“nn” things every Java Developer should know about AI/ML/DL

Mani Sarkar
@theNeomatrix369

About me

- Member of the ***London Java Community***
- Java / JVM, Polyglot Developer
- Code quality, testing, performance, DevOps, deep affinity for AI/ML/DL, NN, topics
- ***Strengthening*** **teams and helping them** ***accelerate***
- Involved in F/OSS i.e. GraalVM, AdoptOpenJDK and other projects and developer communities

Mani Sarkar
@theNeomatrix369

Also...

- Java Champion
- And Oracle Groundbreaker Ambassador

Click here for more info on Freebies!

Disclaimer

- Sharing my ideas, but ***YMMV***
- Possibly missed one or more things or made mistakes, *I ask for forgiveness*
- Not clear for some, *my apologies*
- Lots of info and resources
- Lots of takeaway, please go home and do some more research
- Please contribute and share, please tweet!

Get the slide deck (just now) at

http://bit.ly/nn-things-java-dev-ai-ml-dl

Thank you!

Martin Toshev and
Martin Patsov!

Agenda

Why “nn” things?	Demo, maybe some code!
Timeline: how it started for me...	Insights, shares and tips...

What we won't cover...

Pros or cons of ...

What is the best library / framework to do "xyz" in?

We may name some topics but we won't go into the details of them. We name them or express an opinion so we have a starting point in our learning journey!

Celebrating 25 years of Java

Lots of apps we all want to know about...

"The 25 greatest Java apps ever written" blog post by **Alexa Weber Morales**

...for more go here

Why this talk or such talks?

They are all doing an awesome job and a service to us!

Many can show how to install or configure “xyz”!

Many have shown us how to code super fast and at a presentation!

Many have talked about frameworks, APIs, algorithms, theories, and the works!

But not many talks about how to get there, how each one of us can become better?

I could just show you some code

Share some links and videos

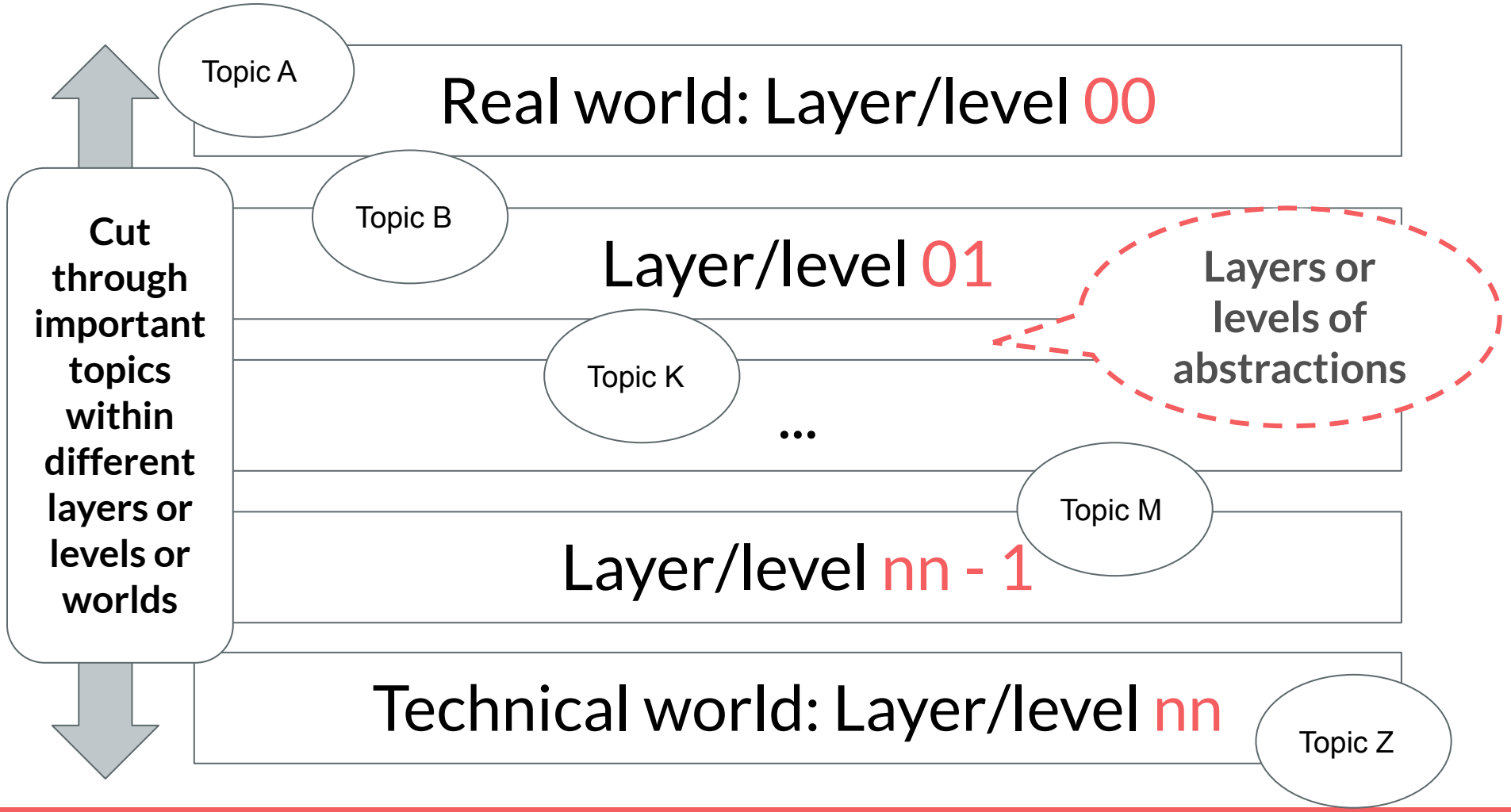
Show some demos!

And share some tips!

Instead I'll try to share perspectives, on how things sit together in my experience!

Share approaches and how I try to go around things!

Share methods and maybe some of my work! In additional to all other useful resources!


Topic A
Real world: Layer/level 00
Topic B
Layer/level 01
Layers or levels of abstractions
Topic K
...
Topic M
Layer/level nn - 1
Technical world: Layer/level nn
Topic Z
Cut through important topics within different layers or levels or worlds

Why “nn” things?

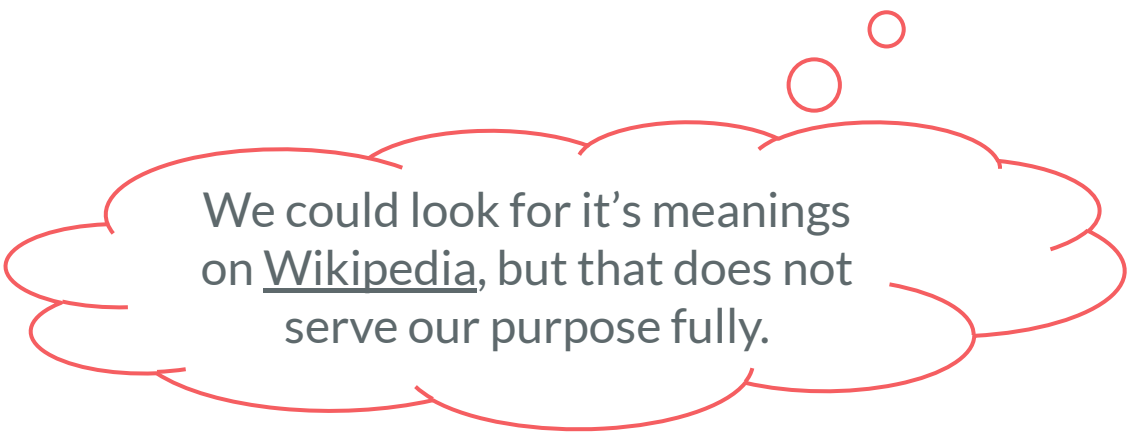
Each time we know “n” things...

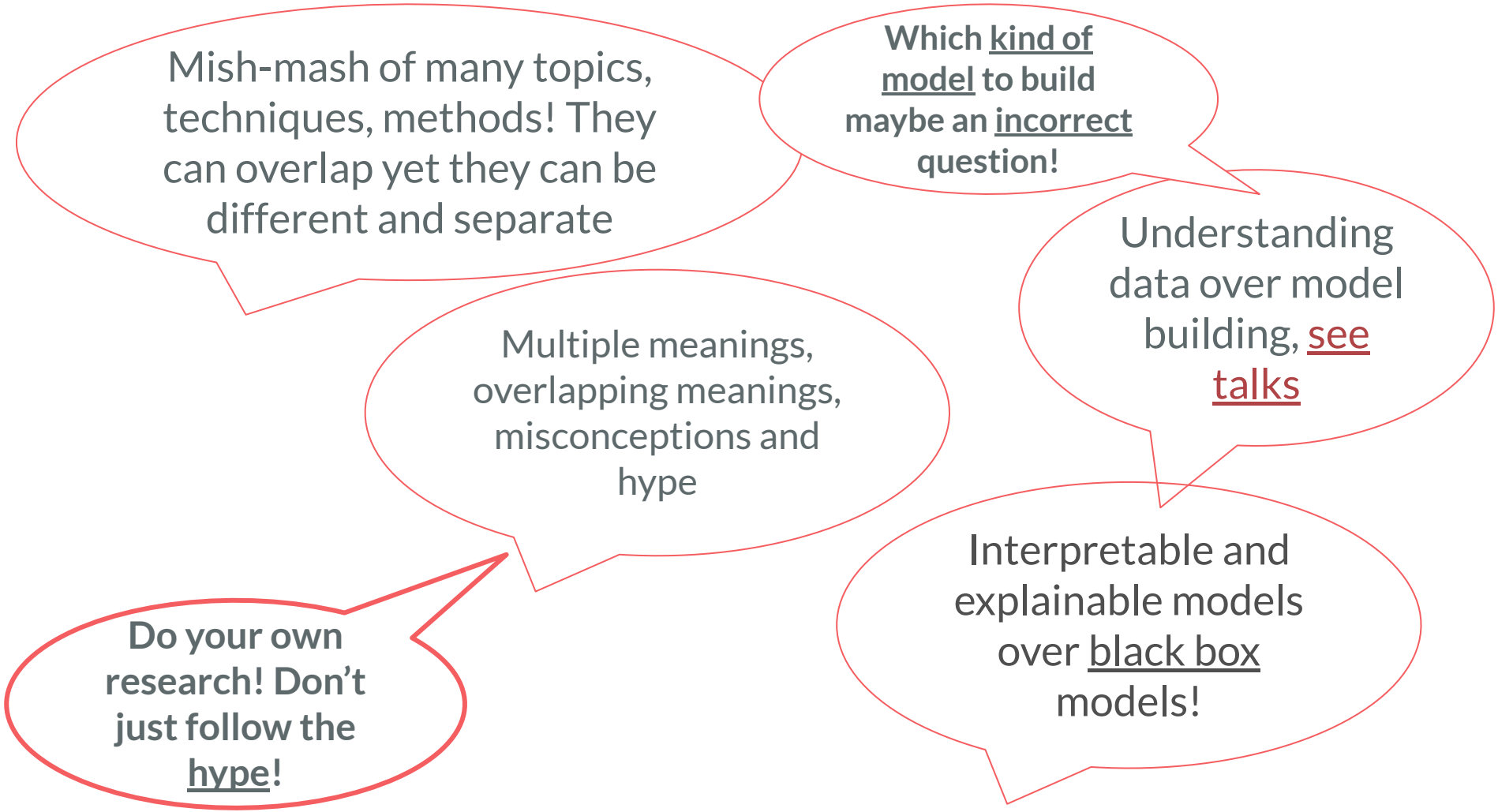
We find that there are “n” plus “more” things to know

It can give rise to something like “imposter syndrome” or just “overwhelm” us!

Only way forward is to remove such impediments!

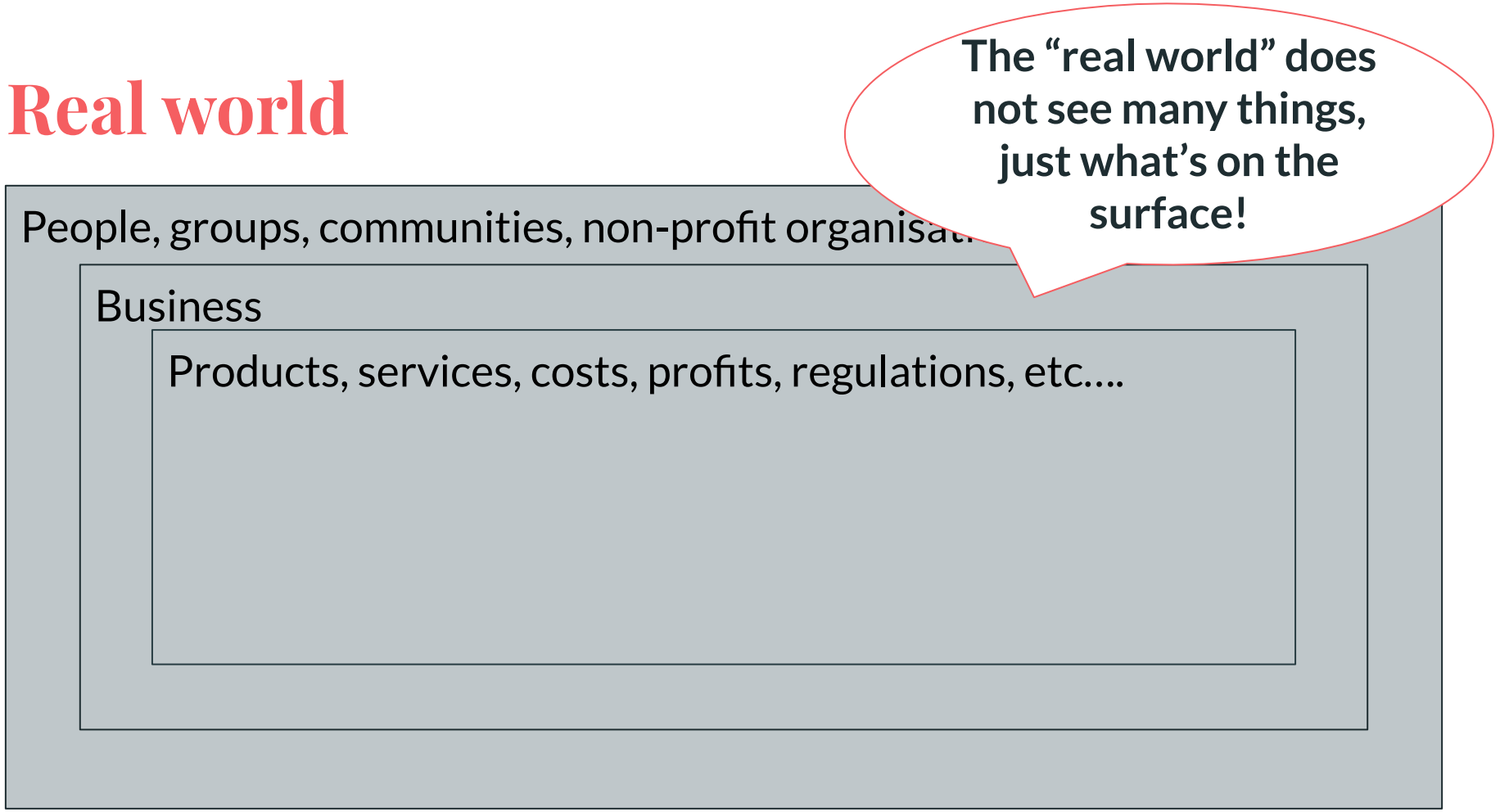
What is AI / ML / DL ?



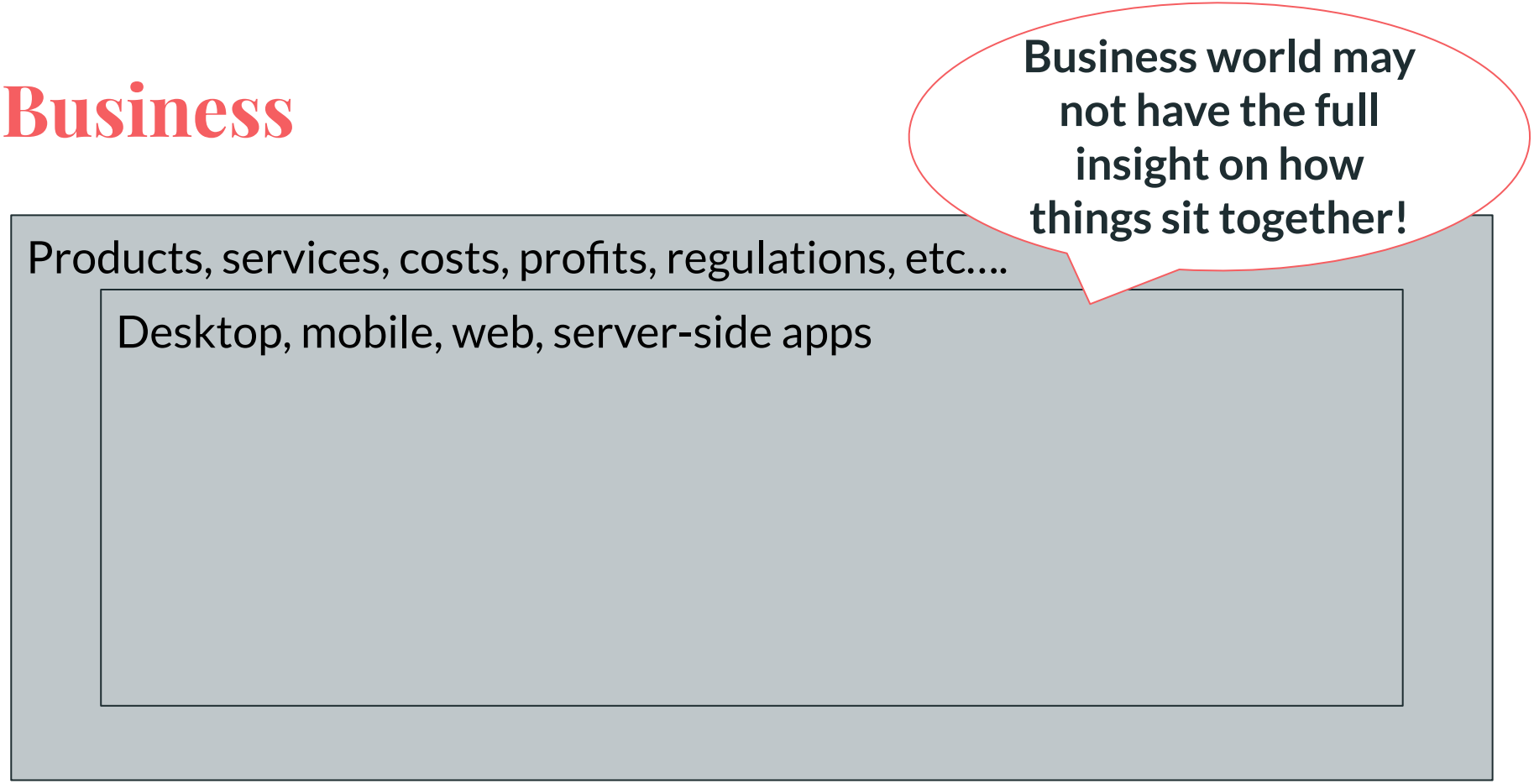

Mish-mash of many topics, techniques, methods! They can overlap yet they can be different and separate
Which kind of model to build maybe an incorrect question!
Understanding data over model building, see talks
Multiple meanings, overlapping meanings, misconceptions and hype
Interpretable and explainable models over black box models!
Do your own research! Don't just follow the hype!

Levels of understanding

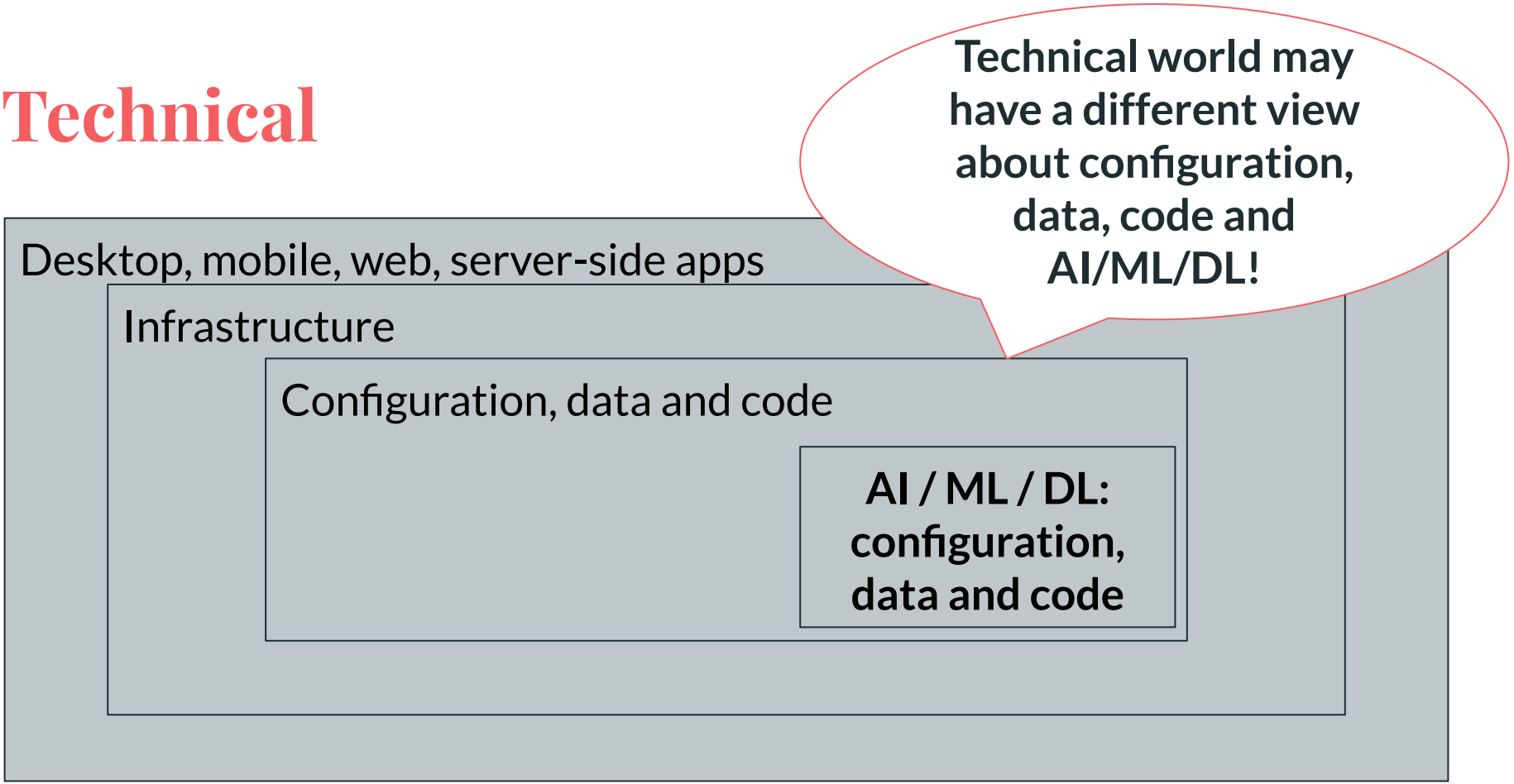
Real world


The “real world” does not see many things, just what’s on the surface!
People, groups, communities, non-profit organisat
Business
Products, services, costs, profits, regulations, etc….

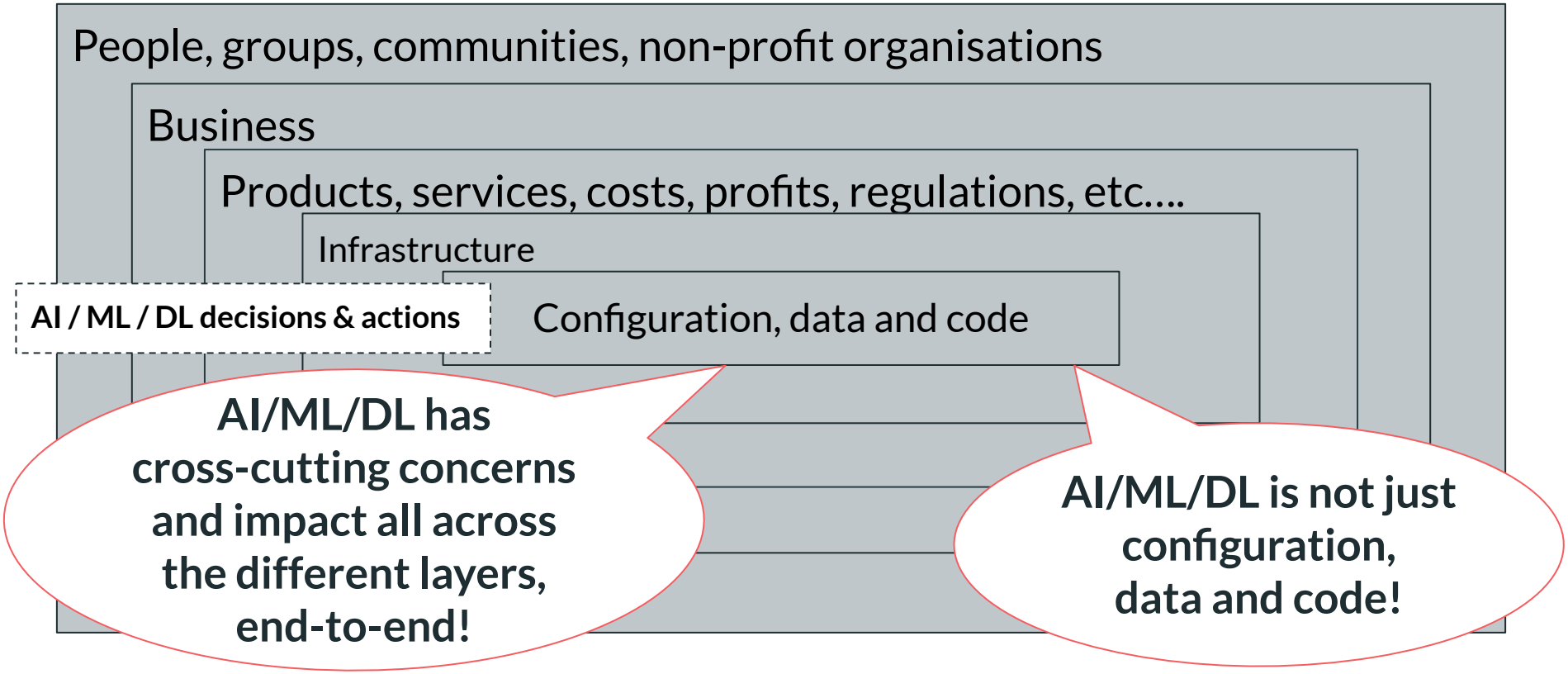
Business


Business world may not have the full insight on how things sit together!
Products, services, costs, profits, regulations, etc....
Desktop, mobile, web, server-side apps

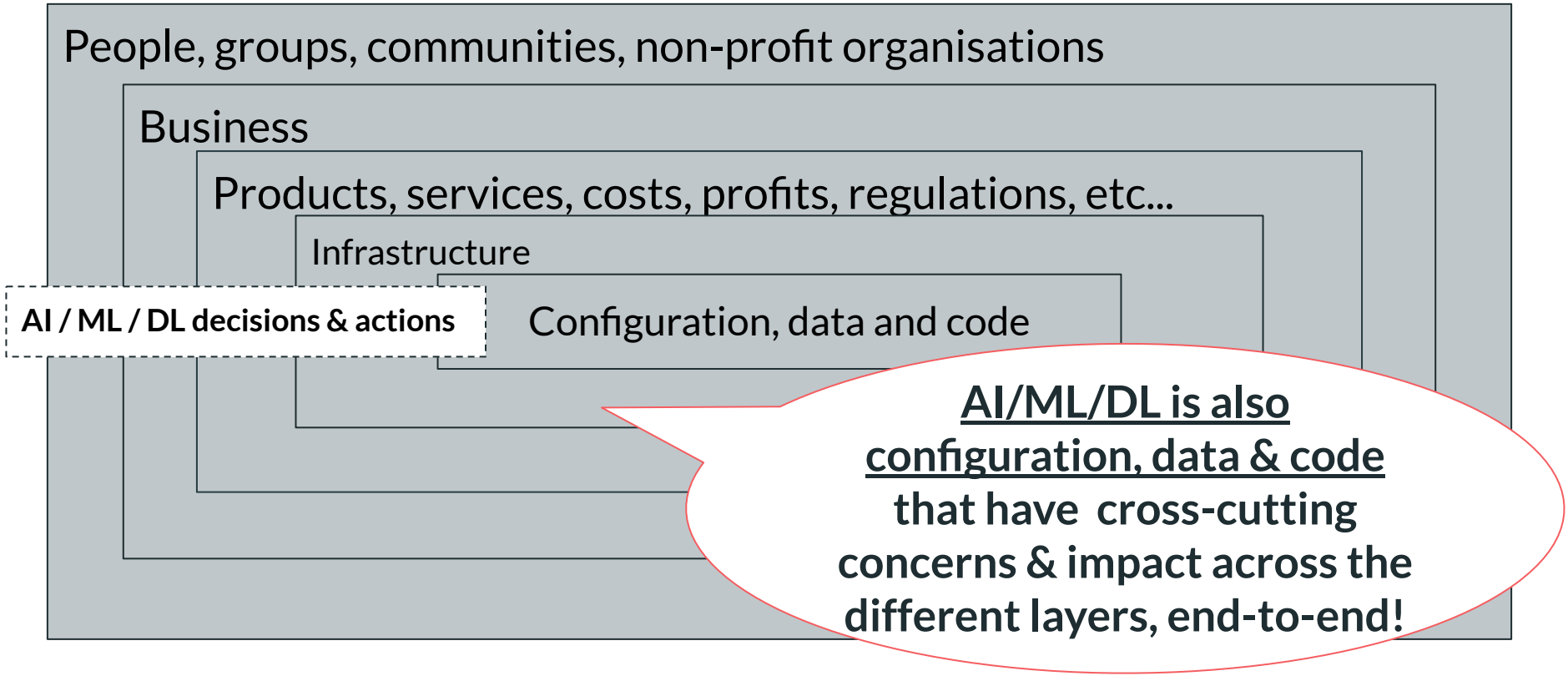
Technical


Technical world may have a different view about configuration, data, code and AI/ML/DL!
Desktop, mobile, web, server-side apps
Infrastructure
Configuration, data and code
AI / ML / DL: configuration, data and code

In reality


People, groups, communities, non-profit organisations
Business
Products, services, costs, profits, regulations, etc....
Infrastructure
Configuration, data and code
AI / ML / DL decisions & actions
AI/ML/DL has cross-cutting concerns and impact all across the different layers, end-to-end!
AI/ML/DL is not just configuration, data and code!

In reality


People, groups, communities, non-profit organisations
Business
Products, services, costs, profits, regulations, etc...
Infrastructure
Configuration, data and code
AI / ML / DL decisions & actions
AI/ML/DL is also configuration, data & code that have cross-cutting concerns & impact across the different layers, end-to-end!

And then there are other realities!

How it all started for me?

Timeline

- January to November 2018: collecting links
- 2018 December: creation of **awesome-ai-ml-dl**
- Early 2019: Better NLP
- 2019: collaboration with Virgilio (reviewing guides)
- May 2019: talk on data, see talks slide
- Mid 2019: another Better NLP presentation
- Rest of 2019: blogs on AI/ML/DL and example projects
- Late 2019, early 2020: online competitions
- 2020: more links added to **awesome-ai-ml-dl**

Thanks to Yolande Poirier for posting all those AI/ML/DL links on @java between 2016 and 2018

Thanks to the Meet-a-Mentor initiative by LJC during 2018-19, for holding the ML Study Group!

Thanks to Yaz for the Tensorflow meetups in London, UK during 2018 and 2019

How I learn?

Top-down, outside in learning, see Rachel Thomas' talk (fast.ai)

Mind-maps, gathering reading lists, organising them, categorising them

Nodes of Knowledge (see Appendix, previous talks)

Automatic chunking, Learning to Learn course by Coursera

List building: helps know the landscape and understand the context: birds-eye-view

List building: helps prioritise, order, remove or add topics you like, or don’t like or don’t know

AI/ML/DL git repo

https://github/neomatrix369/awesome-ai-ml-dl

README.md

Awesome AI-ML-DL

awesome | License CC BY-SA 4.0

Better NLP: docker pulls 55

NLP Java: docker pulls 310 | NLP Clojure: docker pulls 8 | NLP Kotlin: docker pulls 9 | NLP Scala: docker pulls 6 | NLP using DL4J (cuda) docker pulls 88

Dataiku DSS: docker pulls 26 | Grakn: docker pulls 1.7k | Jupyter-Java: docker pulls 25 | MLPMNist using DL4J: docker pulls 46 | Zeppelin: docker pulls 62

Awesome Artificial Intelligence, Machine Learning and Deep Learning as we learn it. Study notes and a curated list of awesome resources of such topics.

This repo is dedicated to engineers, developers, data scientists and all other professions that take interest in AI, ML, DL and related sciences. To make learning interesting and to create a place to easily find all the necessary material. Please contribute, watch, star, fork and share the repo with others in your community.

Watching the repo will keep you posted of all the changes (commits) that go into the repo.

Also, please SPONSOR us, find out how-to!

- General
- Artificial Intelligence
- Automation
- Competitions
- Courses

No packages published
Publish your first package

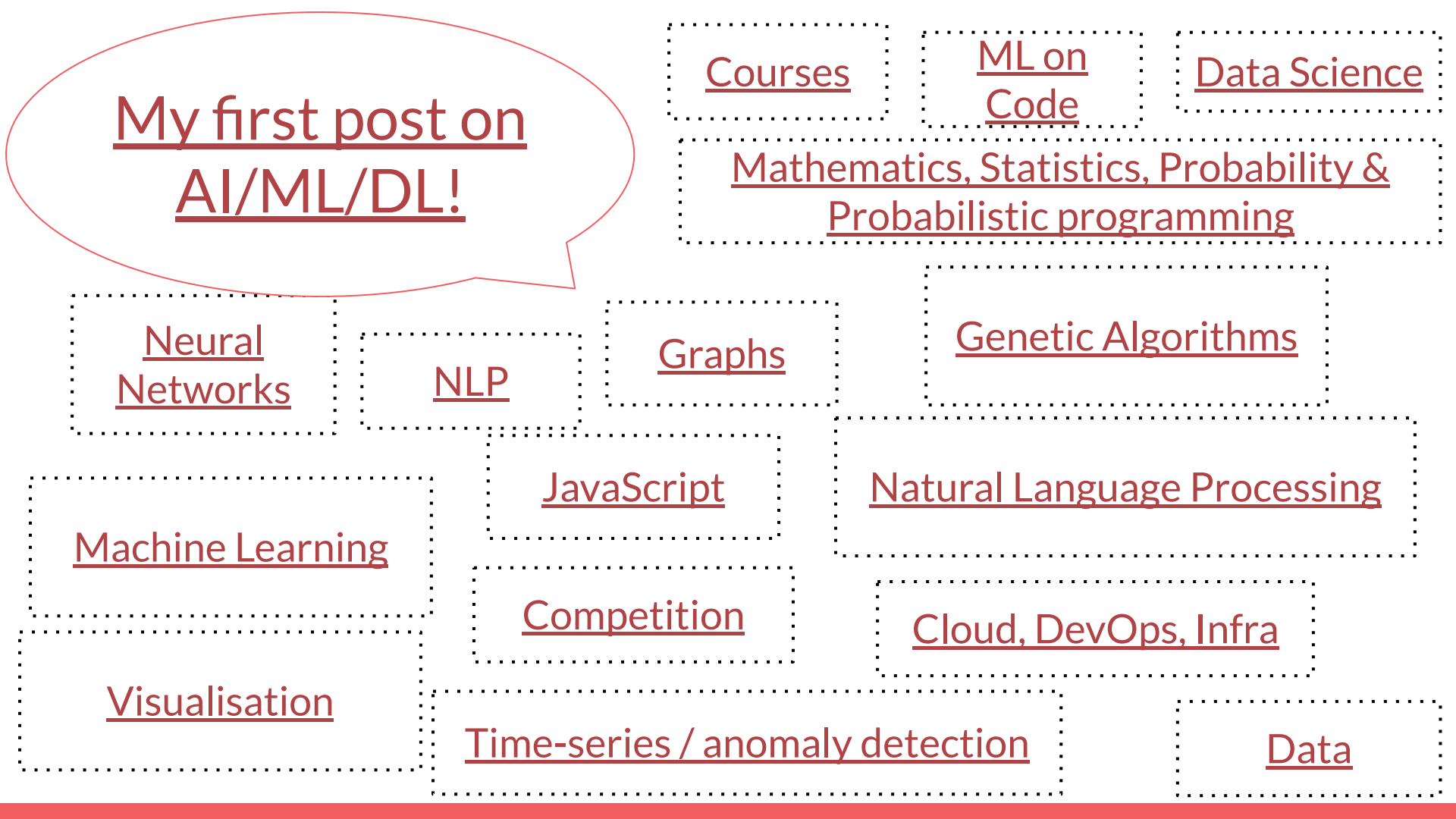
Contributors 15

+ 4 contributors

Languages

Jupyter Notebook 97.8%
Python 2.0% Shell 0.2%

AI/ML/DL highlights


My first post on AI/ML/DL!
Courses
ML on Code
Data Science
Mathematics, Statistics, Probability & Probabilistic programming
Neural Networks
NLP
Graphs
Genetic Algorithms
JavaScript
Natural Language Processing
Machine Learning
Competition
Cloud, DevOps, Infra
Visualisation
Time-series / anomaly detection
Data

AI/ML/DL & Java highlights

Classifiers and Decisions trees

Data Science

Deep Learning

Neural Networks

Tools, libraries & resources

Genetic Algorithms

Machine Learning

Natural Language Processing

JSRs

Clojure resources

Scala resources

Java related projects and technologies

AI/ML/DL artefacts

Cheatsheets

Notebooks

Examples

Virgilio

Presentations

Docker containers

Guides

Study Notes

Things to know list

Learning by example

See Appendix: Learning by example

- DL4J example
- NLP example
- Jupyter Notebook example
- Apache Zeppelin example
- graql-to-english, english-to-graql
- grCuda example
- grPython examples

12+ months worth of coding work

Demo

Running Java inside a Jupyter Notebook

- Blog post
- Github repo nlp-java-jvm-example

Making chatbot have a conversation

- Github repo: https://github.com/neomatrix369/chatbot-conversations
- Questions
  - Is this a real conversation?
  - Would this pass the Turing test?
  - What is good about this example/demo?
- Puzzles to solve
  - What are the different things that can be improved?
  - What new ideas come to mind when you see this?

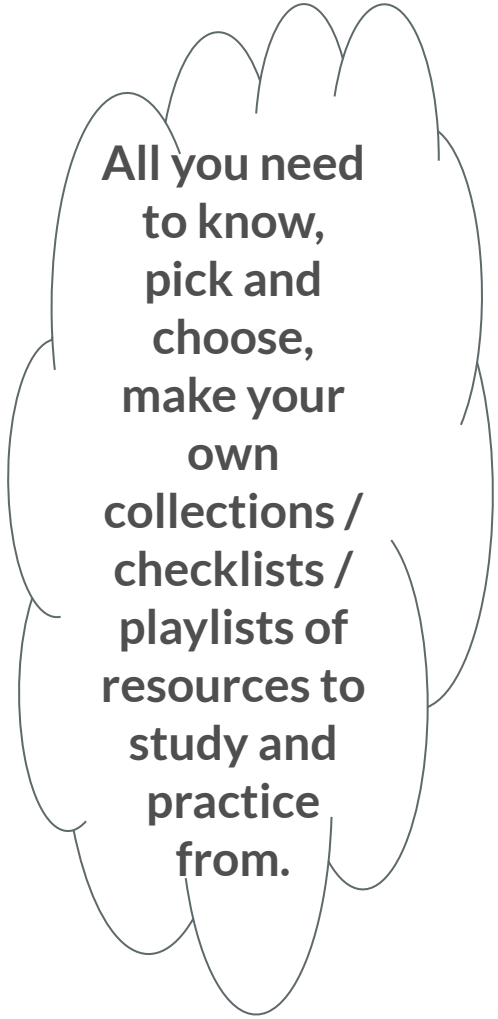
Java is a tool, a programming language.

We can use it to create anything that are created by other programming languages!

Resources

Learning websites

- Awesome AI/ML/DL
- Better NLP
- Awesome Graal
- Virgilio | GitHub
- MadeWithML



More resources

See Appendix section

Summary

- My journey and how I learnt the topics
- We get better results when we organise ourselves
- Create something simple from scratch, similar to our demo and other examples
- A simple idea can go very far, inspire others with creative solutions

Summary

- Java is versatile and can be used to create amazing things just like others are creating
- We didn't cover even more number of Java and AI/ML/DL related topics, libraries framework as the resources shared here hopefully will lead up to them
- No need to feel overwhelmed and let's try to avoid "imposter syndrome"

Summary

The pathway to mastery or championing a subject means we must be different and take different approaches, in addition to the ones that have already been taken!

Thank you!

Martin Toshev and
Martin Patsov!

Questions & feedback

Please share your questions and feedback at

@theNeomatrix369

or on the video stream

Keep in mind...

It’s your turn next to share and inspire!!!

Appendix

Freebies!

Get $500 worth of free cloud credits on Oracle Cloud

People doing some great work with AI & Java

Eyal Wirsansky, Zoran Sevarac, Suyash Joshi, Adam Pocock, Johan Vos and many more...

(please share more names and examples with me so I can add them here)

Accelerated Processing

Java on the GPU by Mitia

Plenty of resources on 

Nvidia's Developer site | Community | Research | Blog site

Parallel processing

GNU Parallel

https://www.gnu.org/software/parallel/

Parallel / Async programming

Talk by Venkat Subramanian

ASYNC

Reactive programming

https://community.oracle.com/docs/DOC-1006738

Java and AI/ML/DL

Machine Learning Best Practices

o

Top 5 machine learning libraries for Java

o

10 Popular Java Machine Learning Tools & Libraries

o

What are machine learning libraries in Java?

Levels of abstractions

- Higher to lower
  - From generic to more specific
- Always know the level of abstract
  - when reading
  - when writing
  - when speaking

Try to analyse this presentation, it's structure, the different levels explored...

See how we can get to know many topics and yet stay high-level, and go deep only when required.

Book references

- How to think like a Scientist!
- Deep work book
- Getting Things Done: Book | Free Resources

Previous talks

- Java N.n: What to know? How to learn?
- Do we know our data as well as our tools?
- Naturally, getting productive, my journey with Grakn and Graql

Create something simple from scratch

- even if it's as simple or silly idea - like the demos i have shown
- *(it can also be putting existing components together from scratch)*
- this can have a massive impact on us and our communities

Learning by example

DL4J example

- Github
- Blog post

NLP examples

- Example 1
  - Github
  - Blog post
- Example 2
  - Blog post
- Example 3
  - Blog post
- Better NLP

Jupyter Notebook example

- Example 1
  - Github
  - Blog: Exploring NLP concepts using Apache OpenNLP inside a Jupyter notebook
- Example 2
  - Blog post
- Example 3
  - Github
  - Blog post

Apache Zeppelin example

- Github
- Blog posts
  - Apache Zeppelin: stairway to notes* haven!
  - Running Apache Zeppelin on Oracle Cloud Infrastructure

graql-to-english, english-to-graql example

- Presentation
- Github

grCuda example

- Blog posts
  - grCUDA: A Polyglot Language Binding for CUDA in GraalVM. NVIDIA Developer Blog, November 2019.
  - grCUDA: A Polyglot Language Binding. Presentation at Oracle CodeOne 2019, September 2019.
  - Simplifying GPU Access. Presentation at NVIDIA GTC 2020, March 2020
  - Optimizing Machine Learning Performance at Netsuite with GraalVM and NVIDIA GPUs
- Github

graalPython examples

- Blog posts
  - Introduction to the Python implementation for GraalVM
  - Moving from Jython to GraalVM
  - Running Python on GraalVM
- Github