Machine Intelligence

Vision, Research & Use cases

Artificial Intelligence

Allow machine to accomplish task human execute with intelligence

Mc Carthy & Minsky

Artificial Intelligence

3 Key Factors

Everything is a connected device producing **Data**

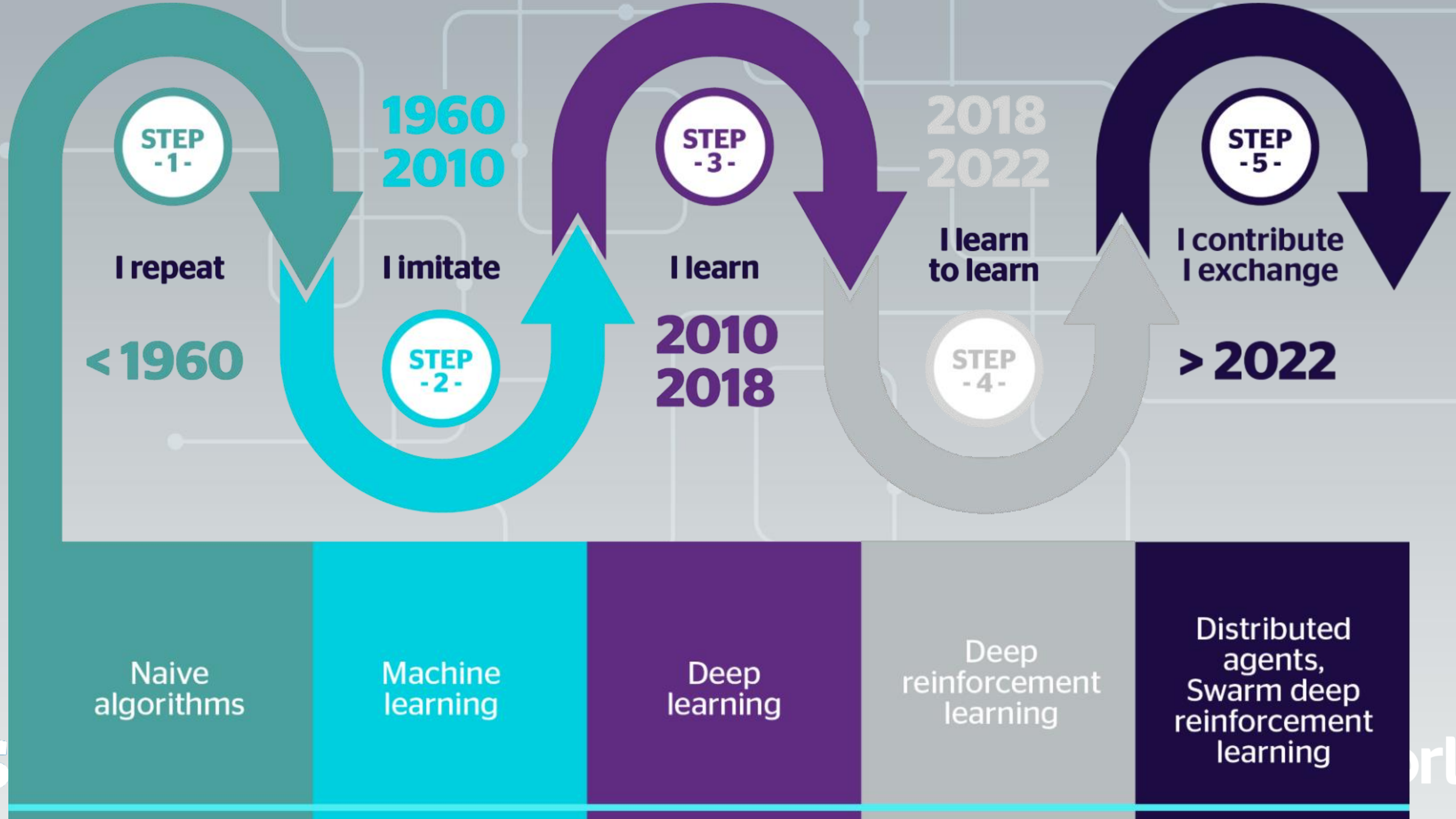
Structured and unstructured data are unified as **Knowledge**

Machine Learning is affordable with today computing power

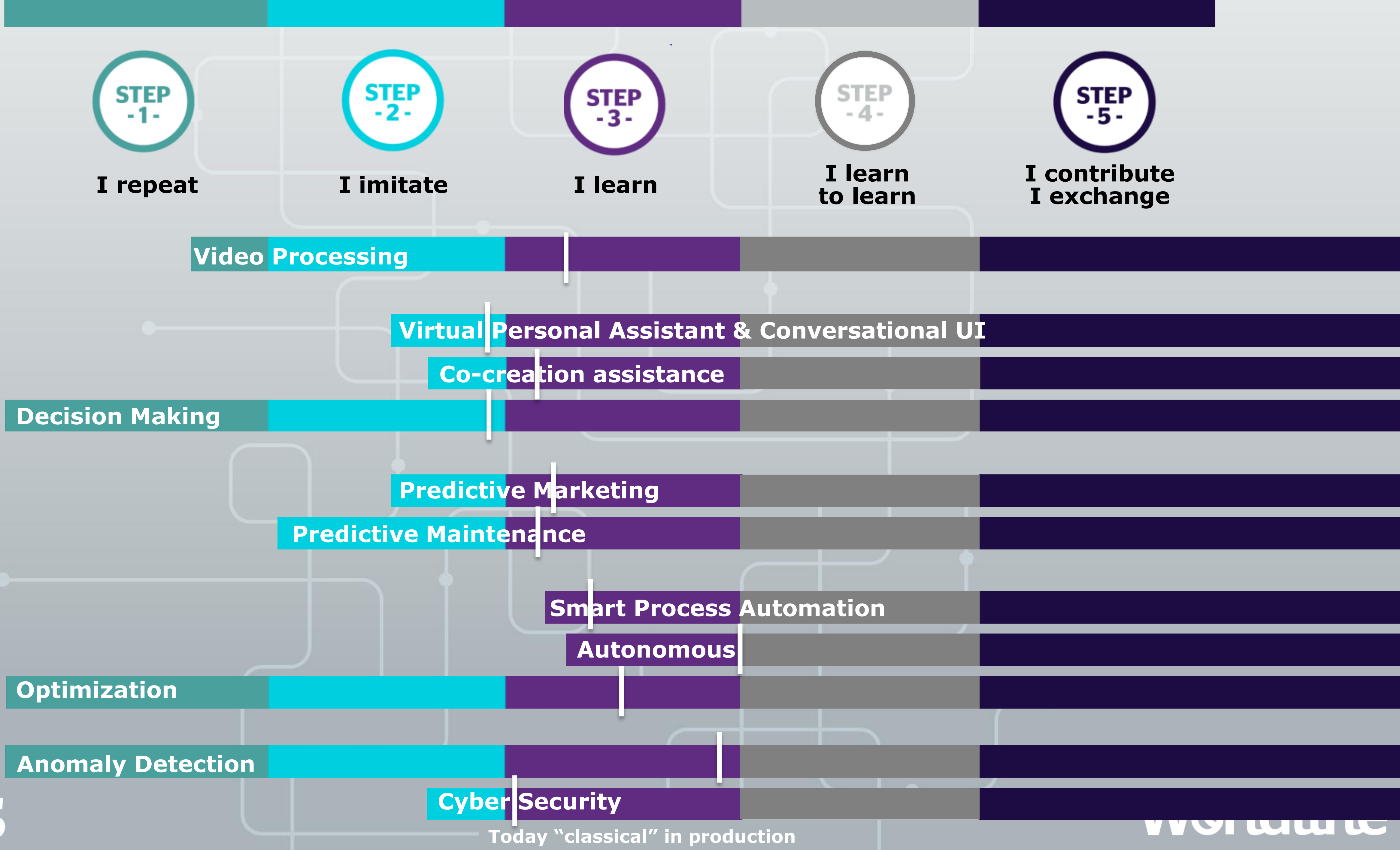
Atos

Worldline

AI & academic history



Business Usages



Get ready for tomorrow

→ Atos / Worldline **Research** programs

Free University of **Brussels**
- Scalable fraud detection techniques in big data context
- Time Series

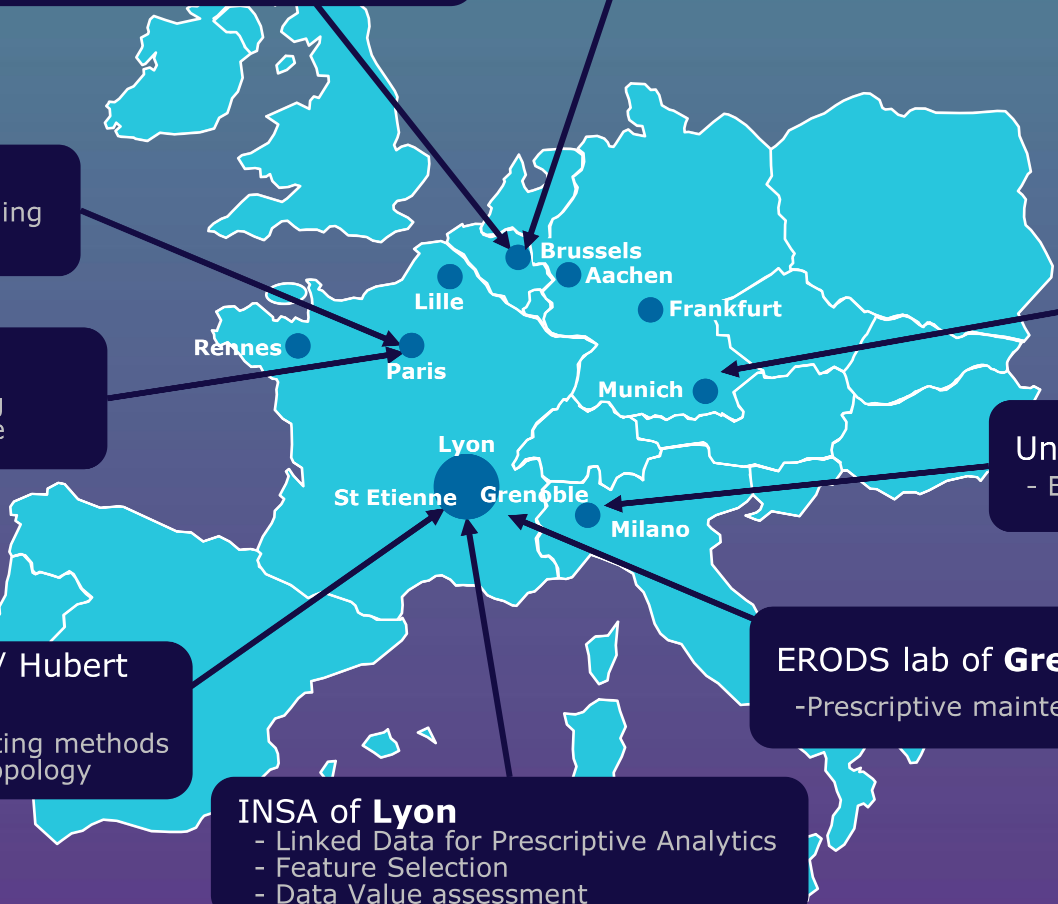
Catholic University of **Louvain**
- Graph mining for fraud detection

ENS – **Saclay**
- Advanced Machine Learning
- Transfer Learning

CEA – **Saclay**
- Advanced Machine Learning
- High performance hardware

University of **Passau**
- Data Value Assessment
- Certificate Program
- Deep Learning on Data Streams

University of **Milano**
- Expert rules ranking methods



Jean Monnet University / Hubert Curien - St Etienne
- Online learning and boosting methods
- Massive Deep Learning topology

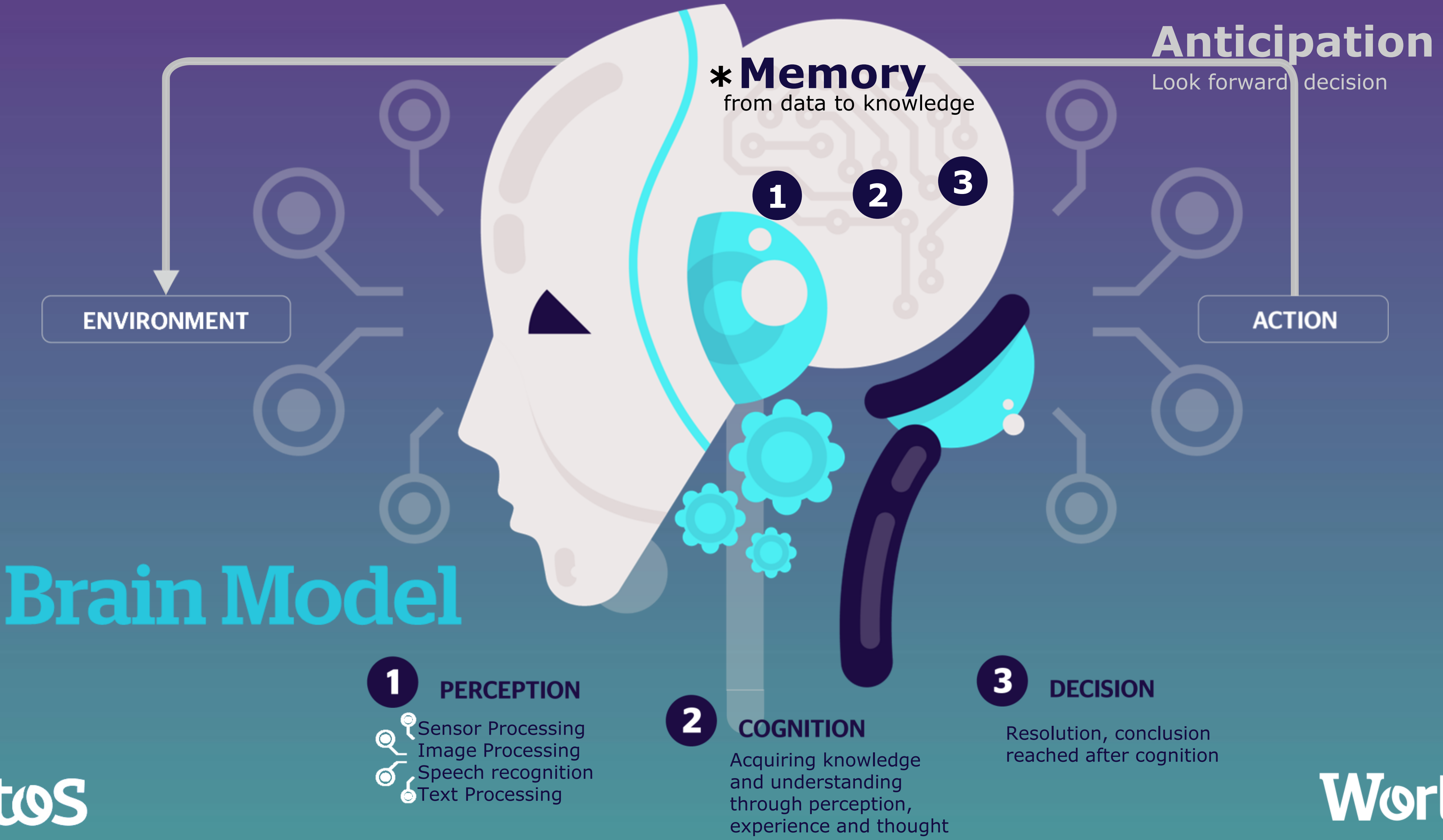
ERODS lab of **Grenoble** (LIG)
- Prescriptive maintenance for exascale HPC

INSA of **Lyon**
- Linked Data for Prescriptive Analytics
- Feature Selection
- Data Value assessment

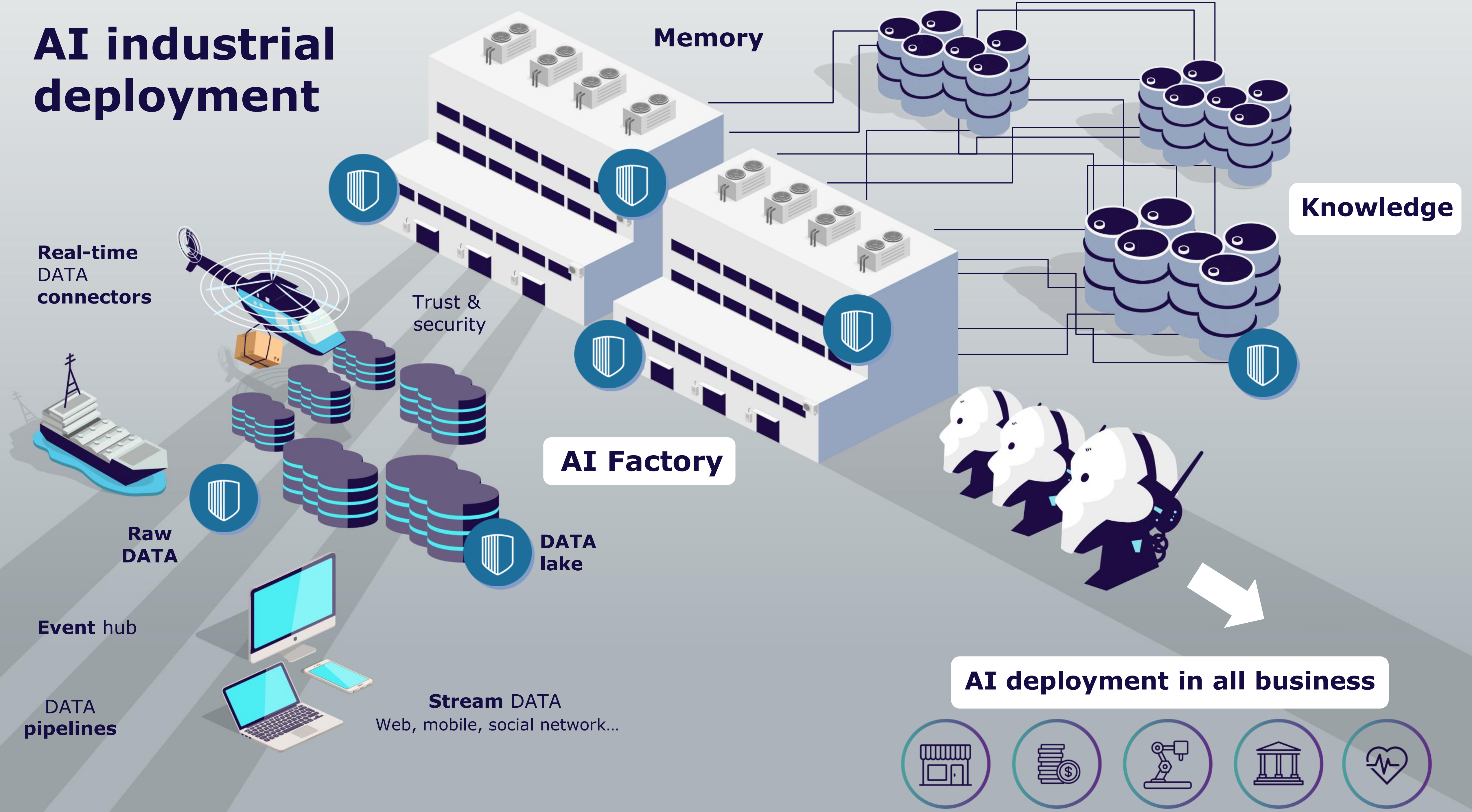
Atos

Worldline

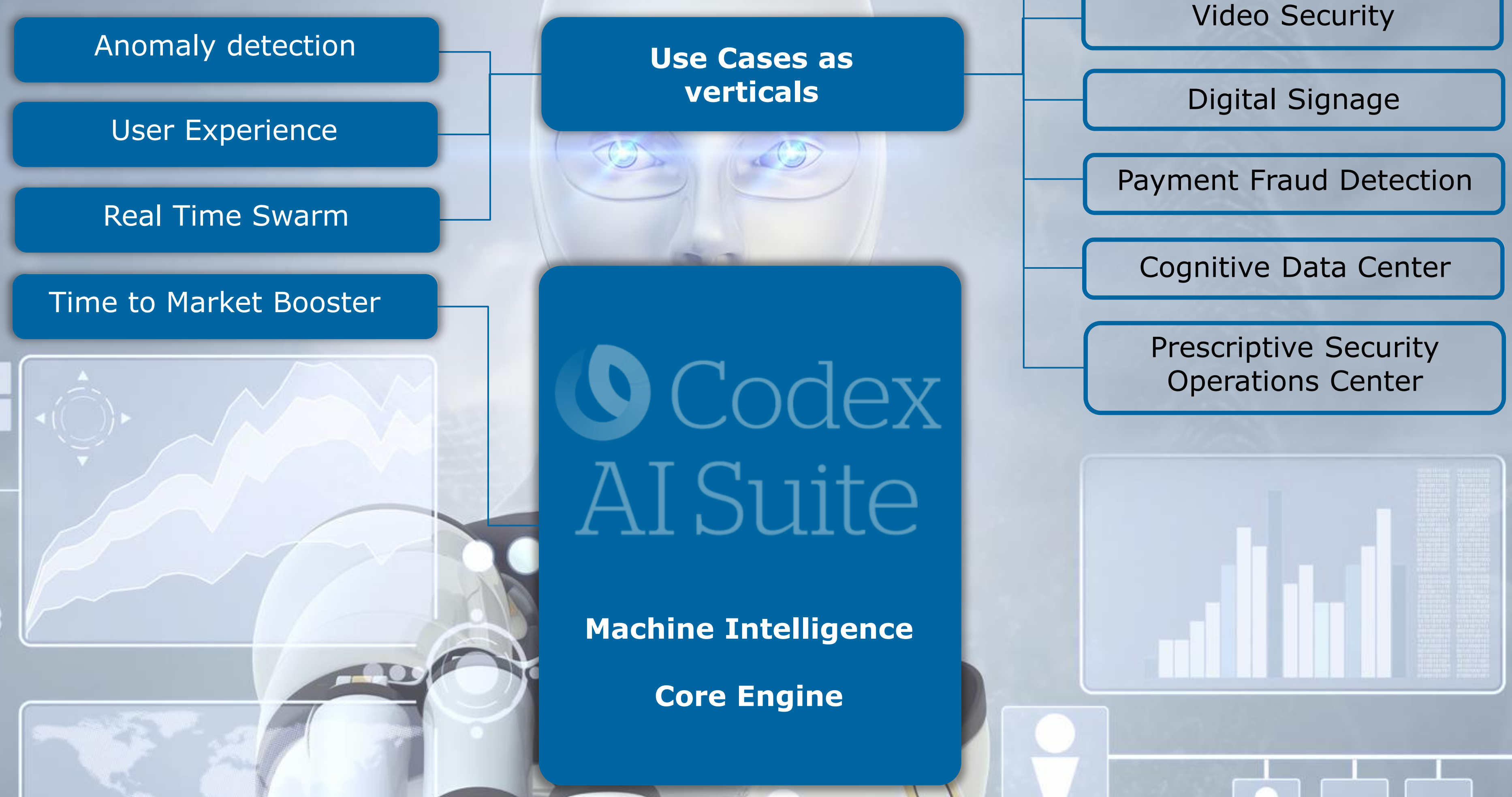
From Artificial Intelligence to Machine Intelligence



AI industrial deployment



Make it happen today



Atos

Worldline

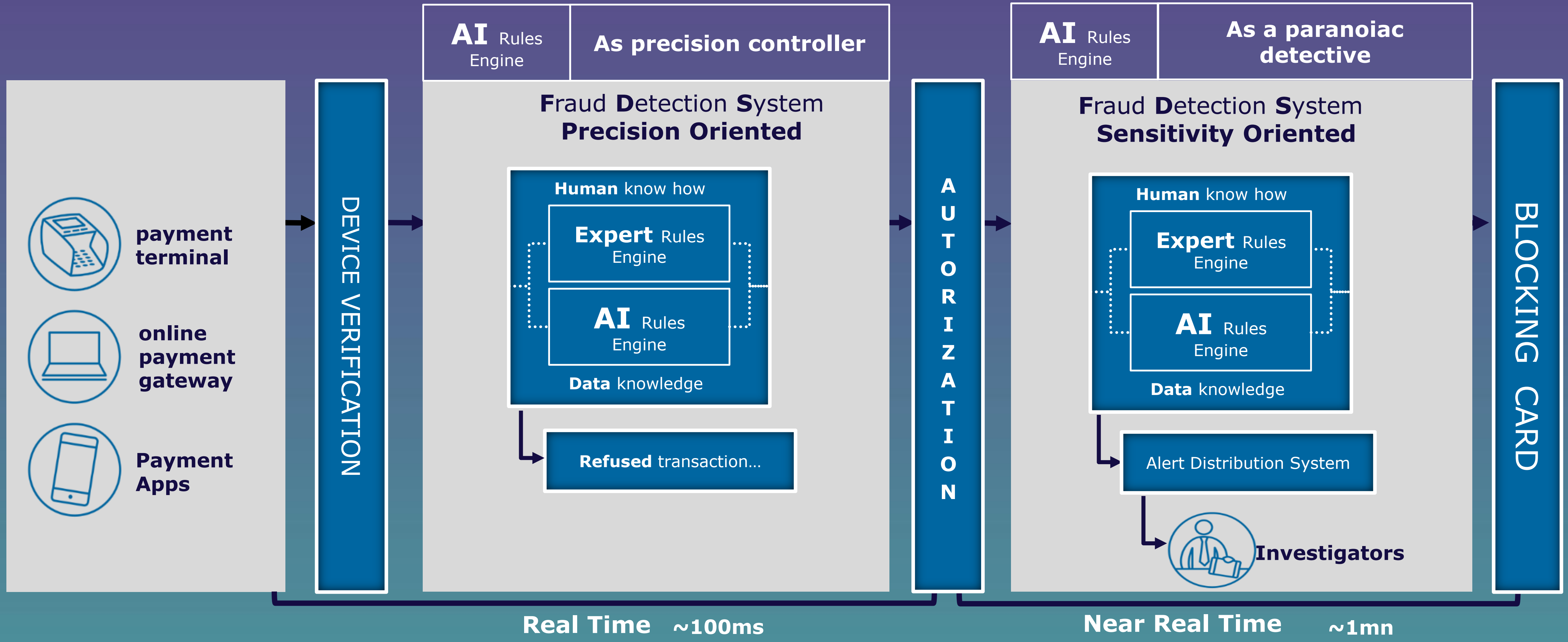
Payment Fraud Detection

Fraud can be done on
the **issuing** side or on the **acquiring** side

The 4-corners model:

- **Cardholder:** the customer (you, me, any individual)
- **Issuer:** The bank of the cardholder
- **Acquirer:** The merchant bank
- **Merchant:** selling a product

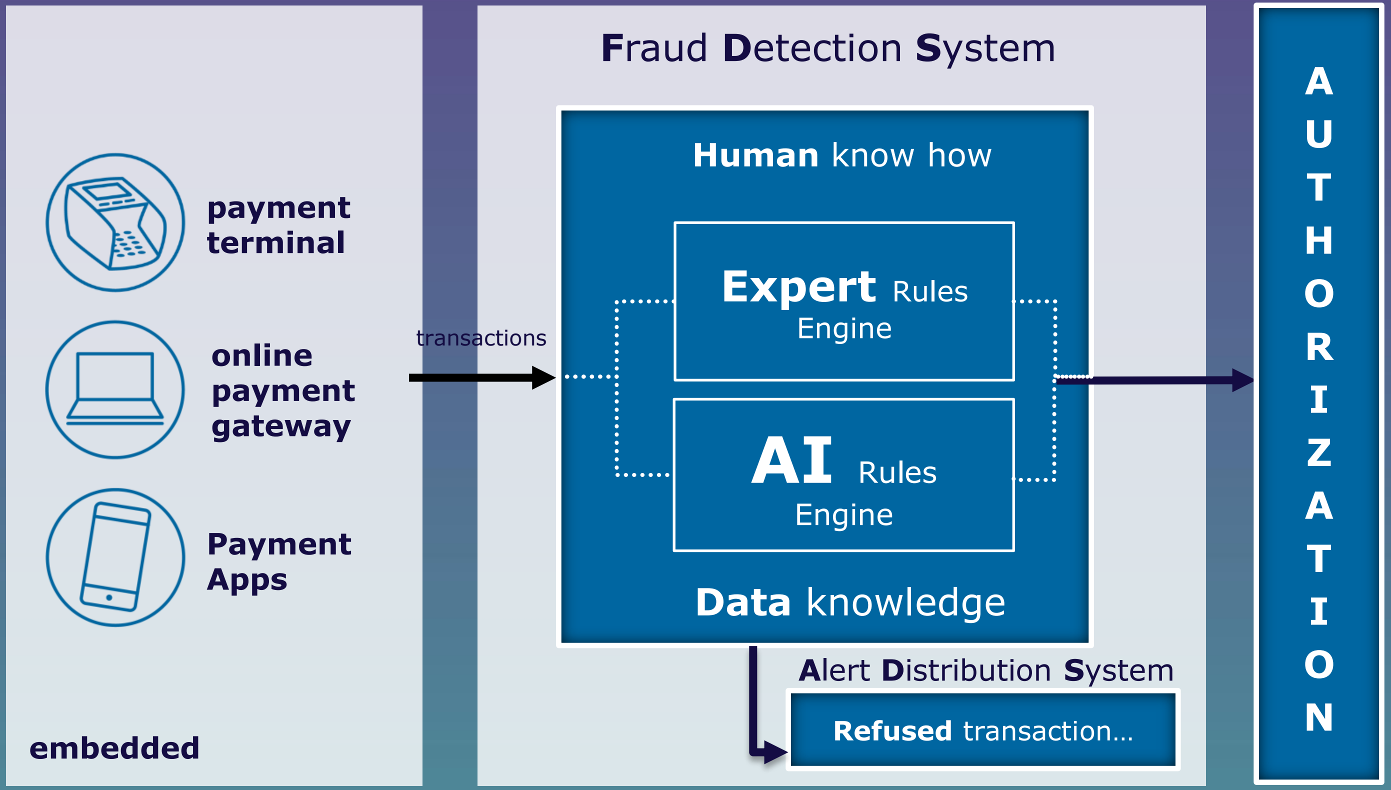
Fraud Detection: a triple blades system
Applied to payment



Atos

Worldline

Payment Fraud Detection

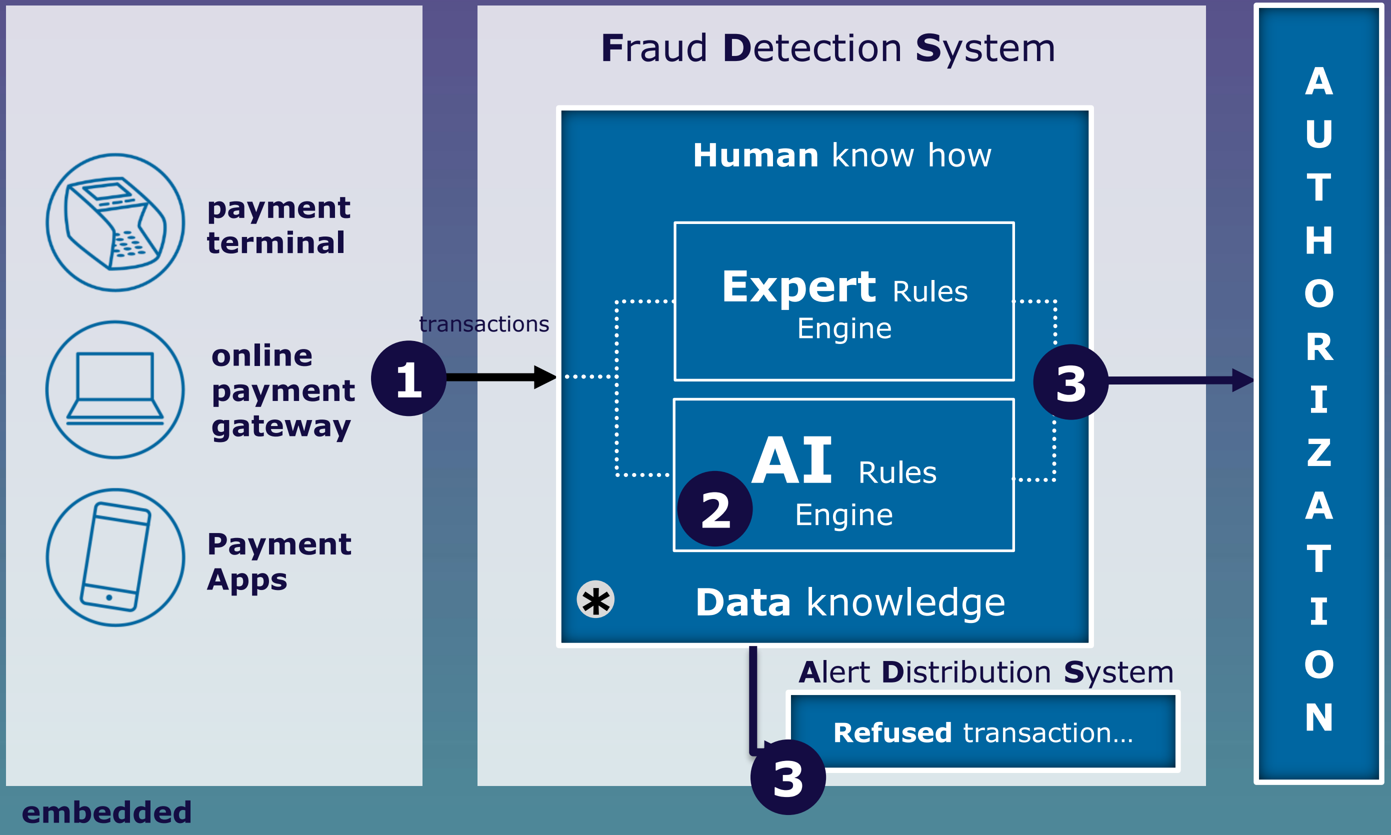


Machine Learning challenges

- Continuous **streams** of transactions
- **High volume** of data and **fast react**
- Highly **unbalanced** distribution
- Change in **pattern** for fraud
- **Overlapping** classes
- Rules management **automation**

Atos Worldline

Payment Fraud Detection



from Machine Learning to

→ **Machine Intelligence**

Brain Model

1. **PERCEPTION**
2. **COGNITION**
3. **DECISION & ACTION**

* **MEMORY**

Atos

Worldline

Fraud Detection as a machine learning challenge

Class imbalance

Class overlapping

Fast detection

Input → Objective function optimization (Feature engineering, Classifier) → Output

- **Hand made designed features,** e.g. aggregations
- **Representation learning,** e.g. Deep Learning, auto encodeur, RBF

- **Random Forest**
- **SVM**
- **Boosting**
- **Neural Network**
- **LSTM**
- **...**

Concept drift

Delay Feedback

Fraud Detection: Machine Learning challenge

Data Description

Transaction ID	Card ID	TERM COUNTRY	Amount (euros)	TX DATETIME	MCC		Fraud
1	1010	'FRA'	28	2018-04-04 12:00:00	5992 (Florists)	...	0
2	1234	'BEL'	5	2018-04-05 13:20:05	6011 (ATM)	...	1
3	3456	'DEU'	12	2018-04-05 13:30:19	5814 (Fast Food)	...	0
4	1234	'BEL'	500	2018-04-05 23:50:00	6011 (ATM)	...	1
5	23	'FRA'	1200	2018-04-06 00:05:00	3007 (AIR FRANCE)	...	0

In total, more than **40 variables** for each transaction.

	Categorical	Numerical	Other
Transaction	Currency, is Ecom, …	Amount	Datetime
Terminal	Country, activity sector, …		ID
Card	Type, brand, …	Credit Limit	PAN ID
Card Holder	Broker, gender, …	Age	

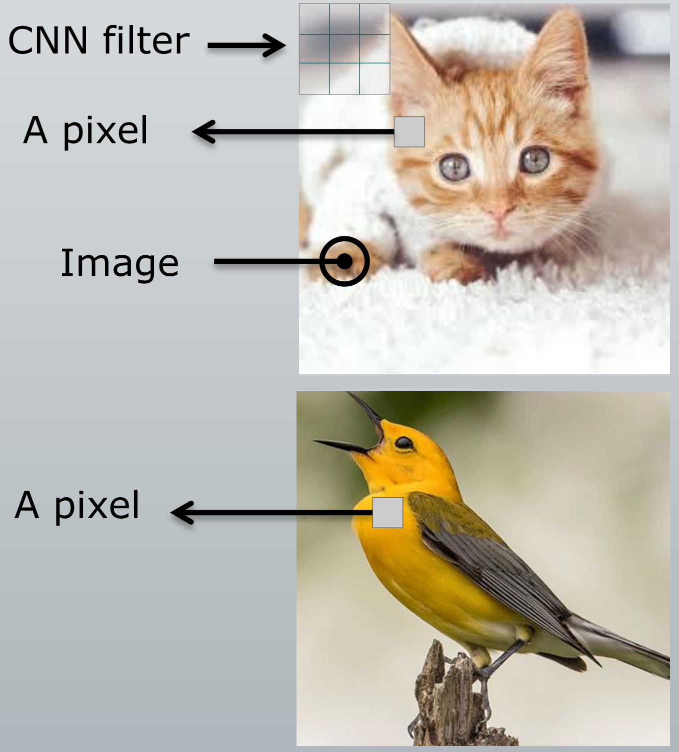
Atos

Worldline

Fraud Detection: Machine Learning challenge

Comparison with CNN applied to image

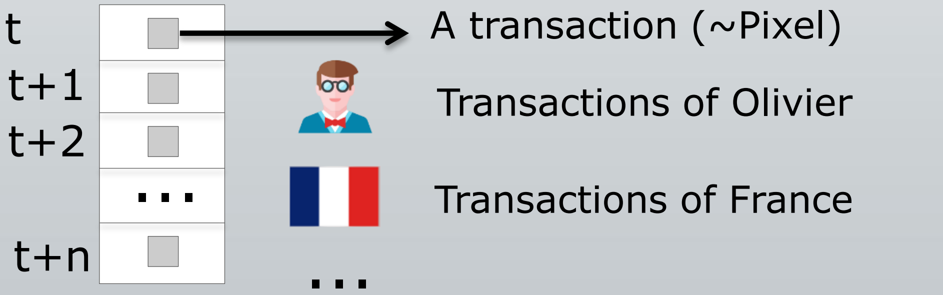
An image = vector of pixels

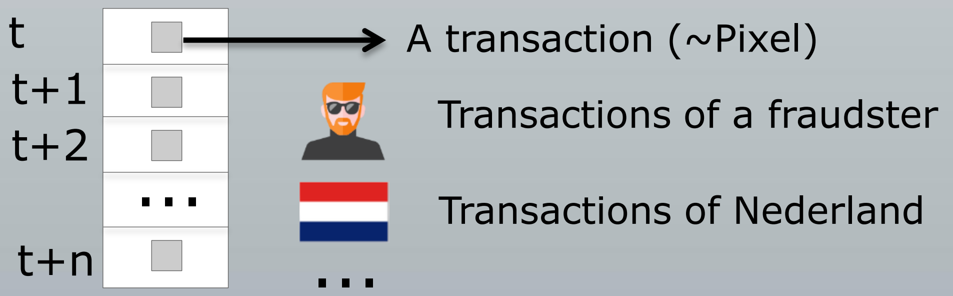


- 2 dimensions (color, space)
- Spatially correlated
- Continuous values
- Constant resolution

VS

A pool of transactions = vector of vectors





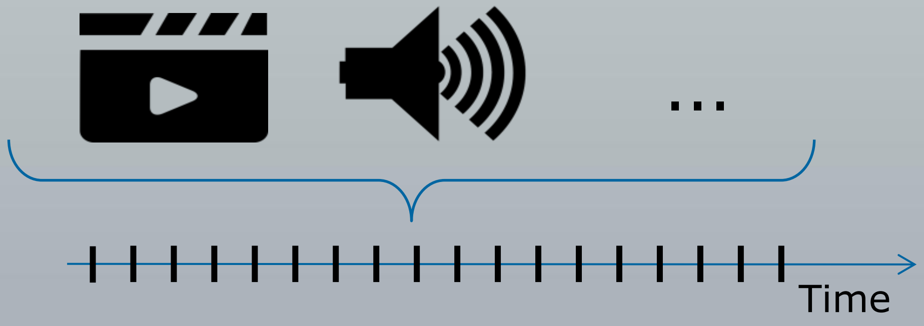
- More than 40 dimensions
- Temporal correlated
- 90% discrete values
- Irregular time gap

Atos

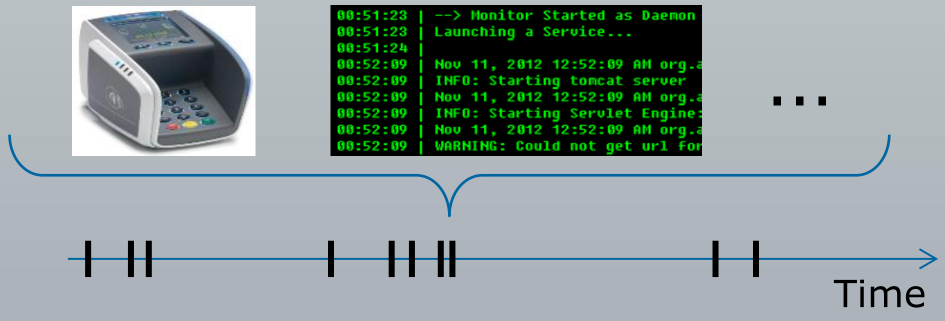
Worldline

Fraud Detection: Machine Learning challenge
Comparison with LSTM

	Videos	Sound	Sensors/ Servers	Payments
Example of application	Security, Tracking	Speech recognition	Predictive maintenance	Fraud detection
Sequence of	Images	Noises	Logs	Transactions
Time-lapse between two events	Fixed	Fixed	Fixed Variable	Variable
Key feature	Time (frequency)	Time (frequency)	Sensor ID, PID	Cardholder



The pace is given by the key feature



Multiple ways to define a sequence according to key features.

Fraud Detection: Machine Learning challenge

Comparison with Word2Vec and Text recognition

Sequences in sentence

Atos **Wins** **Multi-Year** **Contract** **For 120 Million Dollars**

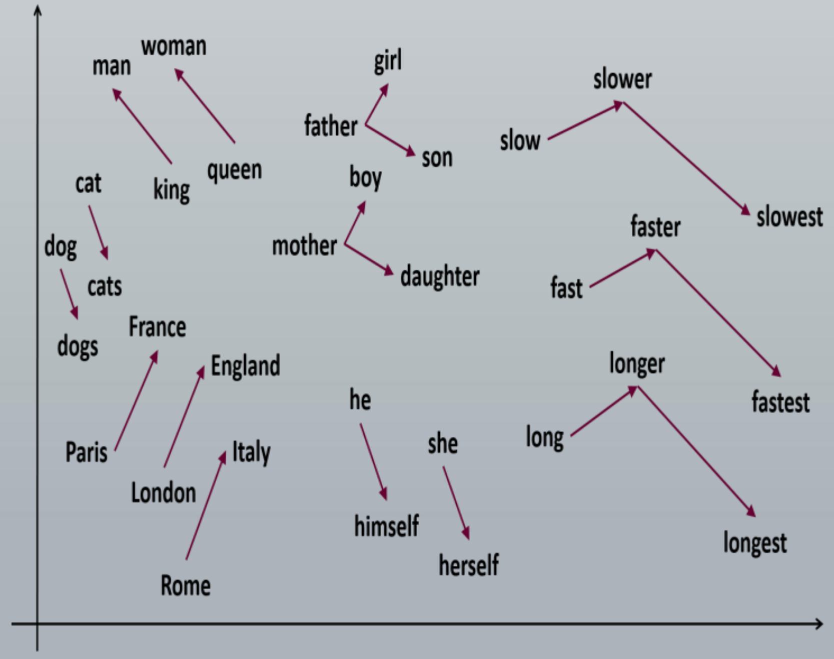
cardID	Country
1010	FRA
25	BEL
1010	BEL
34	USA
25	FRA
1010	JPN

cardID	Country
1010	FRA
1010	BEL
1010	JPN
25	BEL
25	FRA
34	USA

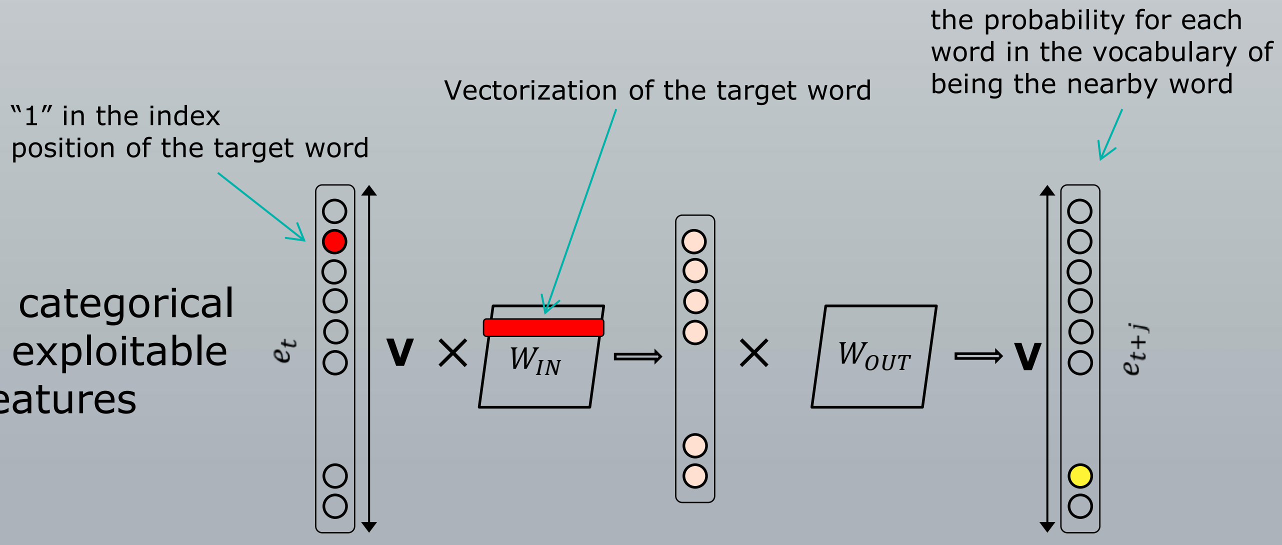
1010: " FRA BEL JPN"

25: "BEL FRA"

34: "USA"



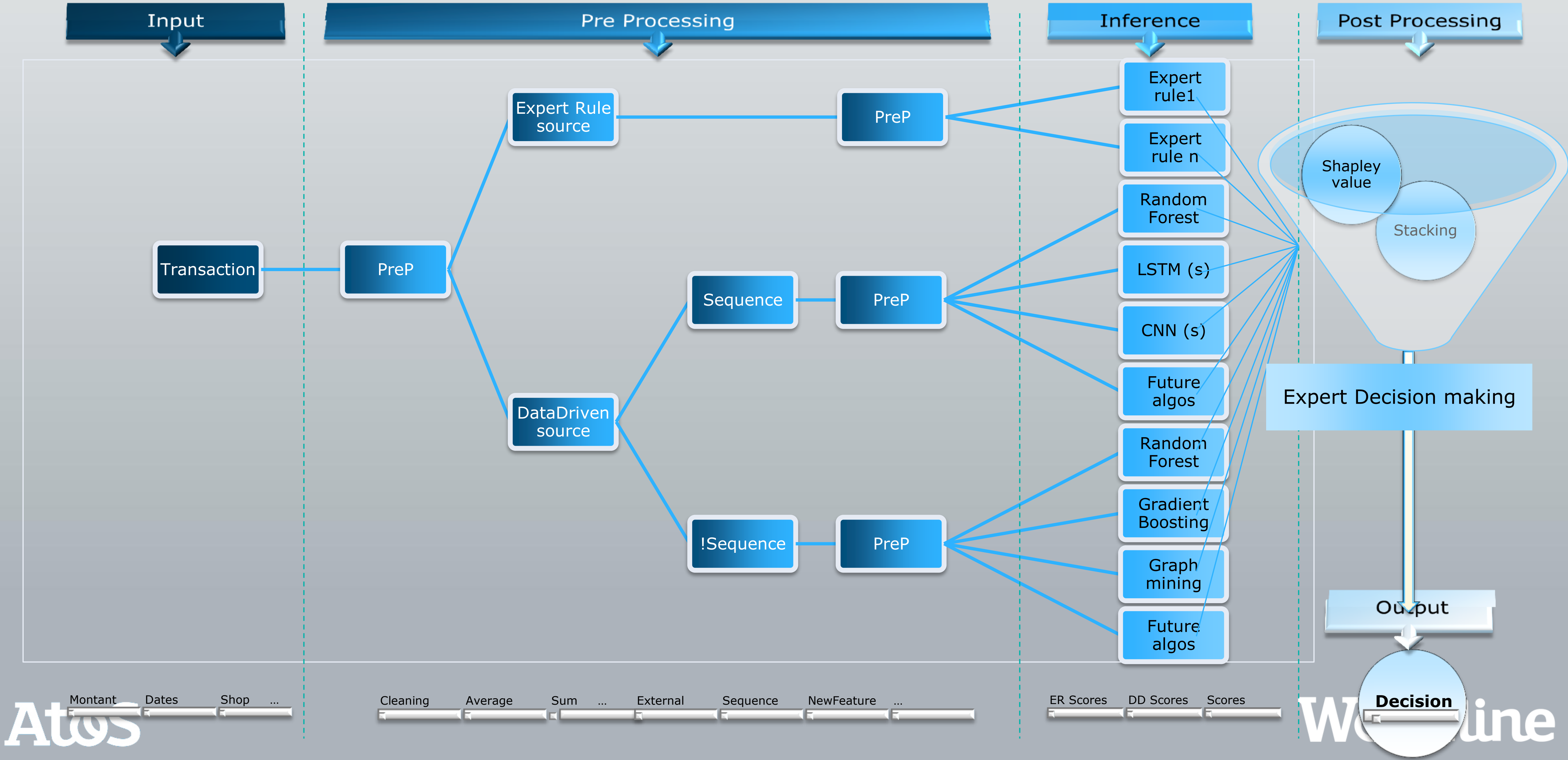
Transforming categorical features into exploitable vector features



One hot encoding of target

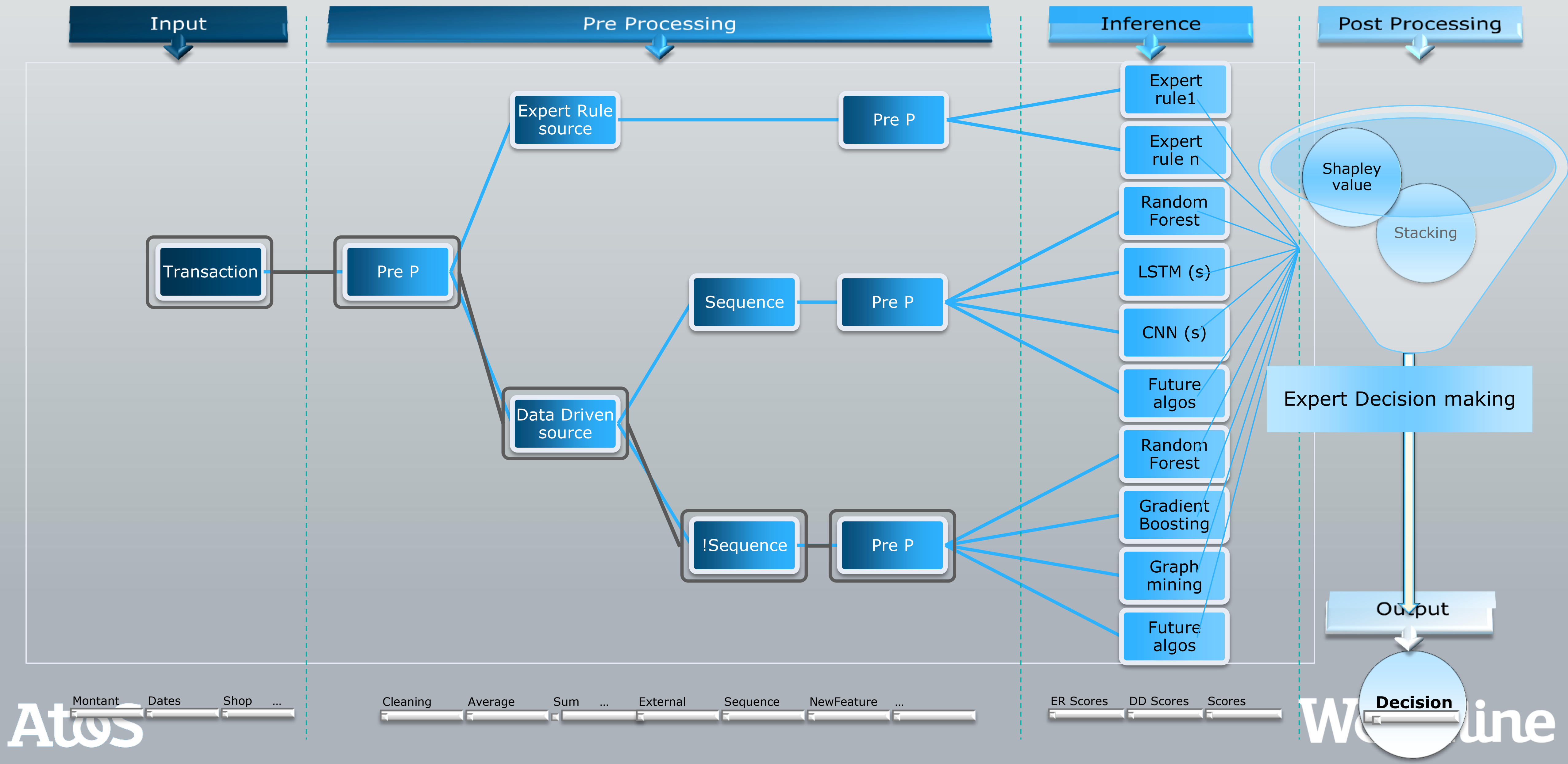
One hot encoding nearby word

Fraud Detection: inference pipeline

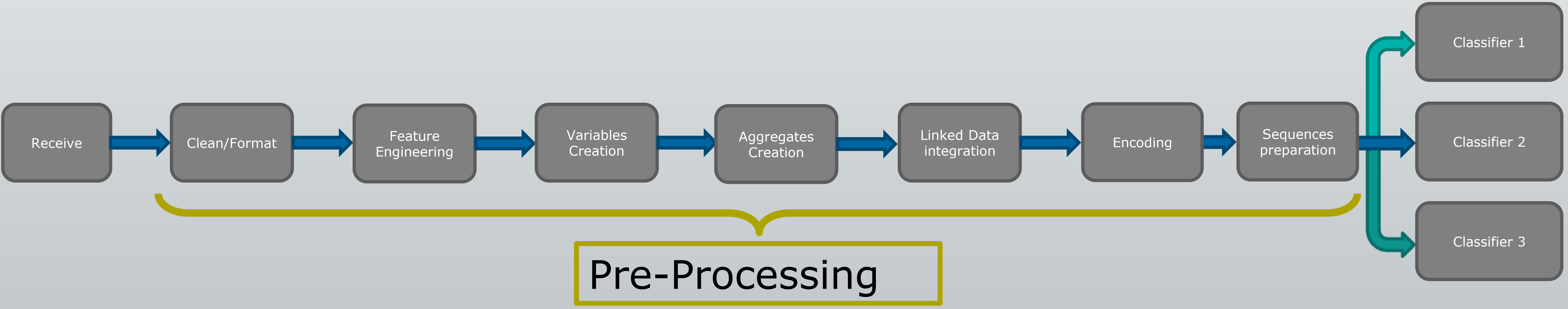


Atos Worldline

Fraud Detection: inference pipeline



Fraud Detection: training pipeline (simplified)

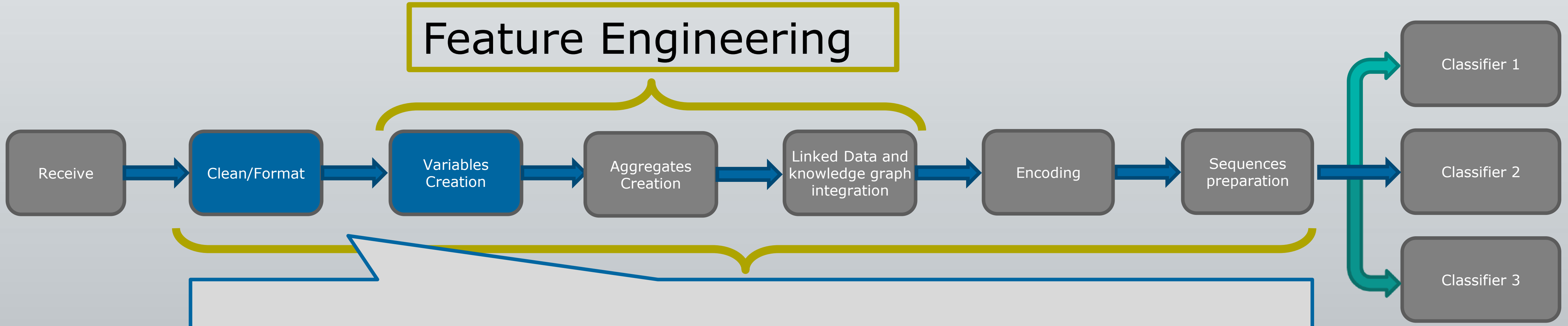


- A “**classical**” data preparation
- A **complexity** because of the **nature** of the problem to tackle
- A **difficulty** because of the **volume** and the **combinations**

Atos

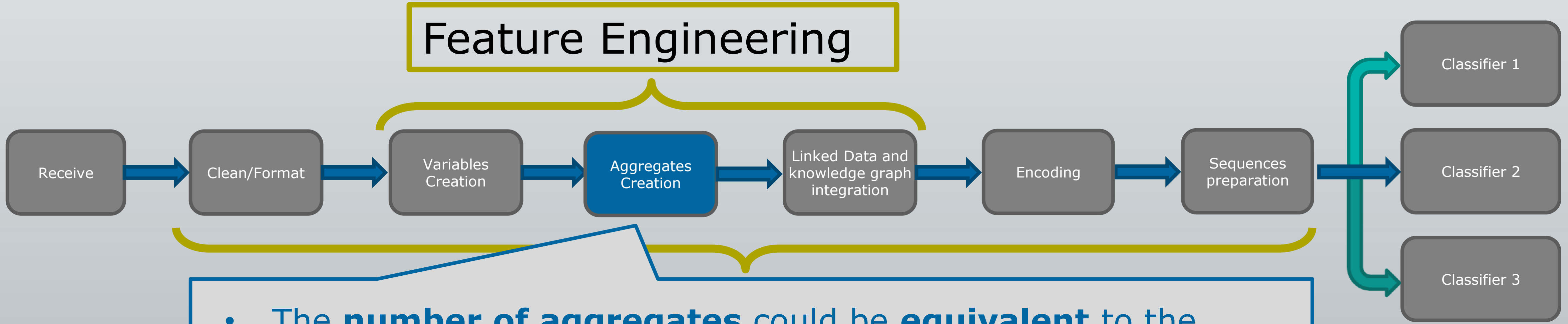
Worldline

Fraud Detection: training pipeline (simplified)



- Mostly a classical data analytics task but on **billion of lines** with more than **40 variables**
- Should **be efficient** enough to be execute during the **inference phase**

Fraud Detection: training pipeline (simplified)

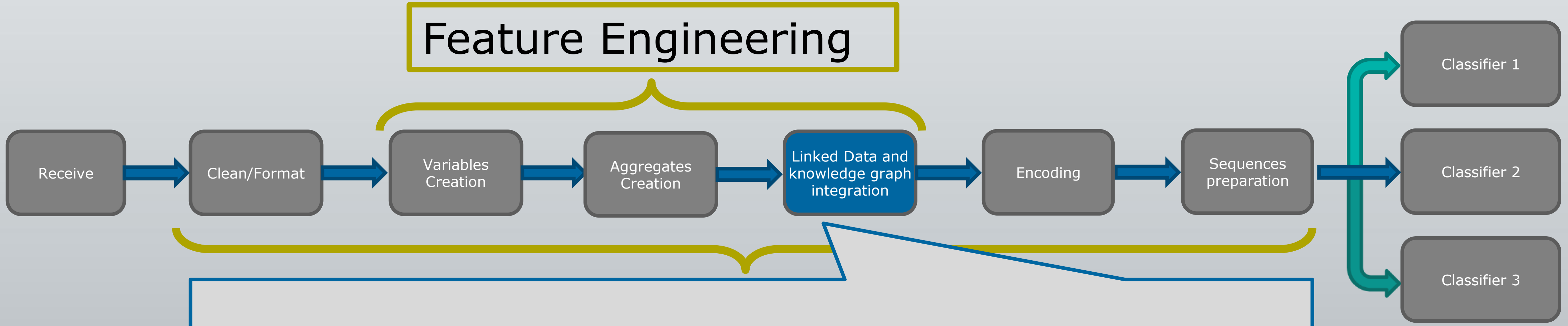


- The **number of aggregates** could be **equivalent** to the **number** of useful **variables**
- **Aggregates** are computed on the **overall historical** and cost a lot in computing
- Some aggregates based on **Hidden Markov Model** are using a **huge amount of computing**
- Some aggregates based on NN (**LSTM**, **CNN**) cost **more** than **HMM**

All of them should be **robust** to the potential **concept drifts**

Atos Worldline

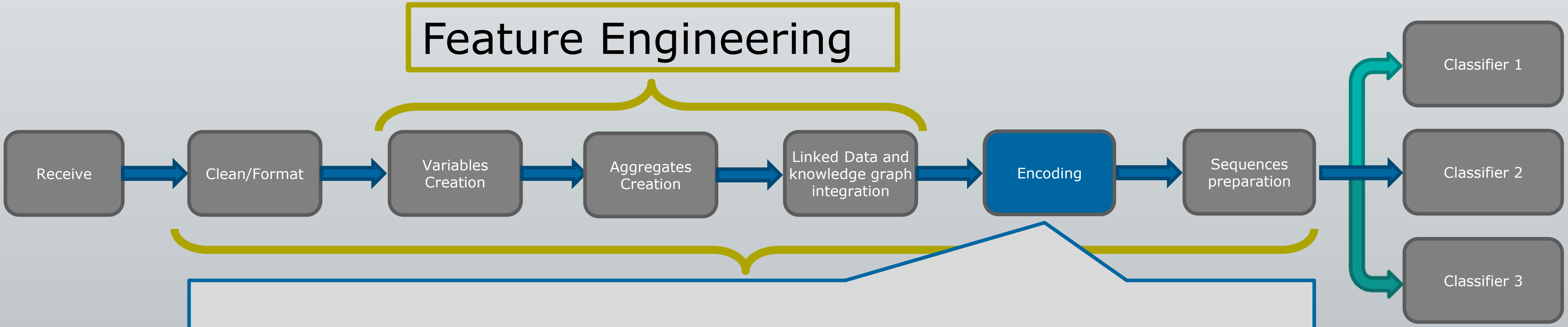
Fraud Detection: training pipeline (simplified)



New specific **variables** could be introduced **as vector**

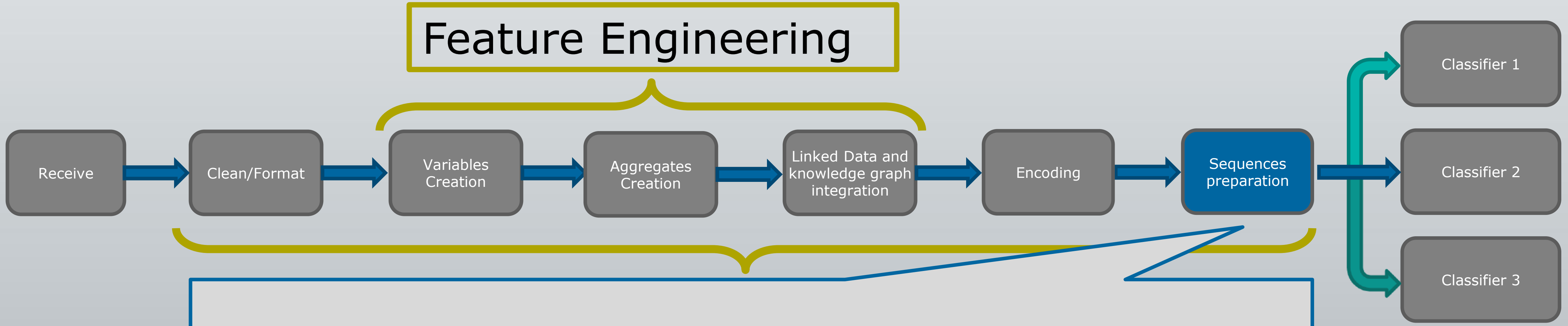
- **Embedding's** based on word2vec technics
- **Knowledge graph** based on the **policy**
- **Linked Data** extended variables

Fraud Detection: training pipeline (simplified)



- **Encoding** issues linked to the **classifier sensitivity**
- **Scaling** issue because of the number of **values** per **variable**
- Should be **robust** enough to be **compute** during the **inference phase** in few **milliseconds**

Fraud Detection: training pipeline (simplified)



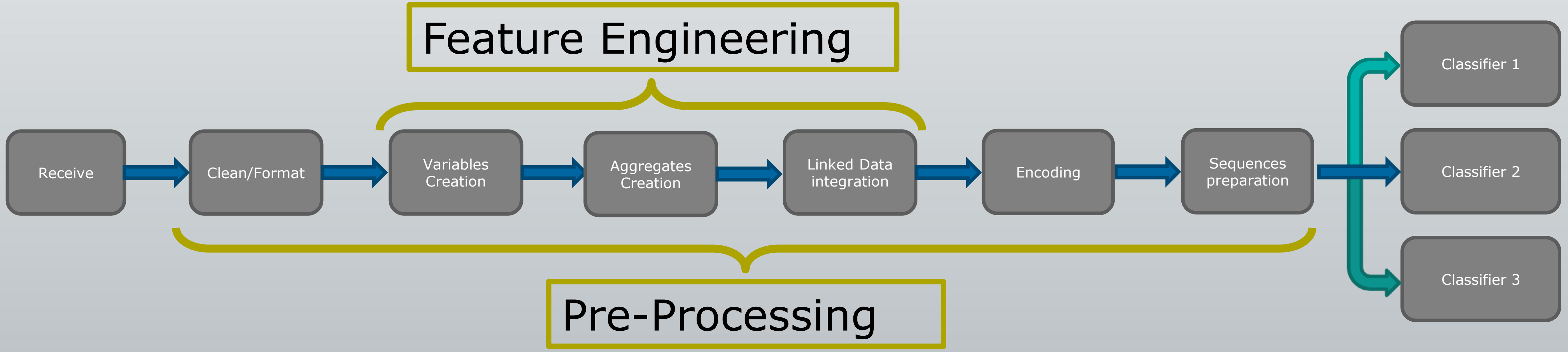
- **Fraud Detection** could be assimilate to weak signal detection, so the time is an important component
- The **creation** of **sequences** of vector means to access quickly to all the **precedent** related **vector**
- Because you have to take care of case without long sequences you have to **double** the **number** of **pipelines**

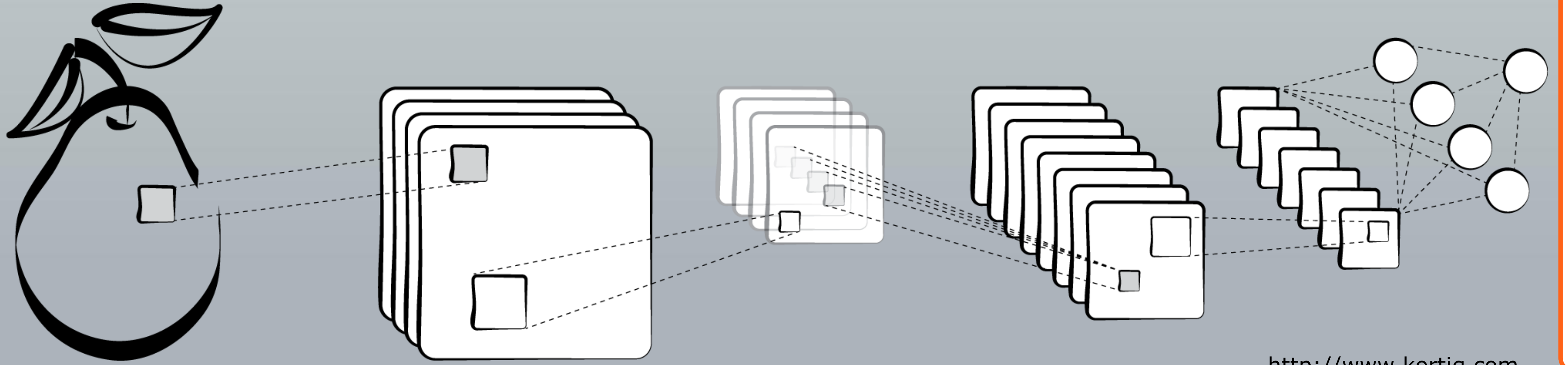
Atos

Worldline

Fraud Detection: training pipeline (simplified)

Comparison to Deep learning for images





http://www.kortiq.com

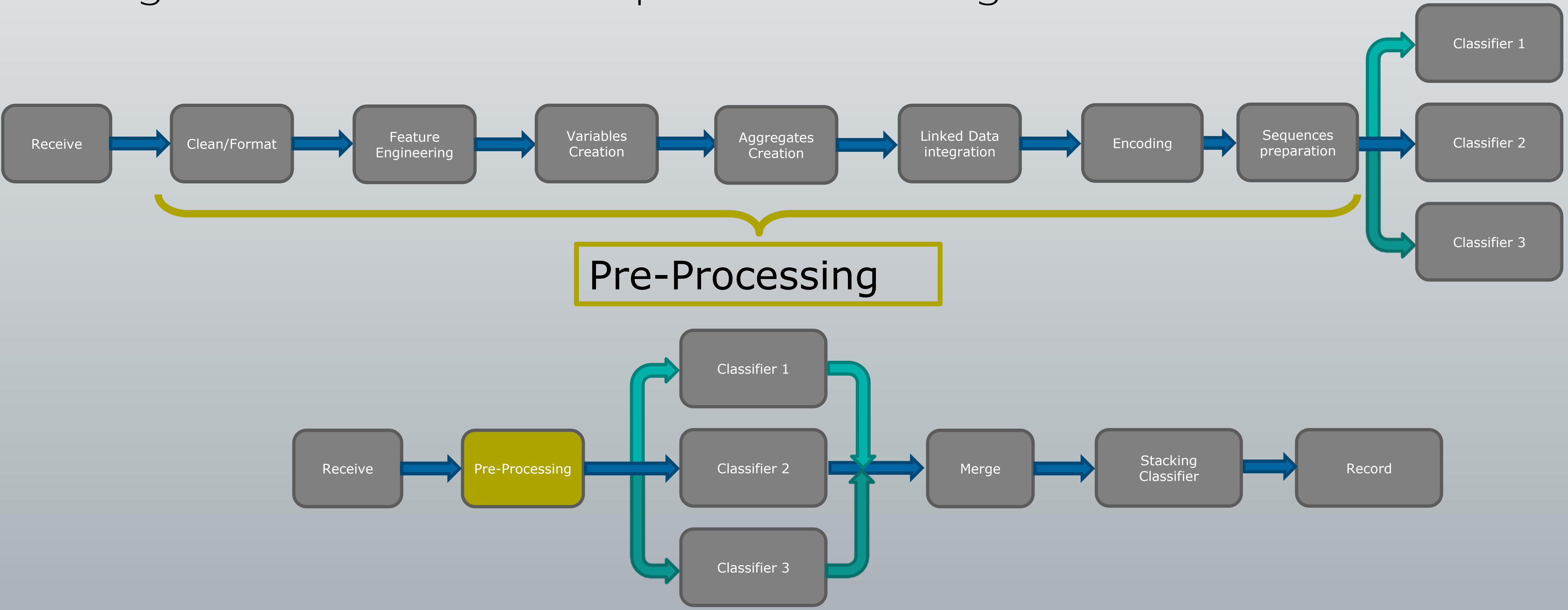
N^{N^N} time the complexity of a CNN for images

Atos

Worldline

Fraud Detection:

Training vs Inference and consequences on training



In less than 100ms per transaction on N pipelines!

Fraud Detection:

How to make it happen ? Ingredient of success

People

- You need **Internal Scientist** able to **discuss** with **Academics**
- You need **Research Engineers** able to **translate** paper into **code**
- You need **System Engineer** to maintain the **infrastructure platform**

Infrastructure

- You need **CPU, GPU, RAM and Storage**
- Agility to install **any kind** of **software**

~ 30 Tflops / User

Time

Months of computing & coding