Artificial Intelligence and Machine Learning with Java

PredAI (JSR 381) and GenAI (LangChain4j)

Gran Sasso

Java-Friendly ML APIs

Zoran Sevarac - Univ of Belgrade - @zsevarac
Frank Greco - NYJavaSIG - @frankgreco

Zoran Severac

@zsevarac

AI Researcher
Univ of Belgrade, Serbia
JC

Frank Greco

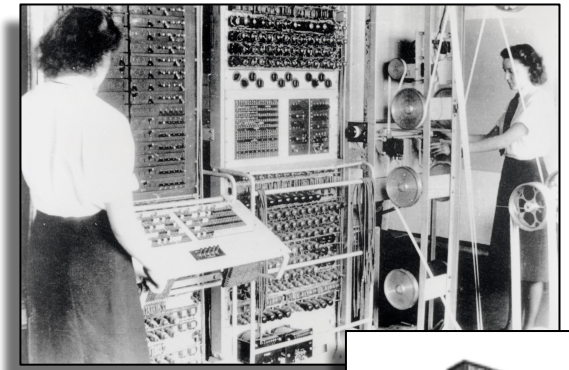
@frankgreco

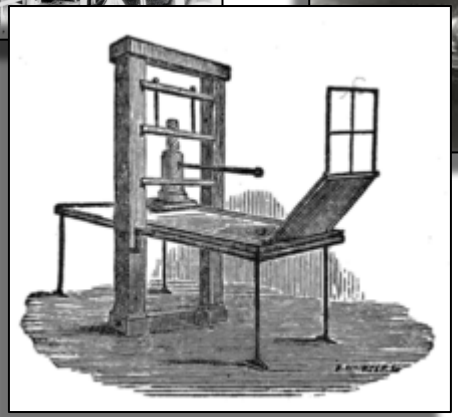
Chair
NYJavaSIG - NY Java User Group
JC, Cloud/Mobile Architect

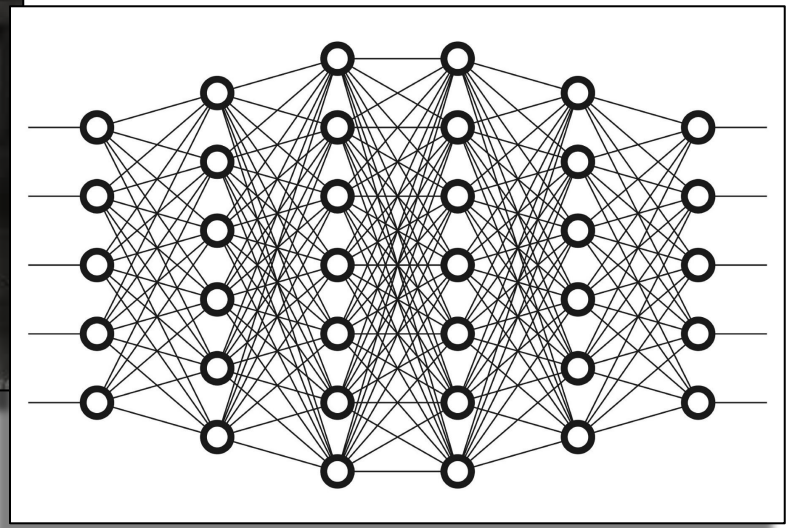
Goals

Instructional suggestions for educators and Java user group leaders on concepts, basic skills, and code walkthroughs that will motivate new Java developers to use the capabilities of JSR 381 (Visual Recognition) for PredAI, and LangChain4j for GenAI.

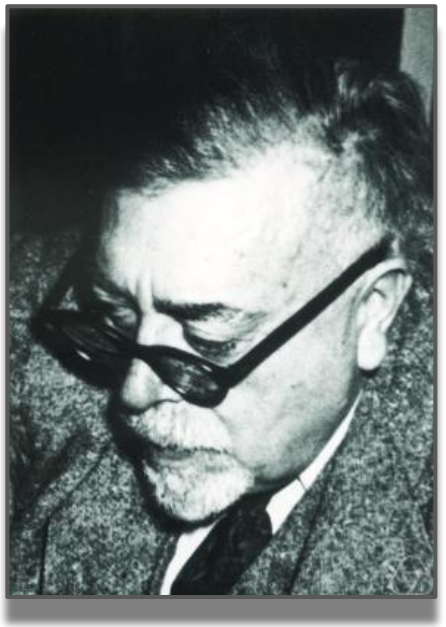
Machine Learning - Huge Impact on Everything...



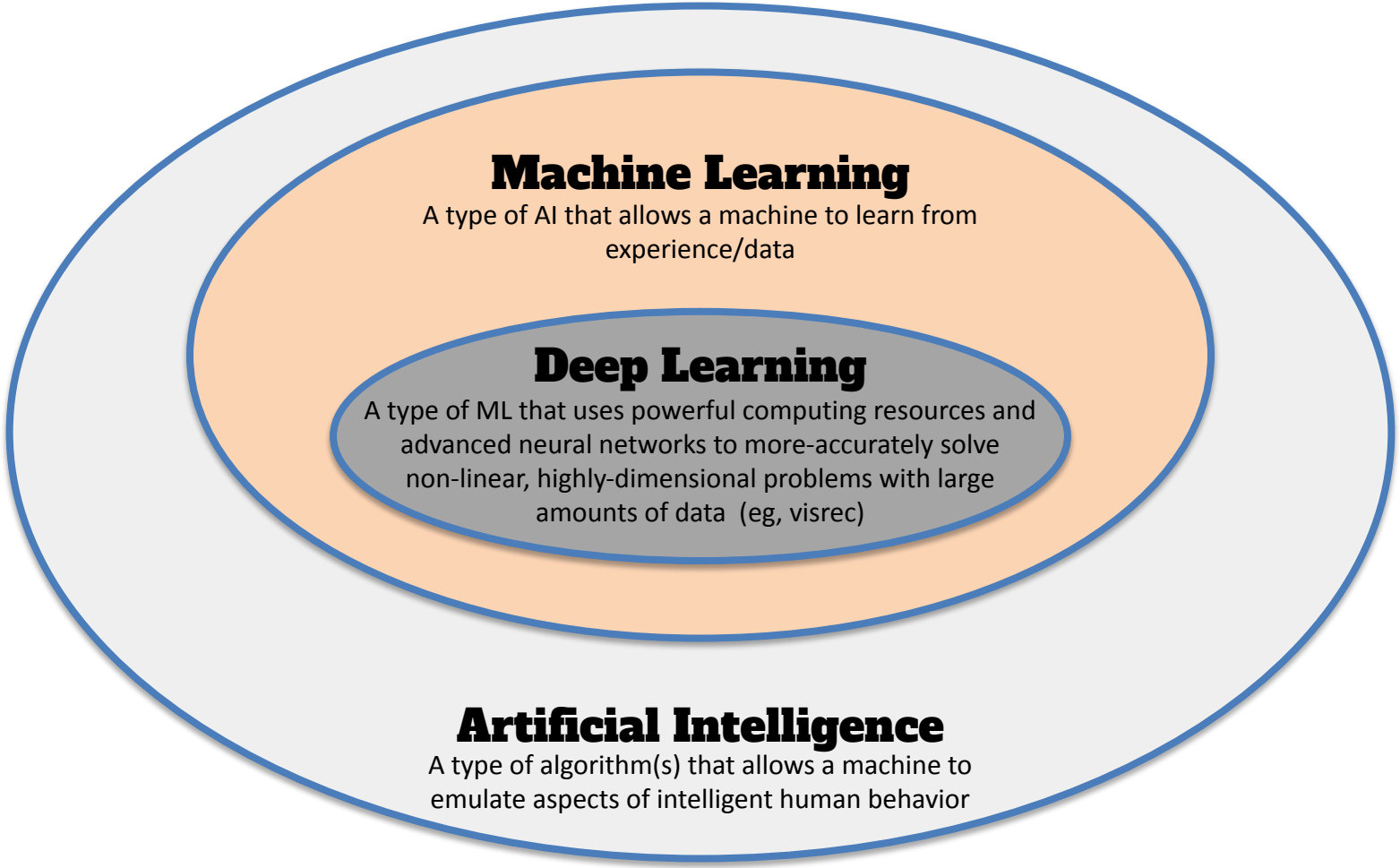


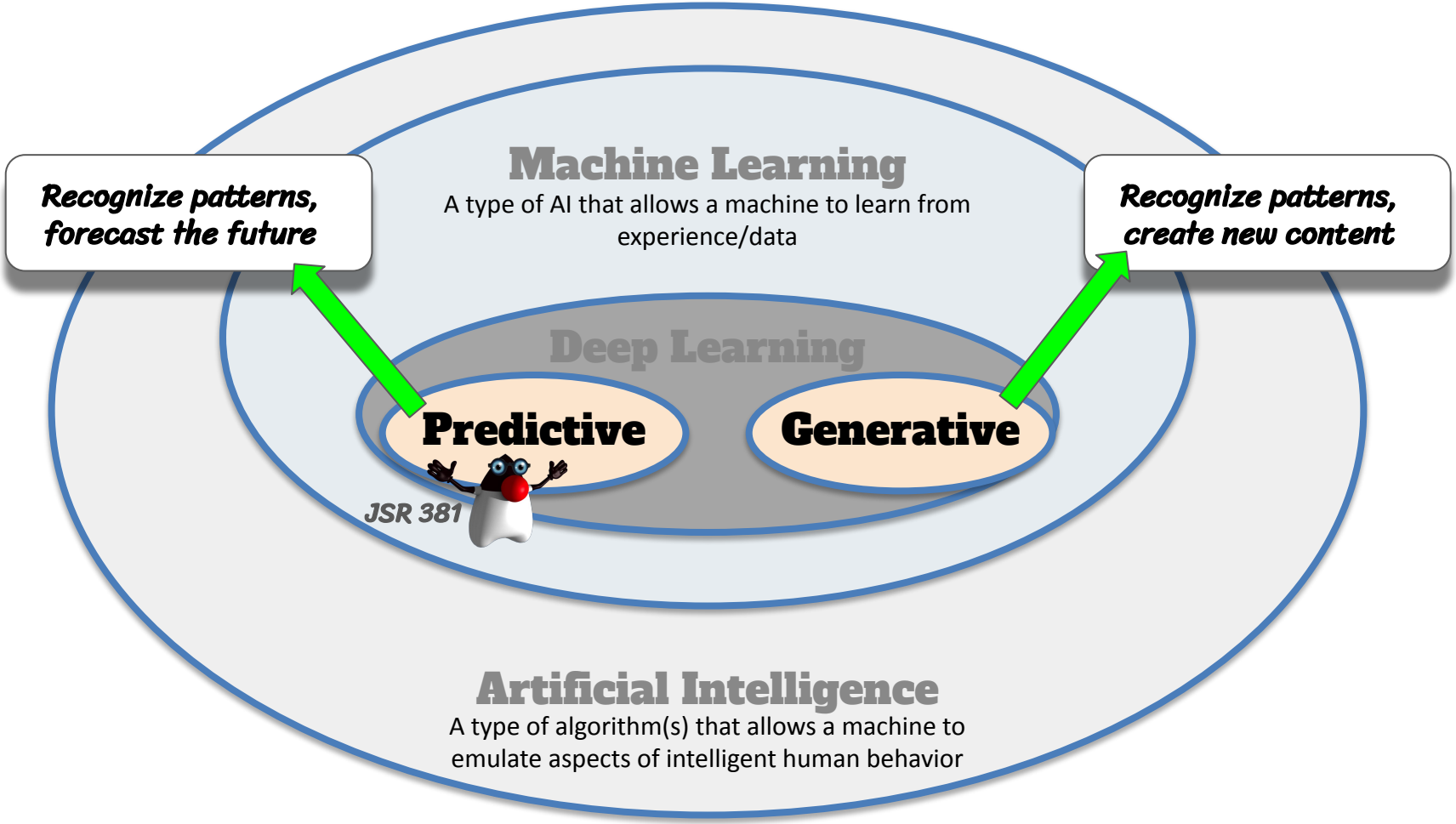


“One of the most interesting aspects of the world is that it can be considered to be made up of patterns”



Norbert Wiener (1948) - 1894-1964 - MIT


Machine Learning
A type of AI that allows a machine to learn from experience/data
Deep Learning
A type of ML that uses powerful computing resources and advanced neural networks to more-accurately solve non-linear, highly-dimensional problems with large amounts of data (eg, visrec)
Artificial Intelligence
A type of algorithm(s) that allows a machine to emulate aspects of intelligent human behavior


Recognize patterns, forecast the future
Machine Learning
A type of AI that allows a machine to learn from experience/data
Recognize patterns, create new content
Deep Learning
Predictive
Generative
JSR 381
Artificial Intelligence
A type of algorithm(s) that allows a machine to emulate aspects of intelligent human behavior

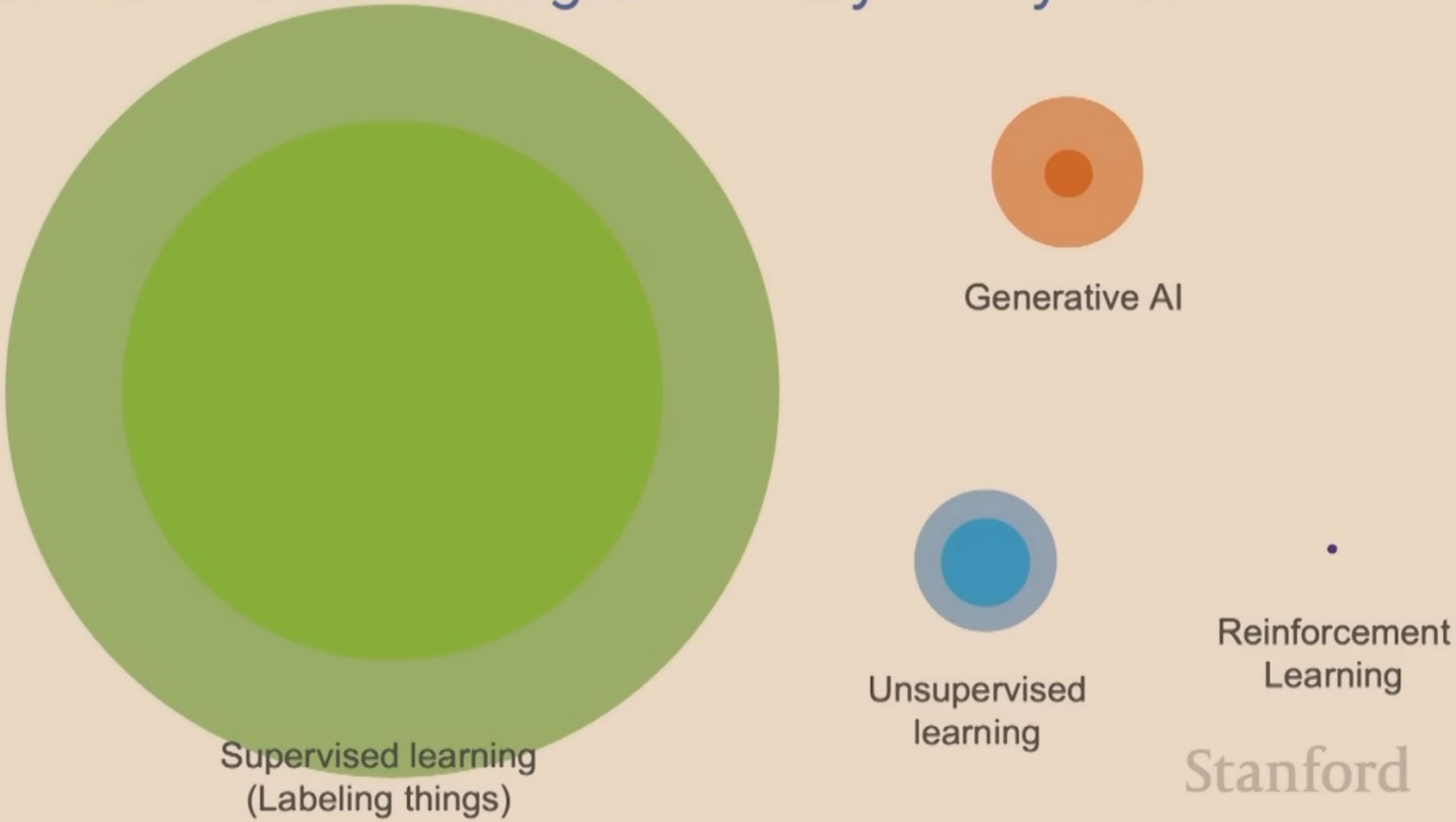
Predictive AI and Generative AI

Most of AI economic value (03/24) comes from PredAI: weather, image detection and classification, financial services, buying behavior, up/cross-selling…

PredAI has been deployed successfully for past 10-15 yrs
PredAI is probably worth at least $100B just to Google (and Amazon/Netflix)

GenAI typically used for more “creative”, content generation
GenAI growth and potential is huge. Market value may match PredAI in 3-5 yrs

Value from AI technologies: Today → 3 years



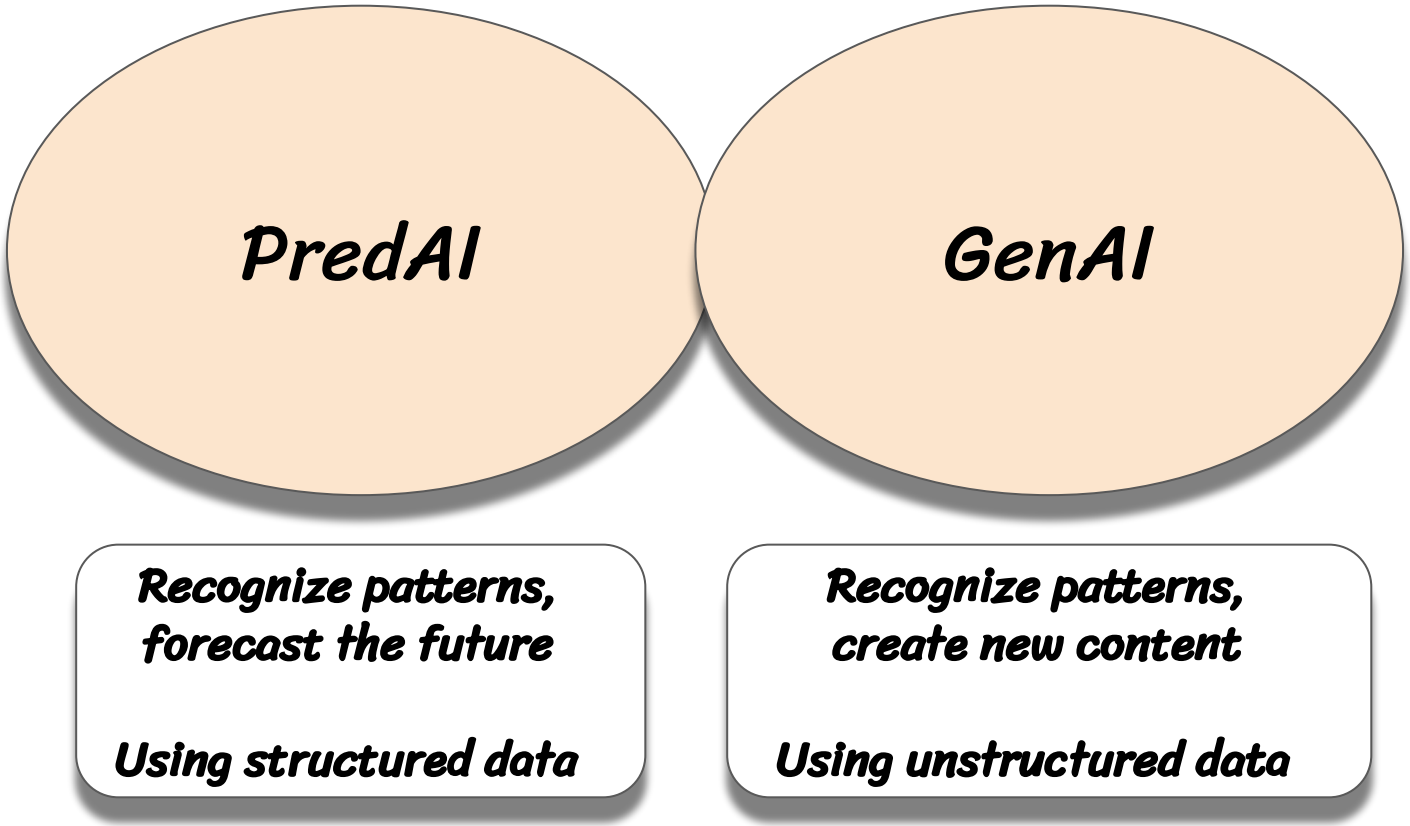
Stanford

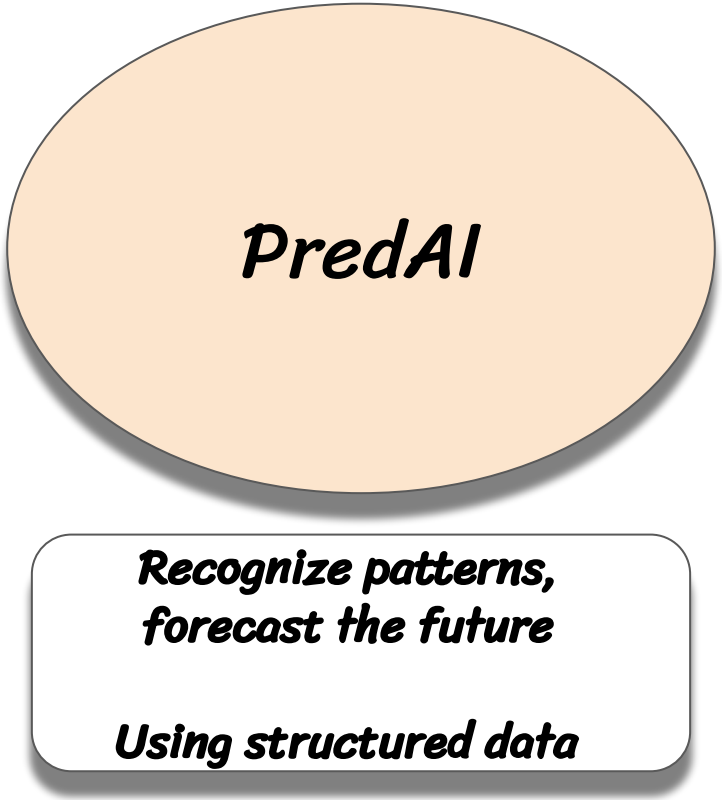
https://www.youtube.com/watch?v=5p248yoa3oE

July 26, 2023
Andrew Ng

Java Needs to be a First Class Citizen for AI/ML

- Don’t want another JavaScript single-language scenario
- Python is a good language
  - There should be others to express the world’s creativity.
- Java - An awesome language with a huge ecosystem with 10-12M developers
- Majority of ML APIs aren’t Java-friendly or AppDev-friendly


PredAI
GenAI
Recognize patterns, forecast the future
Using structured data
Recognize patterns, create new content
Using unstructured data


PredAI
Recognize patterns,
forecast the future
Using structured data

Don’t boil the ocean… [heat up some coffee first]

VisRec - JSR #381 Visual Recognition for Java

Spec Leads and Expert Group

Zoran Sevarac (ML researcher - Univ of Belgrade)
Frank Greco (NYJavaSIG, AI/ML Consultant)
Kevin Berendsen (OpenValue, Senior Software Eng)

JSR 381 - Other Contributors and Advisors

- Frank Liu (Amazon)
- Constantin Drabo
- Amit Nagesh
- Marissa Staller
- Eric Bruno (fmr Perrone)
- Anakar Parida
- Nikita Ivanov (Gridgain)
- James Weaver (IBM)
- Werner Keil
- Jyoti Buddha
- Guillaume Laforge (Google)
- Ed Burns (fmr Oracle)
- Nishant Raut (Mumbai JUG)
- Sandhya Kapoor (fmr IBM)

JSR #381 - Visual Recognition API

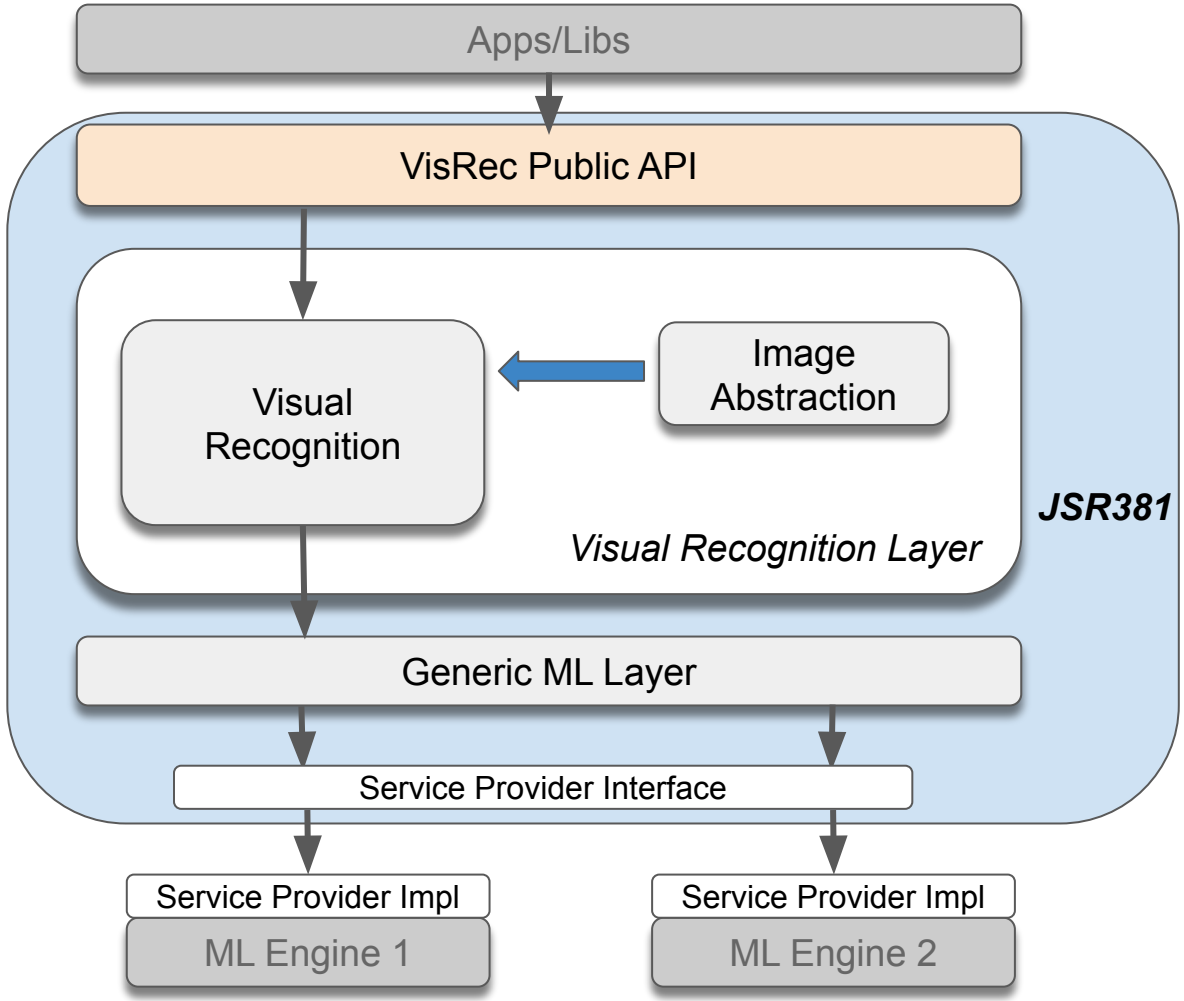
- Java-friendly standard API for computer vision tasks using ML
- Also provides generic ML API
- Designed for Java app devs
- Multiple implementations
- High level abstractions
- More readable and easier to use

VisRec JSR #381

https://github.com/JavaVisRec/visrec-api/wiki/Getting-Started-Guide

VisRec Design goals and principles

- Easy to use by non-experts in VisRec/ML domain
- Enable Java Developers to leverage their Java skills
- Easy to implement for existing ML and imaging libraries
- Reusable design for applied ML applicable to other domains



JSR 381 Architecture

JSR381 API
Code Walkthrough

Pains, barriers and preferences for PredAI in Java

- Many different incompatible data formats
- Many different machine learning algorithms
- Many confusing configuration parameters
- Clear task oriented interface that hides implementation details
- Usage does not require understanding of implementation and configuration (like `Collections.sort()`)
- Simple and portable integration into existing Java apps and devices

VisRec API Reference Implementation

Based on Deep Netts Community Edition

https://www.deepnetts.com/blog/deep-netts-community-edition

A Java based deep learning library provided by www.deepnetts.com

Deep Netts Professional edition

- Free for personal and educational use
- Higher performance implementation
- Advanced visual AI tools
- https://www.deepnetts.com/deep-netts-platform/

How to get started

github.com/JavaVisRec/visrec-api/wiki/Getting-Started-Guide

JSR381 Information

- JCP.org https://jcp.org/en/jsr/detail?id=381
- JSR project on GitHub github.com/JavaVisRec
- Getting Started github.com/JavaVisRec/visrec-api/wiki/Getting-Started-Guide
- API github.com/JavaVisRec/visrec-api
- Reference Implementation github.com/JavaVisRec/visrec-ri
- Examples github.com/JavaVisRec/jsr381-examples
- Mailing list visrec@groups.io

JSR 381
Visual Recognition for Java

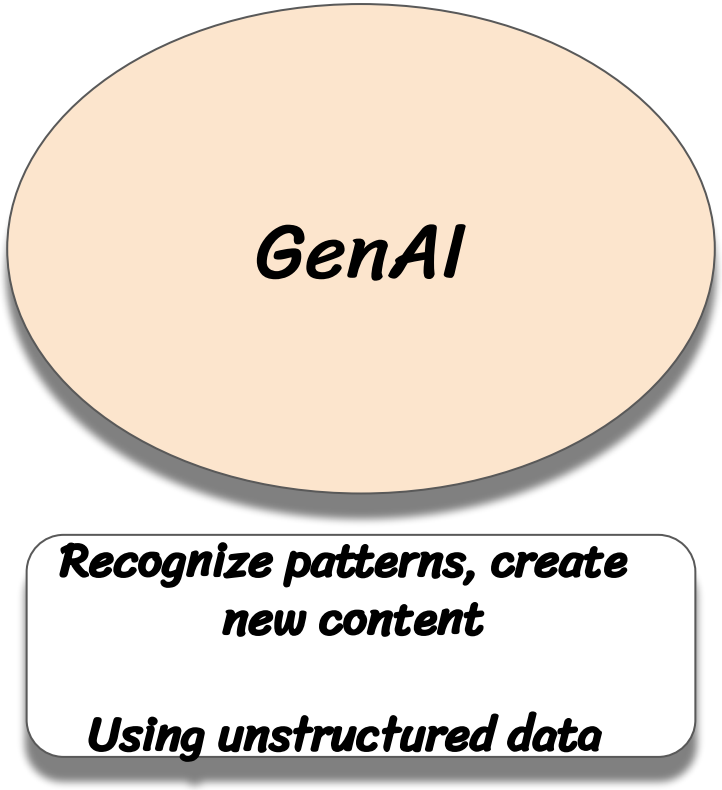
Finally,
A Java-Friendly ML API

Please provide feedback and **contribute**!

Java is a **great** language for AI/ML!

Foundations of AI and Machine Learning for Java Developers

Course for beginners/intermediates


GenAI
Recognize patterns, create new content
Using unstructured data

LangCha... Introduction Get Started Tutorials Integrations Useful Materials Examples Docu chatbot Javadoc GitHub Twitter Discord Search ⌘+K

LangChain4j

Supercharge your Java application with the power of LLMs

Introduction

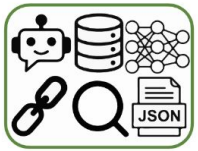
https://docs.langchain4j.dev/

Easy interaction with LLMs and Vector Stores

All major commercial and open-source LLMs and Vector Stores are easily accessible through a unified API, enabling you to build chatbots, assistants and more.

Tailored for Java

Smooth integration into your Java applications is made possible thanks to Quarkus and Spring Boot integrations. There is two-way integration between LLMs and Java: you can call LLMs from Java and allow LLMs to call your Java code in return.

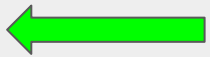


Agents, Tools, RAG

Our extensive toolbox provides a wide range of tools for common LLM operations, from low-level prompt templating, chat memory management, and output parsing, to high-level patterns like Agents and RAG.

```java
public class HelloWorldSystemAnthropic {
    public static void main(String[] argv) {
        ChatModel cmodel = AnthropicChatModel.builder()
                .apiKey(System.getenv("ANTHROPIC_API_KEY"))
                .modelName(CLAUDE_3_5_SONNET_20241022)
                .timeout(Duration.ofSeconds(120))
                .maxTokens(256)
                .build();

        List<ChatMessage> messages = new ArrayList<>();

        SystemMessage sysmsg = new SystemMessage("""
                You are a polite Java expert explaining concepts to a C++ software developer.
        """);
        messages.add(sysmsg);

        UserMessage usrmsg = UserMessage.from("What are Java lambdas?");
        messages.add(usrmsg);

        var answer = cmodel.chat(messages);

        System.out.println(answer.aiMessage().text());
    }
}
```

LangChain4j



Send to LLM

Thanks!

Questions / Comments?

@zsevarac

@frankgreco