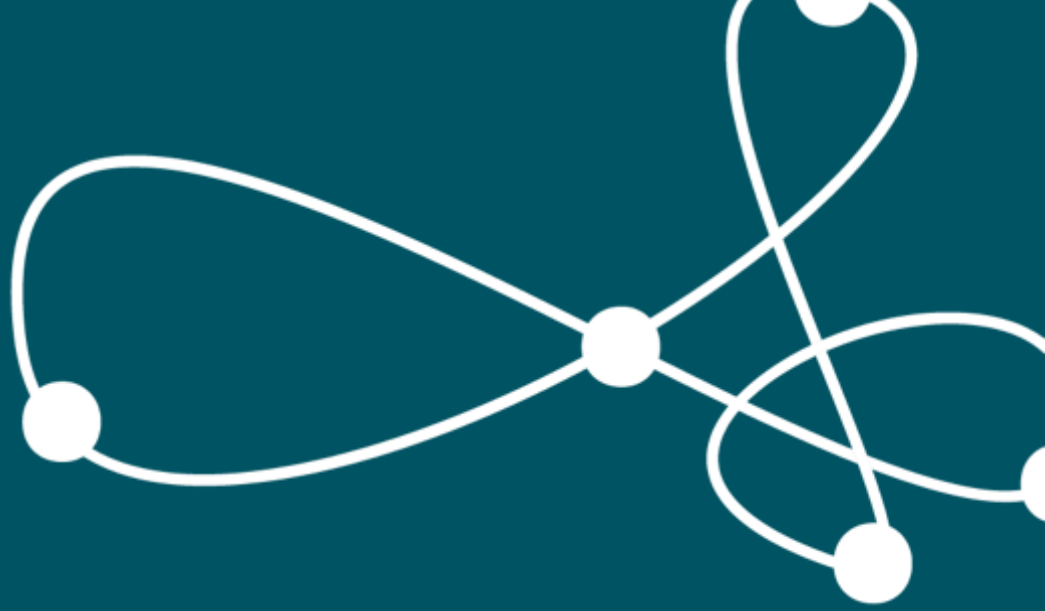


Jasinski’s Rule

Three essential components of a successful AI application

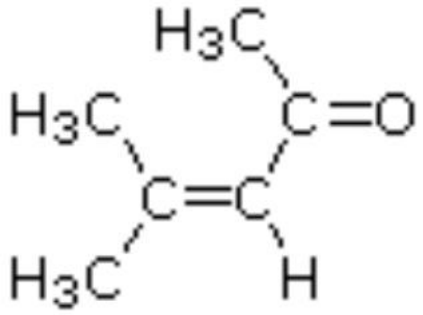
- Access to relevant data
- Subject matter experts to define key performance measures (KPIs)
- Data scientists committed to learning subject matter vocabulary

Joe Jasinski, retired Director of IBM TJ Watson Life Sciences



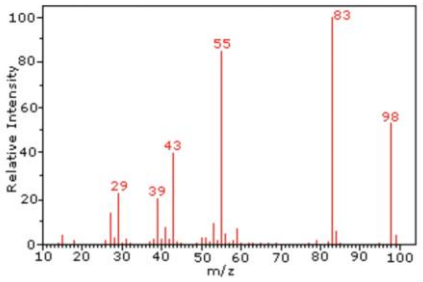
Evolution of knowledge acquisition technologies

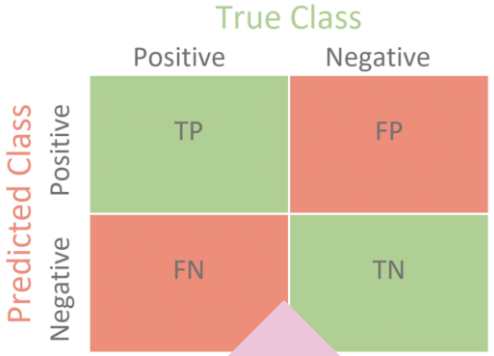
Evolution of knowledge acquisition



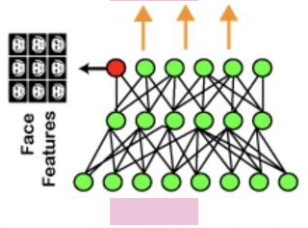
Subject matter expert

"The compound water (H2O), has a molecular weight of 18 since hydrogen has a mass of 1.01 and oxygen 16.00, and its mass spectrum has a peak at 18 units."



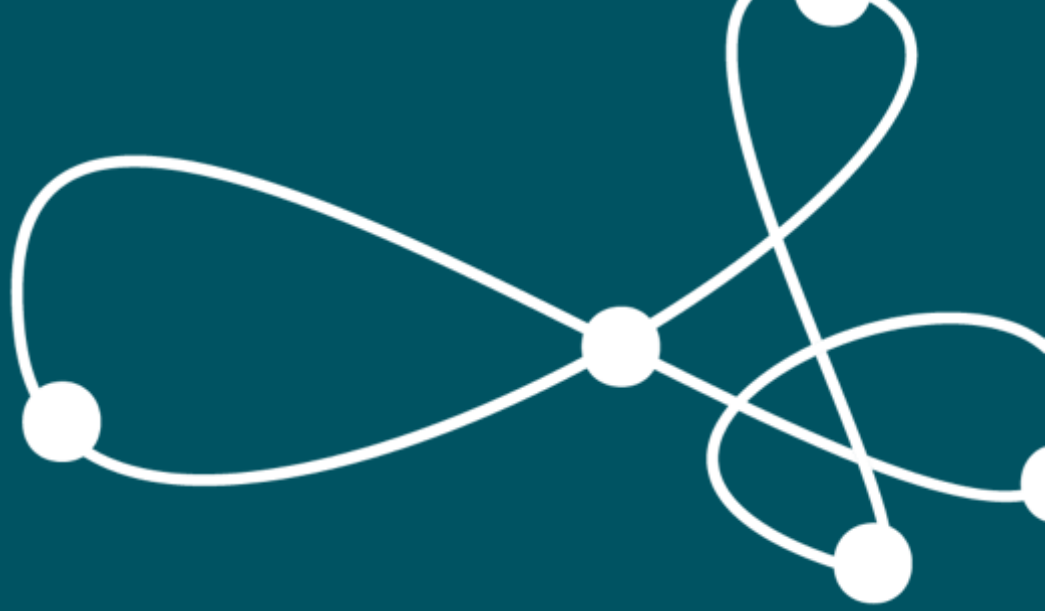


Subject matter expert



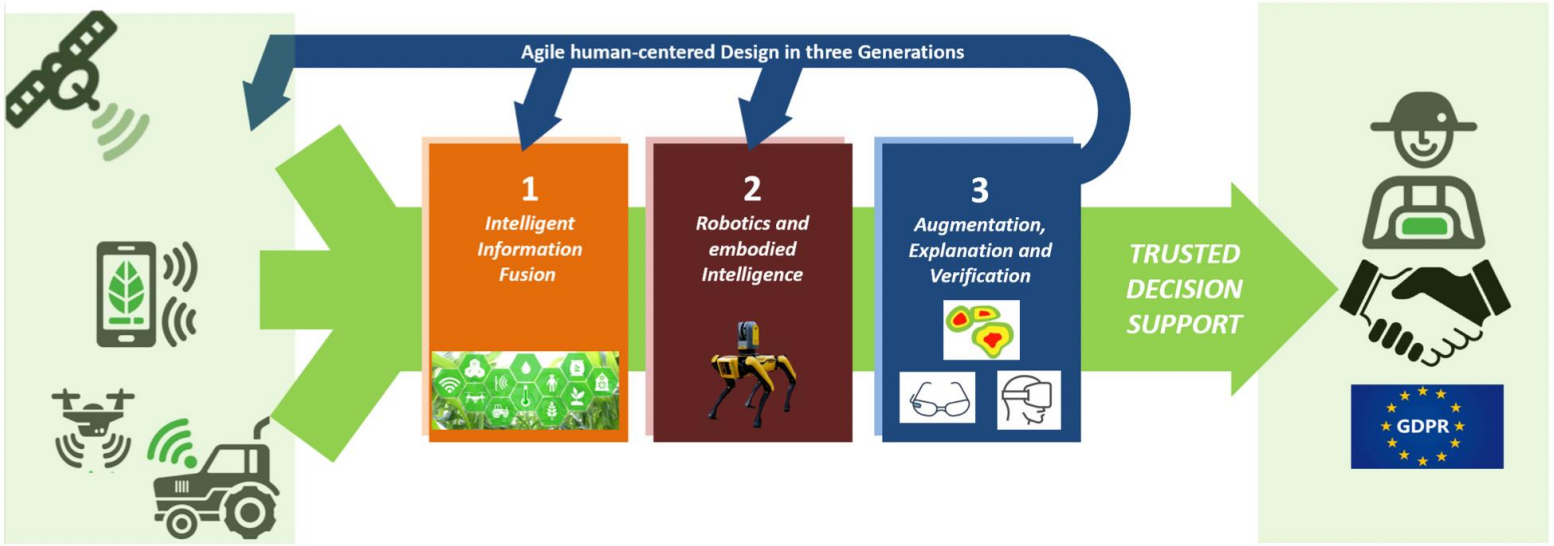
"black box classifier"

Subject matter expert



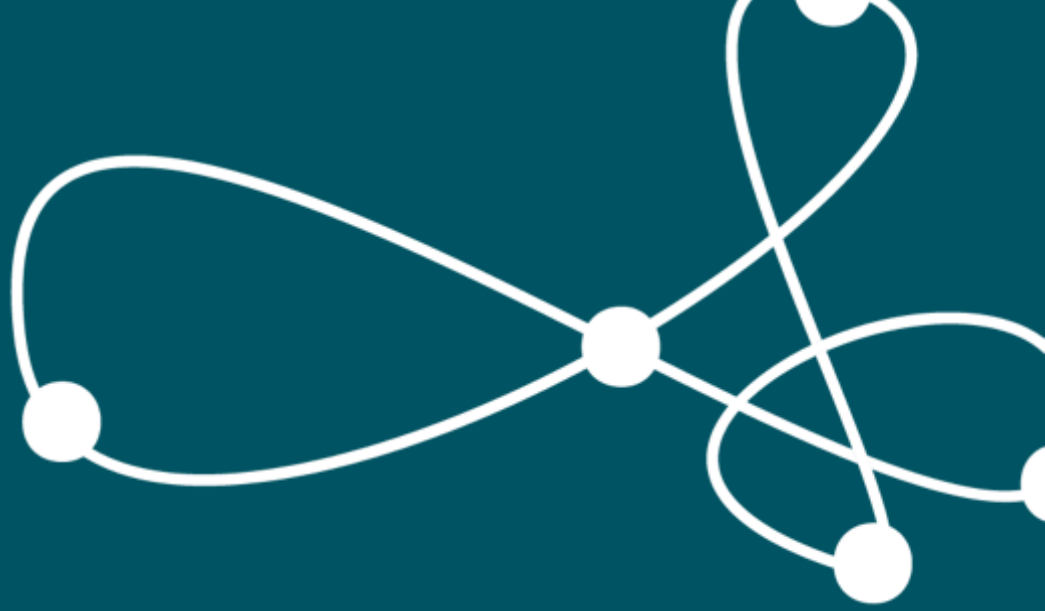
Holzinger Vision

Digital Transformation ... Needs Human-Centered AI.



Andreas Holzinger

- "Digital transformation in smart farm and forest operations needs human-centred AI" (https://pubmed.ncbi.nlm.nih.gov/35459028/)
- Integration of AI methods needs to have a human-centric structure, *by design.*



Applications of AI/Data Science

University of Alberta initiatives

- University of Alberta is ranked 4-5 in Canada, and about 92 in the world (see https://www.ualberta.ca/folio/2022/08/u-of-a-has-best-ever-showing-in-prominent-ranking-of-worlds-top-institutions.html)
- About 40,000 students (10,000 graduate students)
- A full scope university with significant impact in Science, Medicine and Agriculture; College of Natural & Applied Sciences includes the Faculties of Science, Engineering, and Agriculture and Life Sciences.
- Note: Canada is 9,984,670 $km^{2,}$ Alberta is 661,848 km^2, Austria is 83,879 km^2 (density 4.2/ km^2, 6.6/km^2, 107.6/ $km^{2)}$
- A sample of projects related to
  - Environmental infrastructure
  - Biological infrastructure

Mike Flannigan and team … instructions to a data scientist

- Three ingredients in the wildfire recipe - Vegetation (fuel), ignition agent (humans and lightning) and conducive fire weather (typically hot, dry windy)
- Room for improving fuel information (spatial, temporal (updating - succession and changes such harvesting and disturbances), typing - characterization, phenology and moisture content (dead and live fuels using remotely sensed data))
- Forecasting extreme fire weather events as extreme conditions drive the fire world - 3% of the fires in Canada are over 200 ha but they are responsible for 97% of the area burned. Western USA 1% of the fires responsible for 99% of the area burned.
- Fire occurrence prediction - human-caused and lightning-caused; 95%+ by humans.

Mike Flannigan and team … instructions to a data scientist

- Context is that many situations require a decision in 5 minutes or less and there is vast amounts of data/indicators for the decision maker. Automated ML decision support system would be objective and be part of a machine-human decision support system.
- Globally about 420 million ha burn every year – an area the size of India. Most fires are started by humans 95+ %
- Canada 10 year annual average of 2.6 million ha (half the size of nova scotia but could covert to % of Austria). About 6, 000 fires every year – 50% lightning-caused and 50% human-caused.
- Spend billions of dollars every year on direct fire management costs globally. Canada is about 1 billion a year

Using AI models to predict wildfires

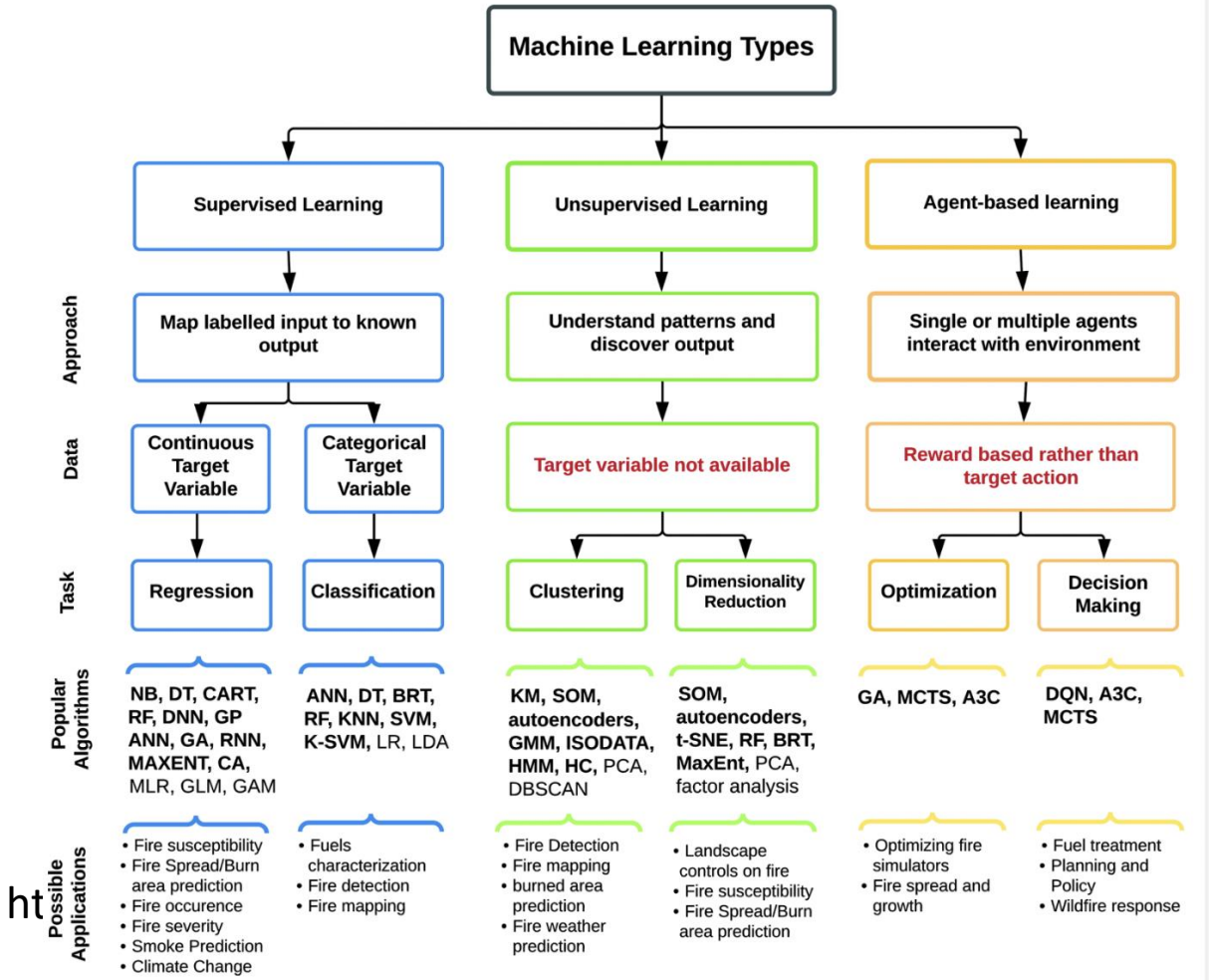
By using artificial intelligence, researchers hope to offer predictions about where wildfires could start based on weather conditions. (Darryl Dyck/Canadian Press)

Mike Flannigan

- Meteorologists can reasonably forecast hot weather and lightning storms, but exactly where lightning will strike — and whether it will spark a wildfire — is nearly impossible to predict.
- That's why researchers are turning to machine learning in a bid to get ahead of catastrophic wildfires that they say are becoming more common and more severe.
- "These machine learning approaches are even better than we are at seeing patterns that we might not see because they have much more processing power than the human brain," said Mike Flannigan, a professor of wildland fires at the University of Alberta.

https://www.cbc.ca/radio/checkup/researchers-are-using-artificial-intelligence-to-help-predict-the-next-wildfire-1.5150103

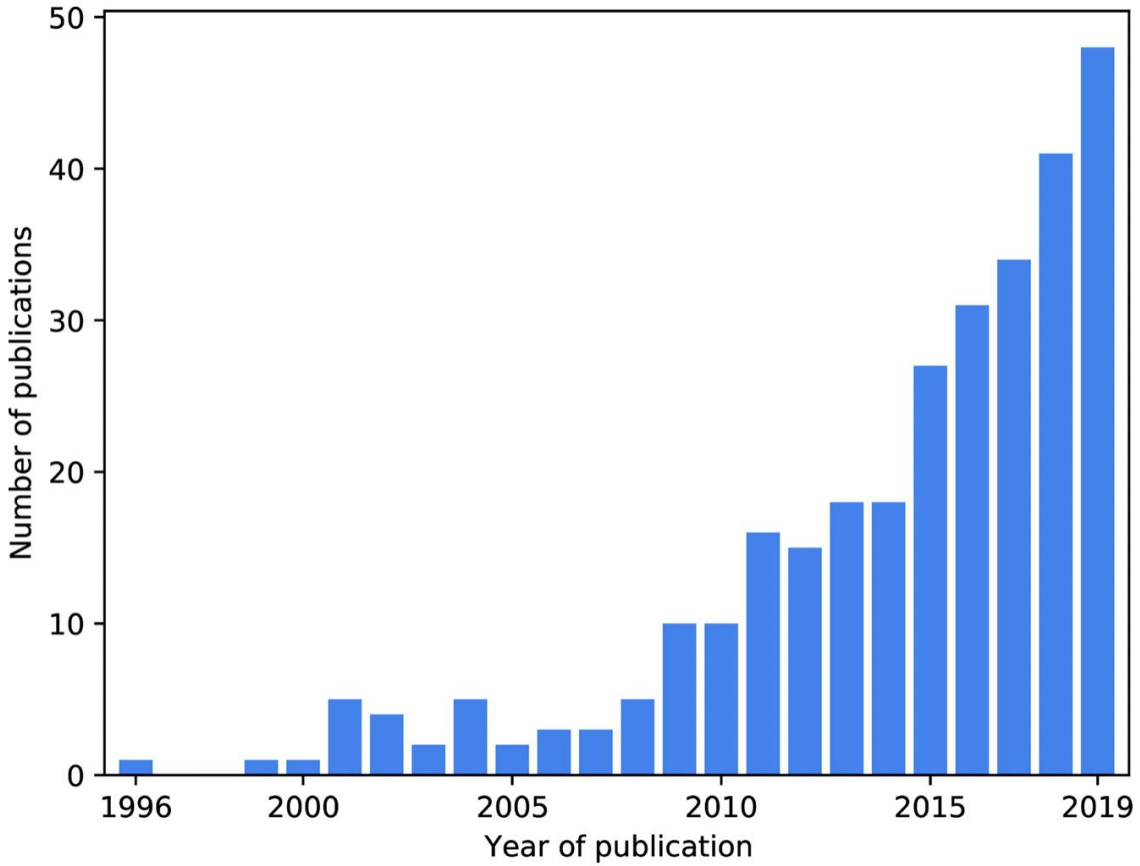
Using AI models to predict wildfires

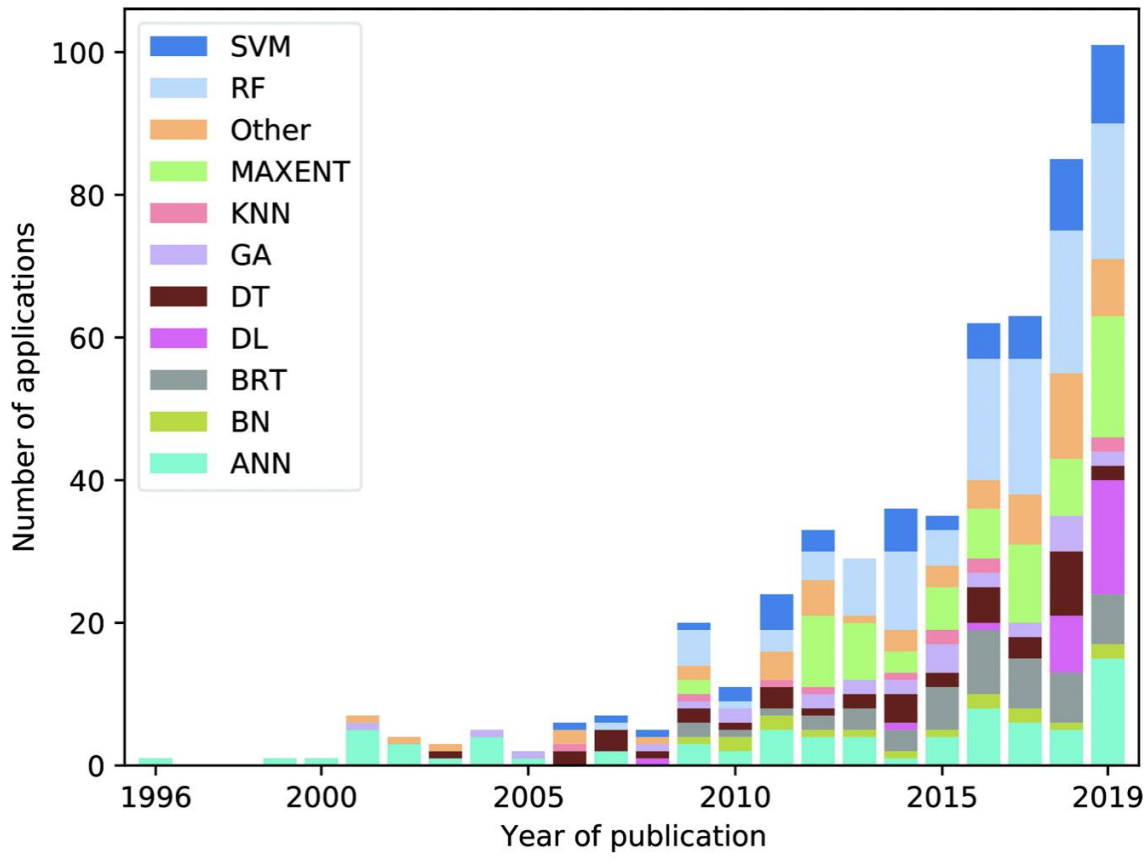


Piyush Jain et al.

- Highly detailed review of the scope and application of machine learning methods.
- **"A review of machine learning applications in wildfire science and management"** *Environmental Reviews* **(2020)** https://cdnsciencepub.com/doi/10.1139/er-2020-0019

Using AI models to predict wildfires

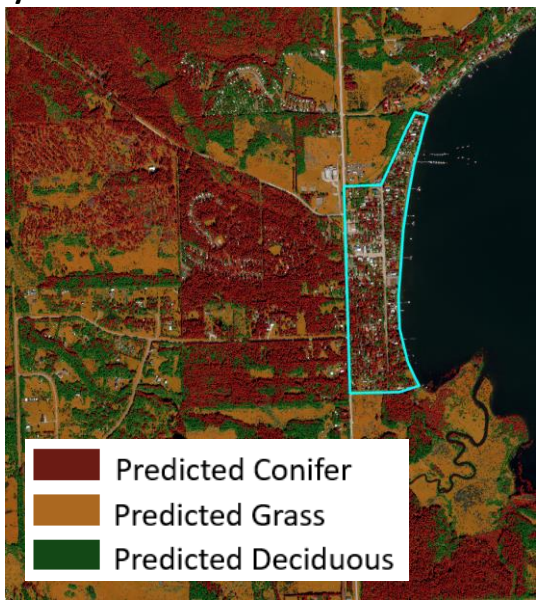




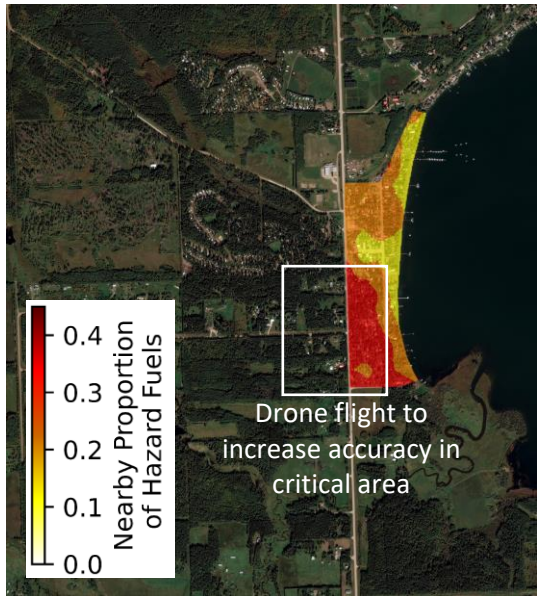
https://cdnsciencepub.com/doi/10.1139/er-2020-0019

Using AI models to manage fire resources

Community highlighted in cyan. Predictions from a CNN



Proportion of nearby hazard fuels ≈ hazard level

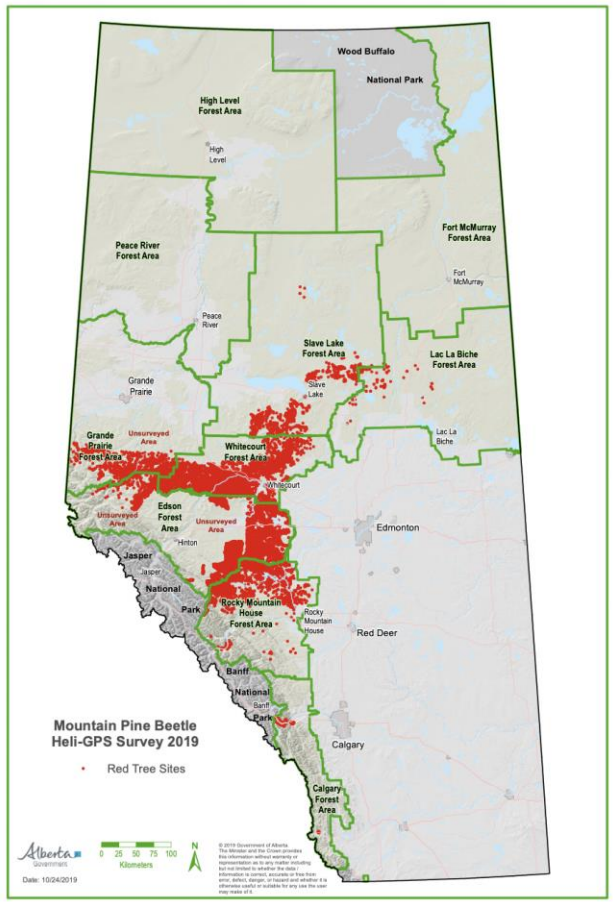


https://sites.ualberta.ca/~jbb/index.html

Jeff Boisvert

- AI/ML for fuel type and attribute modeling to feed fire prediction systems and planning. In this example, conifers are the fuel of interest (deciduous trees are not usually hazardous).
- **Classify fuels** (tree species, grass, surface fuels) and calculate fuel attributes (vertical and horizontal distribution of fuels, fuel load, base of crown height, crown bulk density, etc.)
- Use in fire prediction (slide 21)
- **Integration into workflows**: industry users want to know enough to use the system appropriately. To what degree should the results be trusted and used in decisions?
- **Decisions**: wildland fire hazard/risk assessment, exposure assessment, structure triage during an incident, allocation of resources (crews, equipment, air support)
- **Human centered AI** would involve users in determining the answer, give them some agency, and help build trust in the result, even if the 'black box' is not fully understood.
- **Satellite data**: acquisition easy (purchase) but lower resolution results in higher error than drone-based imagery. Current imagery not always available.
- **Drone use**: Real time (or preplanning) collection of RGB, thermal, LiDAR, hyperspectral, and weather (a vertical profile of temp, humidity, wind) to inform decision making around (or before) fires. This involves collecting data on 'the fire triangle' of fuels/weather/topography.

Using AI to predict dispersion of Mountain Pine Beetles



Ramazi, Kunegel-Lion, Greiner, Lewis

- 2021 Predicting insect outbreaks using machine learning: A mountain pine beetle case study
- Machine-learning algorithms are potential solutions to this challenging problem due to their many successes across a variety of prediction tasks. However, there are many subtle challenges in applying them: identifying the best learning models and the best subset of available covariates (including time lags) and properly evaluating the models to avoid misleading performance-measures. We systematically address these issues in predicting the chance of a mountain pine beetle outbreak in the Cypress Hills area and seek models with the best performance at predicting future 1-, 3-, 5- and 7-year infestations. We train nine machine-learning models, including two generalized boosted regression trees (GBM) that predict future 1- and 3-year infestations with 92% and 88% AUC, and two novel mixed models that predict future 5- and 7-year infestations with 86% and 84% AUC, respectively.

https://www.alberta.ca/mountain-pine-beetle-overview.aspx

https://onlinelibrary.wiley.com/doi/full/10.1002/ece3.7921

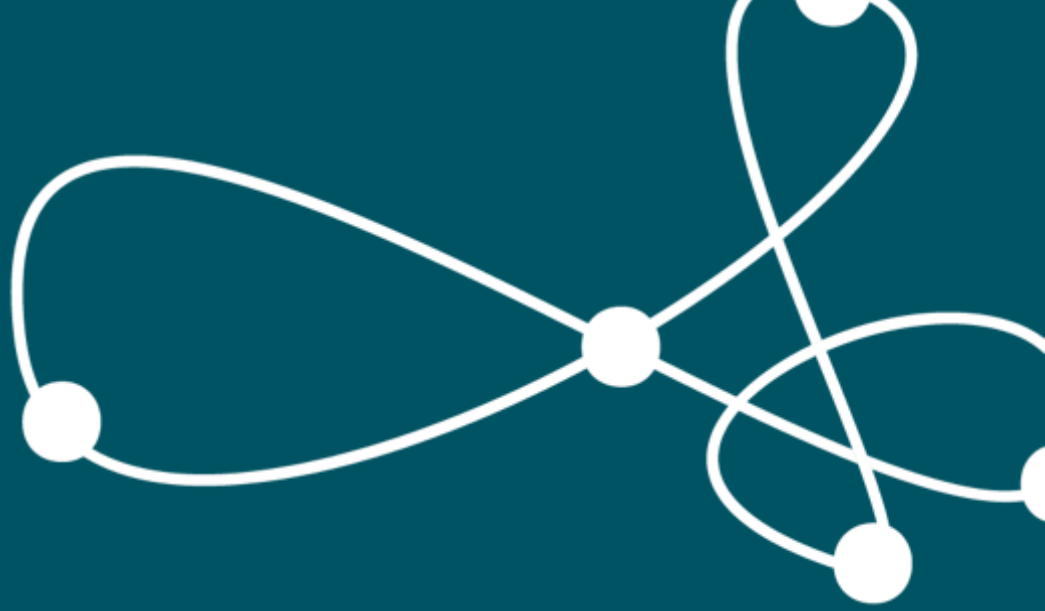
Curating a decade of soil sampling data

Derek MacKenzie

- A new University of Alberta project is, for the first time, harnessing information from thousands of soil samples into one big database to get an idea of how healthy the province's agricultural soils are.
- The two-year initiative brings together decades of data collection on soil samples from 44 benchmark sites across the province, so they can be analyzed through machine learning.
- Once completed, the comprehensive Database on Alberta Soil Health will support farmers in their sustainable land management practices, says project researcher Derek MacKenzie, a soil scientist in the Faculty of Agricultural, Life & Environmental Sciences.

https://www.ualberta.ca/folio/2022/07/new-soil-database-will-help-alberta-farmers-plot-out-sustainable-practices.html

Federal Initiatives



EFFIS

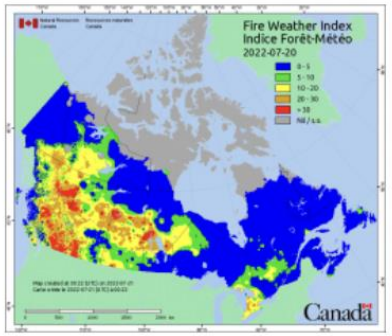
https://effis.jrc.ec.europa.eu

Canadian Wildfire Organization

Canadian Wildland Fire Information System

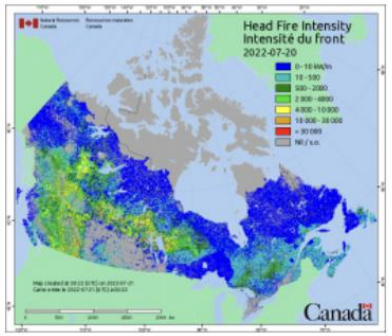
The **Canadian Wildland Fire Information System (CWFIS)** creates daily fire weather and fire behavior maps year-round and hot spot maps throughout the forest fire season, generally between May and September.

Fire Weather



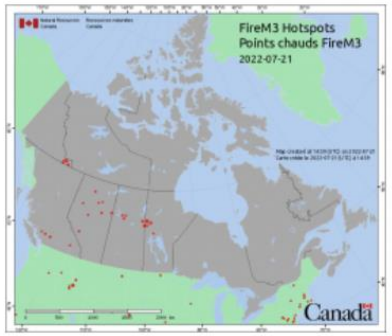
View the most recent Fire Weather map

Fire Behavior



View the most recent Head Fire Intensity

Fire M3 Hotspots



View the most recent Daily Hotspot map

https://www.canadawildfire.org/about

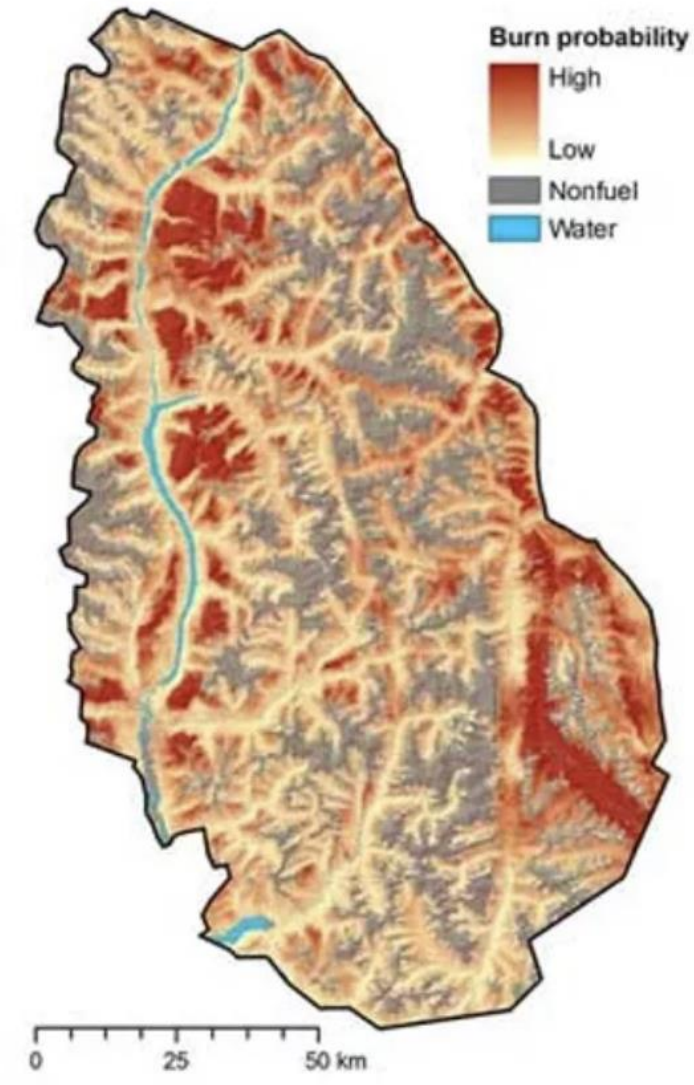
Burn-P3

Burn-P3 (probability, prediction, and planning) is a spatial fire simulation model that is used for land-management planning and wildland fire research. It uses the Prometheus fire-growth engine to simulate the ignition and spread of a very large number of fires. The inputs to Burn-P3 consist of fuels (e.g., vegetation), topography, weather, and patterns of fire ignitions. Its main output is a surface of fire probabilities, or burn probability map.

- Windows-based software application
- Computes burn probabilities for large landscapes
- Produces additional outputs, such as fire intensity maps
- Extracts fire statistics and simulated fire perimeters

System Requirements

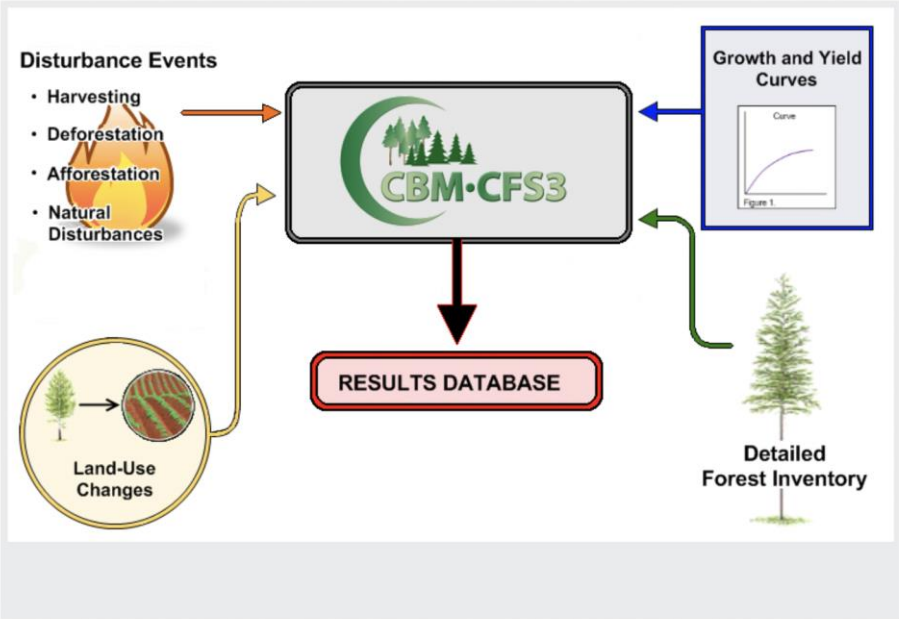
- Processor 4 cores
- 6 GB of RAM
- 10 GB of available disk space
- 64-bit Windows operating system, Windows 7/8/10
- Softwares: Microsoft .NET Framework 4.0, 64-bit Java 8.0 for Windows, and Prometheus 6.2.4



https://www.canadawildfire.org/about

Using AI models to manage fire resources

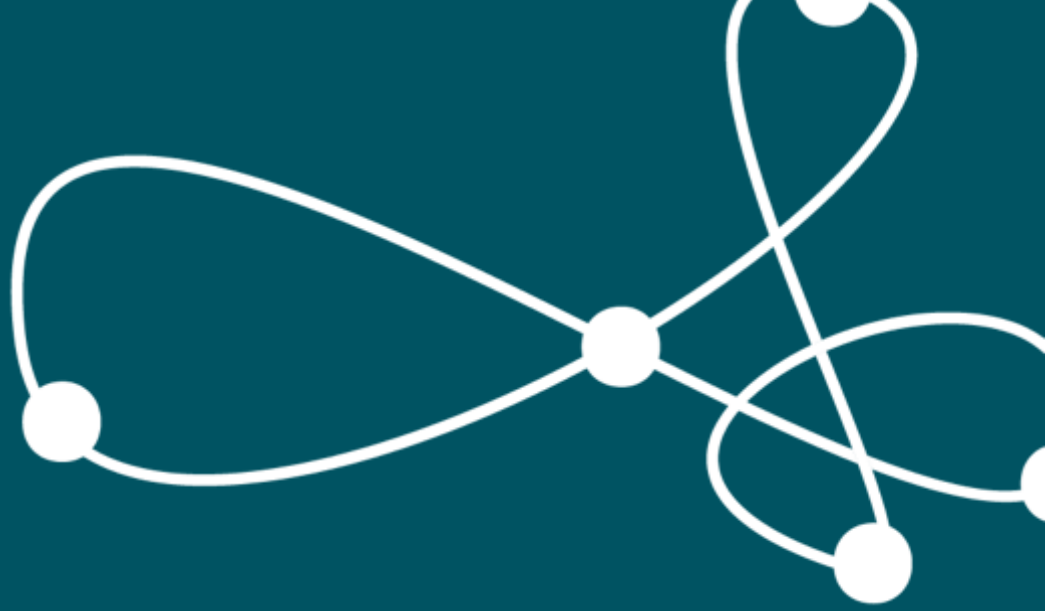
The national forest carbon monitoring, accounting and reporting system



Piyush Jain, Natural Resources Canada

- Meteorologists can reasonably forecast hot weather and lightning storms, but exactly where lightning will strike — and whether it will spark a wildfire — is nearly impossible to predict.
- That's why researchers are turning to machine learning in a bid to get ahead of catastrophic wildfires that they say are becoming more common and more severe.

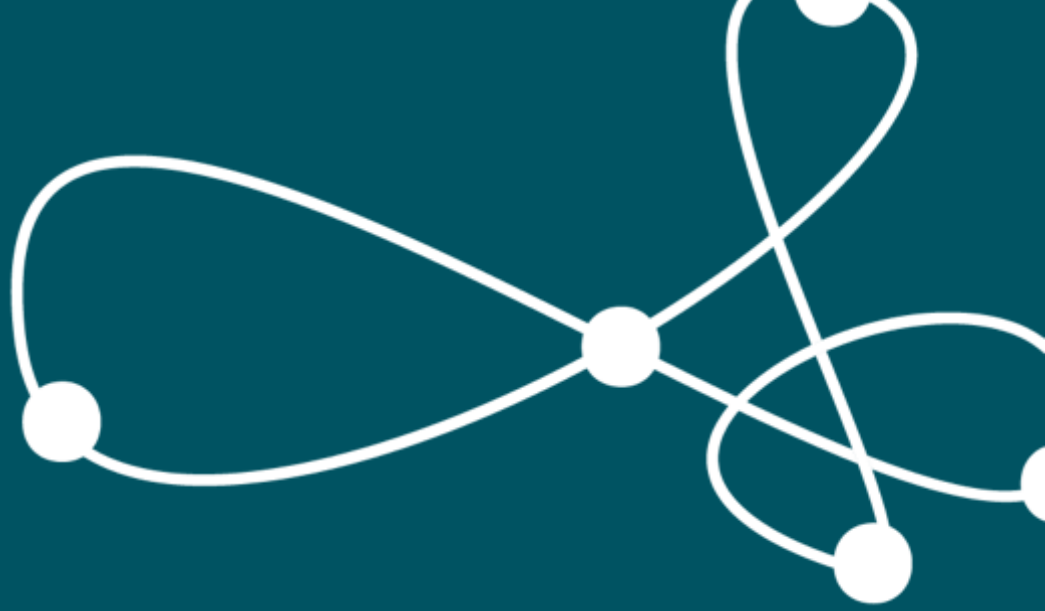
https://www.nrcan.gc.ca/science-data/research-centres-labs/forestry-research-centres/northern-forestry-centre/13485



Summary

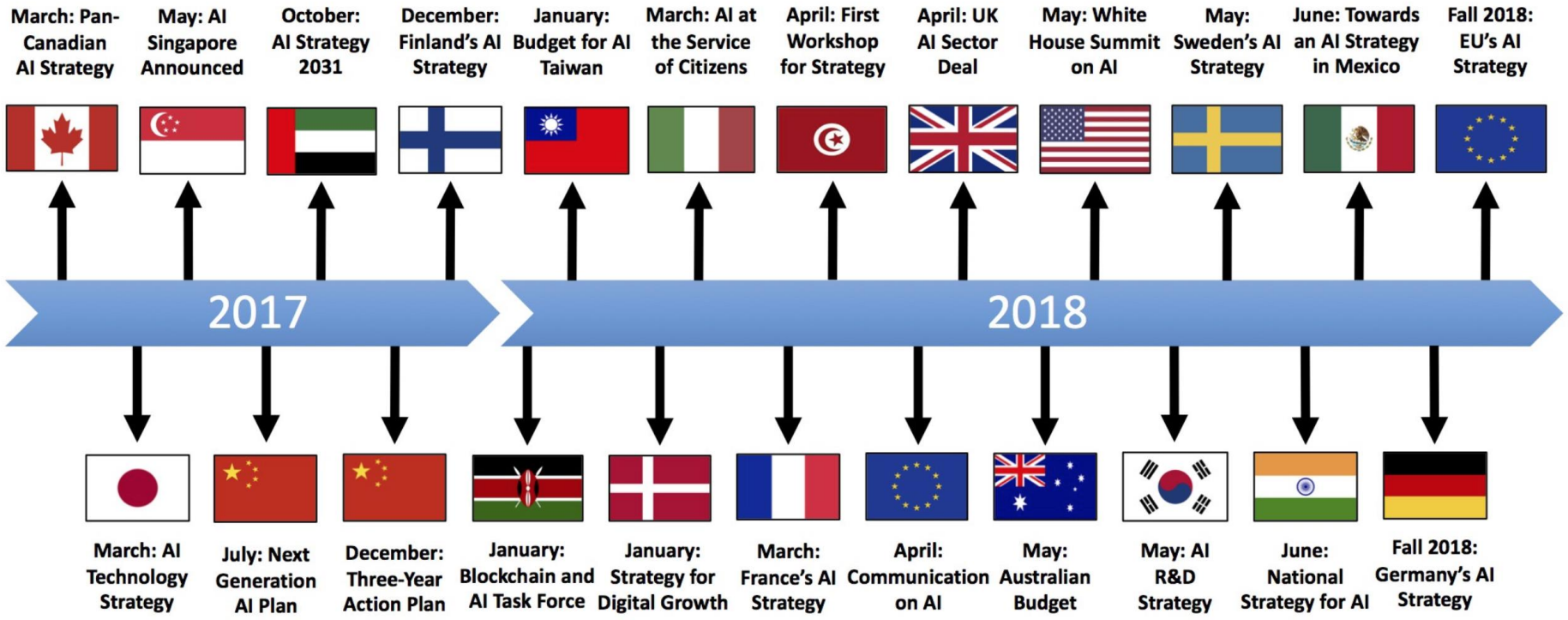
Summary

- Successful application of AI techniques is drive by data, data scientists and application subject matter experts.
- AI technologies are primitive from the viewpoint of capturing subject matter expertise, so all applications have to be designed within a framework of subject matter expert “in the loop.”
- Evaluation of applications is determined by subject matter KPIs, NOT AI technology usage KPIs.



Canadia AI/Amii Summary

Canadian Federal AI Strategy Development



https://medium.com/politics-ai/an-overview-of-national-ai-strategies-2a70ec6edfd

Canada's AI hubs

LED BY:

CIFAR CANADIAN INSTITUTE FOR ADVANCED RESEARCH | ICRA INSTITUT CANADIEN DE RECHERCHES AVANCÉES

Source: Canada - Dots by FreeVectorMaps.com

EDMONTON

amii

UNIVERSITY OF ALBERTA

MONTRÉAL

MILA

Université de Montréal

McGill UNIVERSITY

TORONTO-WATERLOO

VECTOR INSTITUTE

UNIVERSITY OF TORONTO

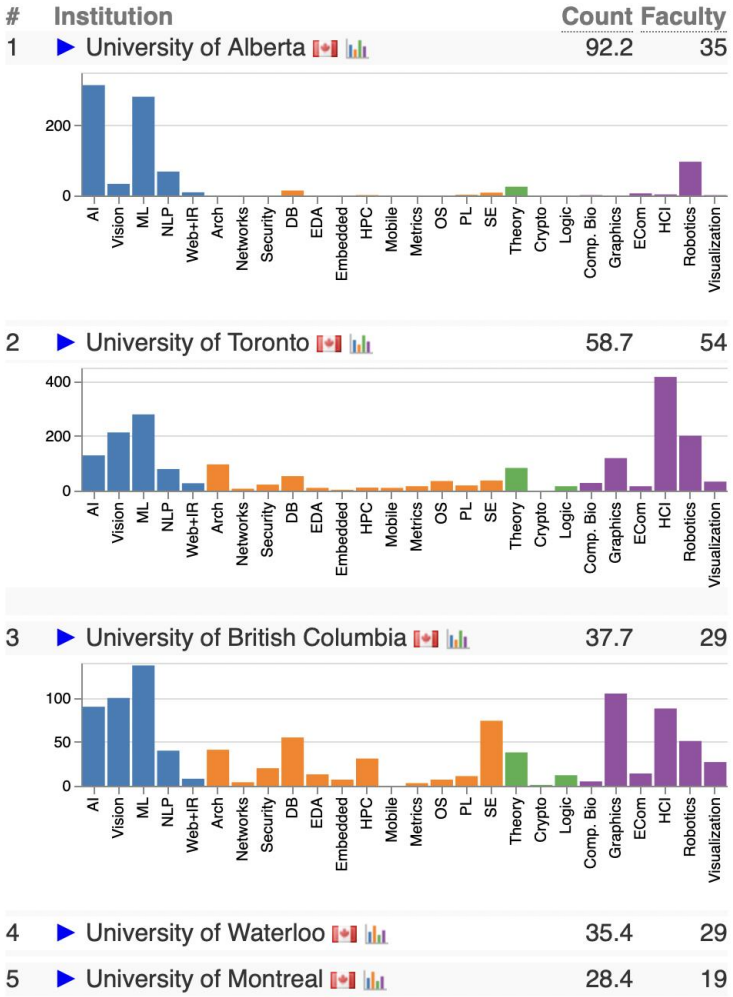
UNIVERSITY OF WATERLOO

CSRankings.Org on AI and Machine Learning

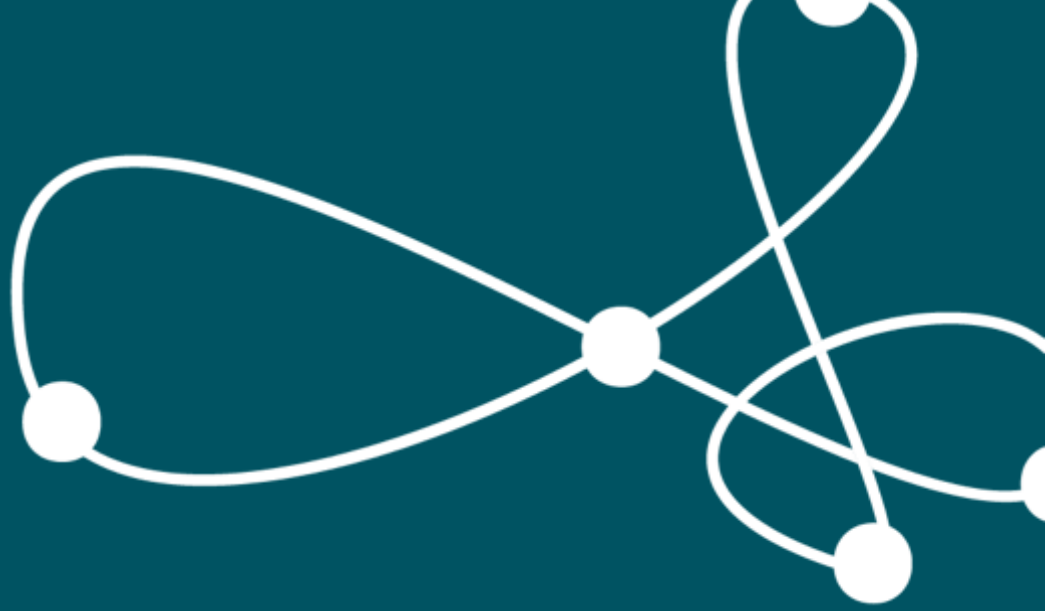
#	Institution	Count	Faculty
1	Carnegie Mellon University	162.8	85
2	Tsinghua University	130.4	80
3	University of Alberta	92.2	35
4	Cornell University	86.8	41
5	Technion	84.7	50
6	Univ. of California - Los Angeles	84.4	27
7	Nanjing University	83.8	52
7	Peking University	83.8	101
9	Massachusetts Institute of Technology	78.0	66
10	HKUST	71.9	31
11	Univ. of Illinois at Urbana-Champaign	65.3	52
12	Univ. of California - Berkeley	64.7	50
13	University of Massachusetts Amherst	62.4	37
14	Stanford University	62.3	50
15	University of Toronto	58.7	54

- There are measure choices to make *any* organization look like a leader (as at August 8, 2022).
- Since 2015, the Chinese national AI strategy has helped Tsinghua move from 10th to 2nd.
- Alberta is also growing, but more slowly, and has been in the top 5 since 2000.
- In 2017 Alberta rose to 2nd, and was then displaced by Tsinghua in 2018.

AMII performance comparisons

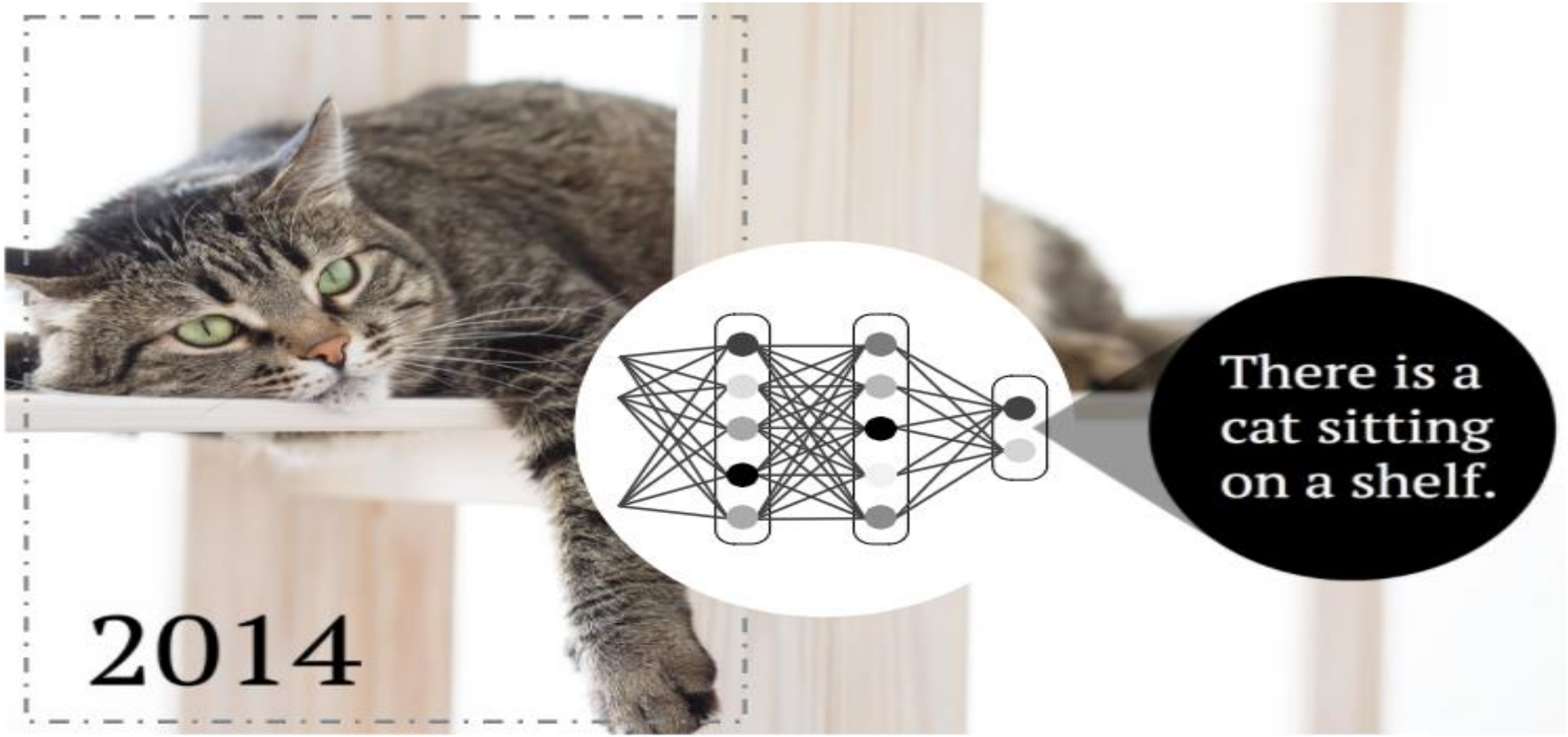


- Top 5 in Canadian AI publication production (as at August 8, 2022).
- Canadian AI excels at academic science, but lags behind USA, Europe, Asia in industrial exploitation.
- Federal programs have invested $125M (2017-2022), $443M (2022-2027) to increase AI activities in talent and commercial activity



Debugging Learned Models

This cat is not sitting



Slide courtesy of Hugo Larochelle, U de Sherbrooke

McCarthy's Advice Taker

IR-00182
July 24, 1958

THE ADVICE TAKER - Preliminary Report

by John McCarthy

The advice taker is a proposed program for solving problems by manipulating sentences in formal languages. The main difference between it and other programs or proposed programs for manipulating formal languages (the Logic Theory Machine of Newell, Simon and Shaw and the Geometry Program of Gelernter) is that in the previous programs the formal system was the subject matter but the heuristics were all embodied in the program. In this program the procedures are described as much as possible in the language itself and, in particular, the heuristics are all so described.

The main advantages we expect the advice taker to have is that its behavior will be improvable merely by making statements to it, telling it about its symbolic environment and what is wanted from it.

The main advantage we expect the advice taker to have is that its behavior will be improvable by merely making statements to it ...

Two alternative “debugging” methods

“The cat is not sitting, it is lying down on the shelf”