eleven | ENPC – Département IMI
Digital Innovation at eleven

To the attention of IMI Department

February 13th, 2024

École des Ponts
ParisTech

eleven
strategy • data • digital

AGENDA

1. About Eleven
2. Digital Innovation projects
3. Competitive market intelligence
4. Natural Language Processing for IT ticket resolution
5. Computer Vision for satellite images quality inspection
6. Closing remarks

TODAY'S SPEAKERS

Emma
Consultant,
Data Scientist

Louis
Senior Consultant,
Data Scientist

Nicolas
Consultant,
Data Scientist

Proposed schedule for the week: the timing may be short, do not hesitate to split the work between the members of the team

	Tuesday 13th	Wednesday 14st
AM session	Presentation of eleven (*9:00 - 9:15*) Presentation of eleven's assignments *(9:15 - 10:15)*	Online Q&A (*9: 00 - 10: 00*)
	Refresher on Supervised Learning (*10:30 - 11:30*)	Working Time (*10: 00 - 12 : 00*)
	Hackaton Presentation and Kick-Off (*11:30 - 12:00*)	
PM session	Working Time (*13: 00 - 14 : 00*)	Working Time (*13: 00 - 16 : 00*)
	Q&A, Set-Up verifications (*14: 00 - 15 : 00*)	Jury (*16:00 - 17: 00*)
	Working Time (*15: 00 - 18 : 00*)	Closing Ceremony (*17: 00 - 17: 30*)

AGENDA

1. About Eleven
2. Digital Innovation projects
3. Competitive market intelligence
4. Natural Language Processing for IT ticket resolution
5. Computer Vision for satellite images quality inspection
6. Closing remarks

4 COMPLEMENTARY OFFERS

DATA AND AI

- Are data science and A.I. relevant to my business?
- How can I implement high value A.I. projects?

DIGITAL STRATEGY

- How does digital disrupt my industry and business?
- To what extent am I prepared to take advantage of this?
- What moves should I make to thrive?

PRIVATE EQUITY

- Is my digital-enabled target attractive?
- How can I drive digital-enabled value from my asset?
- What equity story can I tell?
- How best to position my asset for exit?

IMPACT & CSR

- How can I quantify key environmental business metrics?
- What levers can I activate to improve my CSR performance?
- How should I prioritize environmental actions?
- How can I build a natively sustainable business model?

THE VINCI GROUP OPERATE PRINCIPALLY AROUND TWO ACTIVITIES: CONCESSIONS AND CONTRACTING

Overview of VINCI group's main activities

VINCI
3500+ small to intermediate business units around the globe dealing with various activities

CONCESSIONS		CONCTRACTING			
VINCI CONCESSIONS	VINCI AUTOROUTES	VINCI CONSTRUCTION	EUROVIA VINCI	VINCI ENERGIES	VINCI IMMOBILIER
Specialized in the **infrastructure concession**	Specialized in the **French highways concession, stadium, etc.**	Specialized in **construction** and **engineering**	Specialized in road **conception, construction and maintenance**	Specialized in **building solutions, Infrastructure, Industry, ICT**	Specialized in **real estate**
~20,000 FTEs	~6,168 FTEs	~71,400 FTEs	~43,640 FTEs	~77,300 FTEs	~850 FTEs
~€1,6Billions	~€5,365billions	~€14,2Billions	~€8,9billions	~€12,6Billions	~€1,273billions
46 Airports		~30k projects	~450 BU	~1800 companies	
Subsidiaries	**Subsidiaries**	**Subsidiaries**	**Subsidiaries**	**Subsidiaries**	**Subsidiaries**
VINCI Airports VINCI Railways VINCI Highways	ASF, Cofiroute, Escota, Arcour, ...	VC France VC Grand Projet Freyssinet, etc.	Eurovia CS ...	VINCI Facilities, Omexom, Citeos, Actemium, Axians	

LEONARD REPRESENTS A VINCI TRANSVERSAL UNIT AIMING TO FAVOR INNOVATION, FACILITATE COLLABORATION AND LEAD A STRATEGIC FORESIGHT

Overview of LEONARD position and main activities

Foresight

- Identify long term trends impacting VINCI activities
- Anticipate potential disruptions

Community

- Communicate inside and outside the group
- Spread messages supporting group vision and strategy
- Facilitate interactions inside and outside the group (~3 events per week)

Innovation

- Be the platform for field expert cross-fertilization
- Open the group to external innovation (startups, etc.)

WHY INVESTIGATING AI WITHIN VINCI?

To anticipate and mitigate the risk to be disrupted by other companies

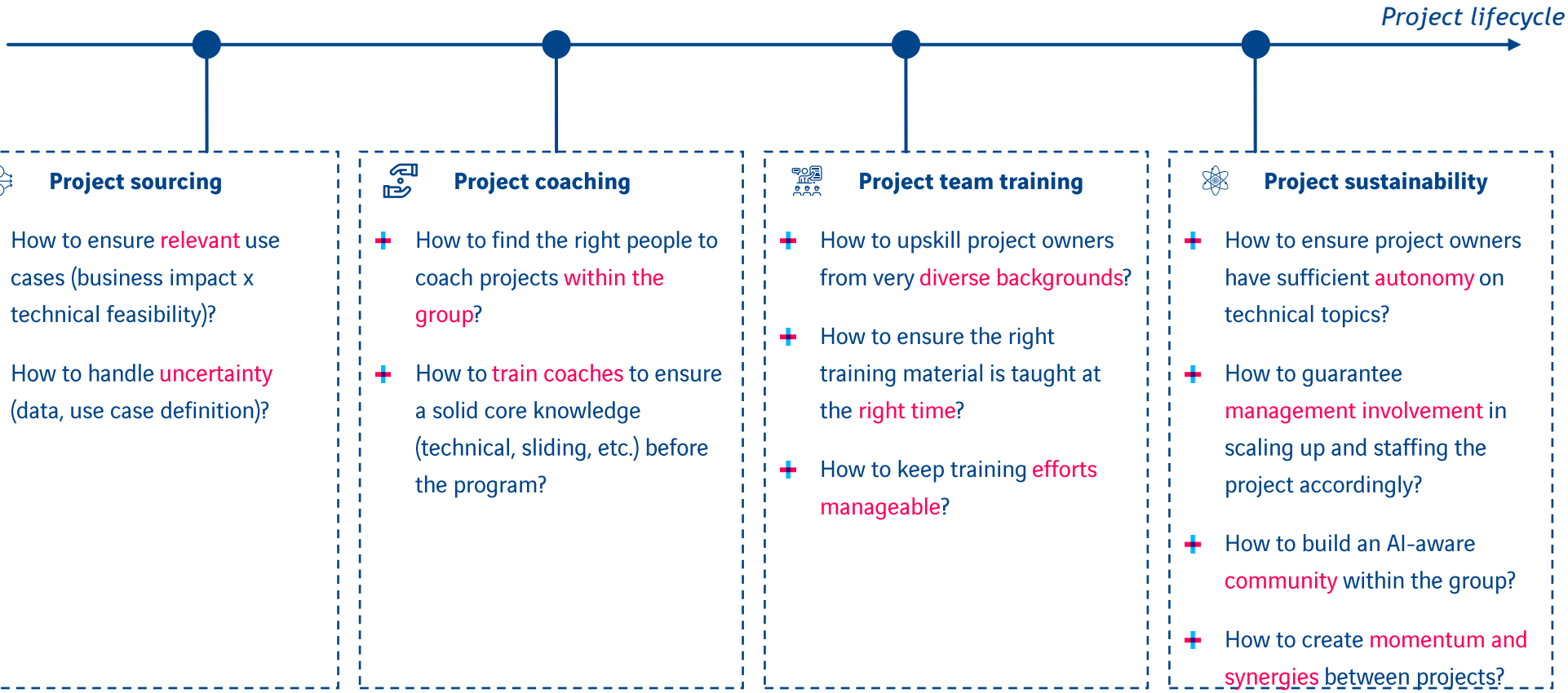
To increase VINCI performance and empower VINCI businesses with innovative AI technologies

To leverage VINCI collaborator ideas around highly potential AI applications for the group and accelerate their realization

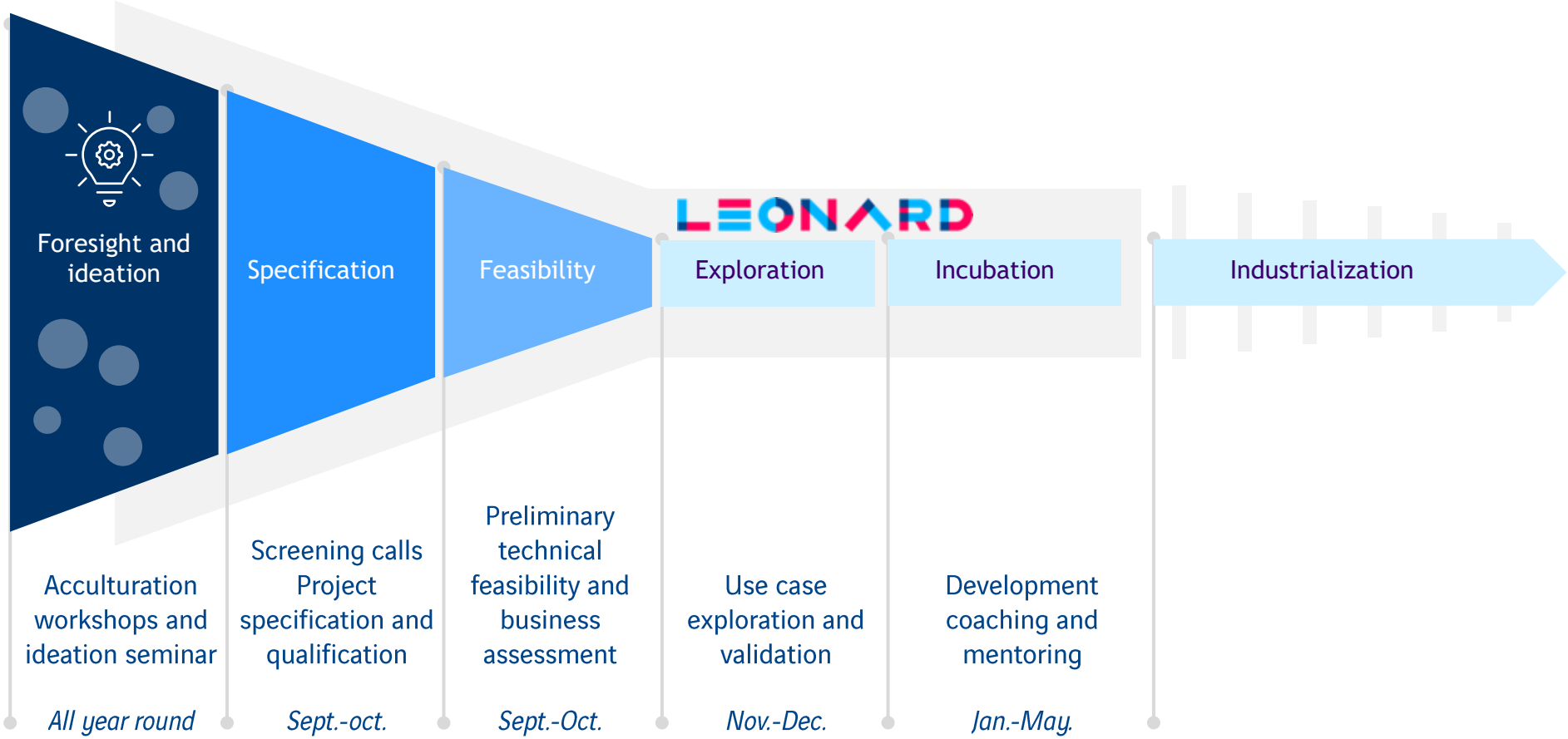
To improve VINCI collaborator skills and knowledge about AI

To share the risk behind highly uncertain projects

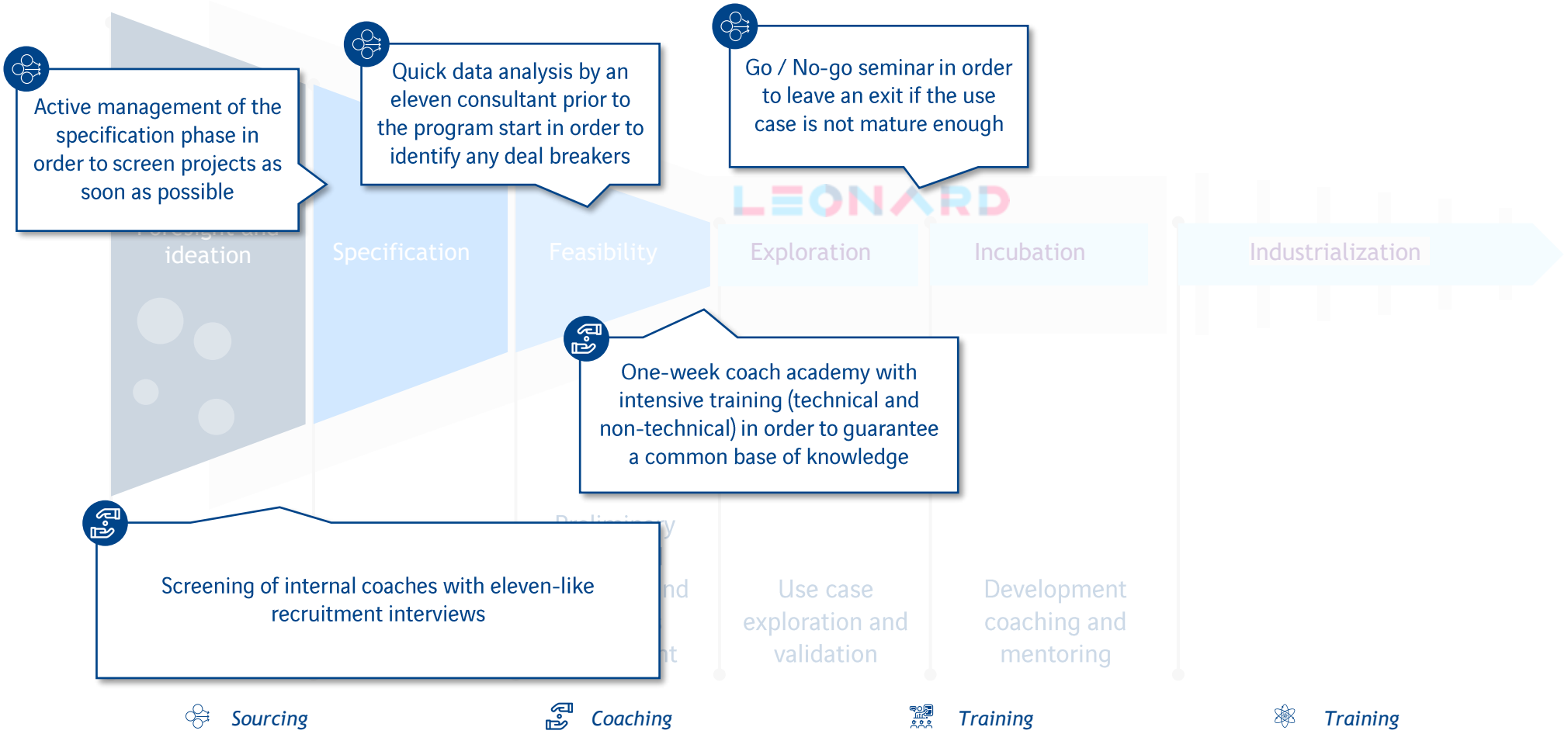
THE DESIGN OF THE AI PROGRAM AS AN INCUBATOR OF SUCCESSFUL PROJECTS ALSO FACED SEVERAL INTERNAL CHALLENGES ALL ALONG PROJECTS' LIFECYCLE



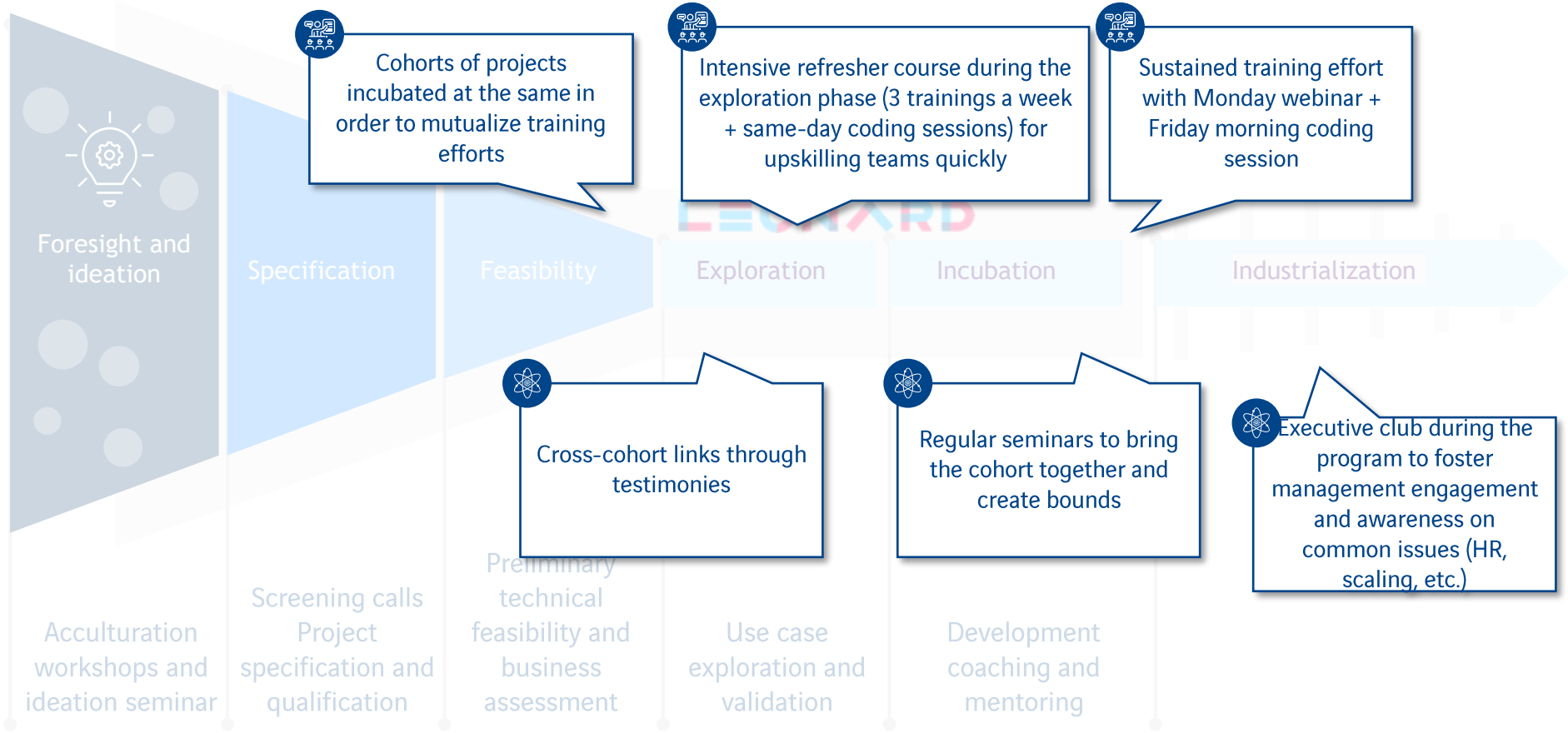
THE SPECIFIC ORGANIZATION AND TIMING OF AI PROGRAM COHORTS WAS DESIGNED TO ANSWER MANY OF THESE CHALLENGES



THE SPECIFIC ORGANIZATION AND TIMING OF AI PROGRAM COHORTS WAS DESIGNED TO ANSWER MANY OF THESE CHALLENGES



THE SPECIFIC ORGANIZATION AND TIMING OF AI PROGRAM COHORTS WAS DESIGNED TO ANSWER MANY OF THESE CHALLENGES



Today we will present two examples of industrial/corporate Machine Leaning projects, involving shallow Machine Learning and computer vision techniques

Markis

Improve public work bidding process thanks to an enhanced competitive intelligence

Business impact

- Anticipate market tension by predicting competitor likelihood to participate to the tender thanks to bidder tender history

AI approaches

- Text extraction
- Binary classification
- Interpretabi-lity tool

Data

- History of all tenders and bids
- Text-based tender description

TicketAI

Improve the efficiency of IT ticket resolution

Business impact

- Accelerate the upskilling of service desk employees
- Reduce the time taken to resolve tickets

AI approaches

- Textual embeddings
- Semantic similarity
- Vector search engines

Data

- Past tickets, with resolution steps
- Internal knowledge base

AInSAR

Automatize quality control of satellite data and extract valuable information from it

Business impact

- Reduce the workload of quality control
- Create high value-added products from raw data

AI approaches

- Image classification
- Edges detection
- Independent Component Analysis

Data

- Intermediary satellite images
- Final ground movement data

AGENDA

1. About Eleven
2. Digital Innovation projects
3. Competitive market intelligence
4. Natural Language Processing for IT ticket resolution
5. Computer Vision for satellite images quality inspection
6. Closing remarks

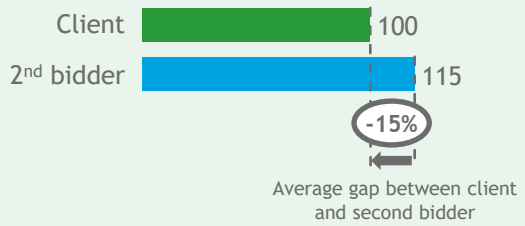
The Markis project aimed at predicting and analyzing the competition on public work tenders in Germany

A competitive market

- Public construction market
- In 2020:
  - **30k** public tenders
  - **10k** bidders
  - **18B€** market size
  - **600M€*** client revenue with public tenders

Bid calculation is key

- Lowest bidder wins the tender
- In average, 15% lost margin to second bidder by the client



Large database

- Existing database with all bids and results for public tenders
- Many potential additional data sources
- Too much information to be analyzed manually

Improving client's bidding process thanks to an enhanced competitive intelligence is key to increase margins and market share

** Client figures were modified*

Bid calculation is a time-consuming process with very uncertain outcome

Tender selection

100+ daily new tenders

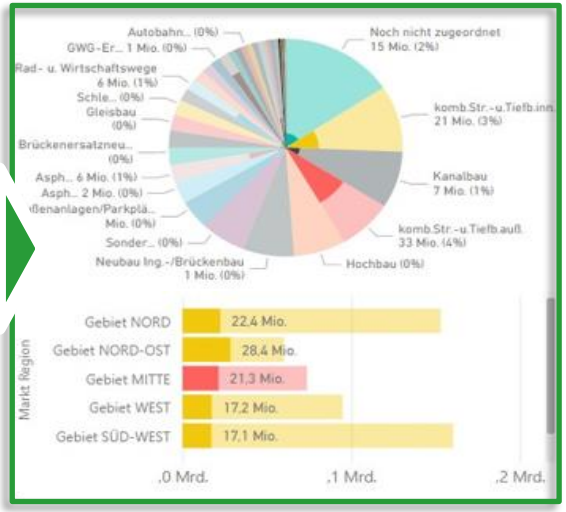
50% drop rate

Offer calculation

2 or 3 weeks process

Lost margin

Results analysis



Time-consuming

Often overlooked

Markis is a solution for automated market competitiveness analysis to support the decision-taking of the calculators and executives

Creation of a support tool for tender selection and competition analysis (1/7)

Key steps

1 A large and complex database

- Complex database of all past tender-results and future tenders with a high granularity
- Enriched with other data sources (duration, text extraction, geographical information [distances], etc.)
- Missing information about bidding behaviors

	ZS-Markt	ObjNr	SubDatum	ImportDatum	Objekt	Plz	BauOrt	Segment	AG	Objarten	Entf Bauort L
+	ZS Langenhorn	1808977	09.04.2021	22.03.2021	Neubau Entwässerungsanlagen - St-Jür	23560	Lübeck	komb.Str.-u.Tiefb.auß.	Entsorgungsbetriebe Lübeck	Straßenbau Kanalbau Schächte/-Abdeck	148
+	ZS Langenhorn	1807205	13.04.2021	15.03.2021	Neubau Rampe NW/SO Versorgungstur	23552	Lübeck	Brückensanierung	Deutsche Bahn AG	Brückensanierung Gleisbau Verbau Spun	146
+	ZS Langenhorn	1808315	13.04.2021	18.03.2021	Instandsetzung Rohrbruch Fernwärmel	24103	Kiel	Kabel-/Schutz-/sonst.Rohre (r	Gebäudemanagement Schleswig-Holstei	Fernwärme	88
+	ZS Langenhorn	1809507	13.04.2021	23.03.2021	San. Asphaltdeckschicht Geh- u. Radwe	23560	Lübeck	Asphaltdecken innerorts	Hansestadt Lübeck - FB 5 Planen und Ba	Straßenbau Fräsarbeiten (Asphalt)	148
+	ZS Langenhorn	1806653	13.04.2021	11.03.2021	Deckenerneuerung Liebigstraße	24145	Kiel	Asphaltdecken innerorts	Landeshauptstadt Kiel - Zentrale Vergab	Straßenbau Pflasterarbeiten Fahrbahnm	91
+	ZS Langenhorn	1810422	13.04.2021	25.03.2021	Tiefbauliche Erschließung B-Plan Nr. 69	23626	Ratekau	B-Plan-/Wohngeb.erschl.	Gemeinde Ratekau	Kanalbau Pflasterarbeiten Beleuchtung	145
+	ZS Langenhorn	1806855	14.04.2021	12.03.2021	Rückbau Gebäude 66, 67 u. Bunker	24848	Kropp	Abbruch	Gebäudemanagement Schleswig-Holstei	Abbruch / Rückbau Rodung / Baumfälla	47
+	ZS Langenhorn	1809409	14.04.2021	23.03.2021	Erneuerung der Verdichter Kläranlage F	24768	Rendsburg	Schleusen,Wehre,Pumpw.	Abwasserbeseitigung Rendsburg	Masch.-techn. Ausrüstung Elektrotechni	64
+	ZS Langenhorn	1809573	15.04.2021	23.03.2021	Neubau Trinkwassertransportleitung in	24229	Schwedeneck	Druckrohrleitungsbau	Wasserbeschaffungsverband Dänischer	Druckrohr/Gas/Wasser Straßenbau Pflas	81
+	ZS Langenhorn	1810715	15.04.2021	26.03.2021	Rüttelstopfverdichtung Julius-Leber-Kas	25813	Husum	Erdbau	Gebäudemanagement Schleswig-Holstei	Erdarbeiten	27
+	ZS Langenhorn	1809969	15.04.2021	24.03.2021	Faschinenbeschaffung 2021 Stützpunkt	25840	Friedrichstadt		Wasser- u. Schiffahrtsamt Tönning	Lieferleistungen	34
+	ZS Langenhorn	1810275	15.04.2021	25.03.2021	Erneuerung RW u. SW-Kanal Möwenstr	24235	Laboe	Kanalbau	AZV Ostufer Kieler Förde	Kanalbau Schächte/-Abdeckungen Straß	92
+	ZS Langenhorn	1811843	16.04.2021	01.04.2021	San. Infrastruktur Scheersberg	24972	Steinbergkirche	Druckrohrleitungsbau	Kreis Schleswig-Flensburg	Druckrohr/Gas/Wasser Kanalbau Kabel/-	54
+	ZS Langenhorn	1810496	19.04.2021	26.03.2021	Entwässerung/Korrosionsschutz B76/KI	24306	Plön	Sonderprojekte	Landesbetrieb Straßenbau u. Verkehr Sc	Abdichtung / Beschichtung Kanalbau	114
+	ZS Langenhorn	1808532	19.04.2021	18.03.2021	Hydraul. RW-Sanierung Ludewig-Meyn	24118	Kiel	grabenl.Verl.TW/Pressz.	Gebäudemanagement Schleswig-Holstei	Kanalsanierung (grabenlos) Kanalbau Sc	86
+	ZS Langenhorn	1810505	19.04.2021	26.03.2021	Geotechnische Erkundung Christian-Alb	24118	Kiel	Wasserbau	Gebäudemanagement Schleswig-Holstei	Brunnenbohrung	86
+	ZS Langenhorn	1808851	19.04.2021	19.03.2021	Rohbau Erw. Mensa u-Klassenräume Gy	24306	Plön	Hochbau	Kreis Plön	Hoch- / Rohbau	114
+	ZS Langenhorn	1812357	20.04.2021	07.04.2021	UI-Dick-und Dünnschichtmarkierung 20	24768	Rendsburg	Sonderprojekte	Landesbetrieb Straßenbau u. Verkehr Sc	Fahrbahnmarkierung	64
+	ZS Langenhorn	1811477	20.04.2021	31.03.2021	Fahrbahnerneuerung K PLÖ 13 Köhn - (	24321	Giekau	Asphaltdecken außerorts	Landesbetrieb Straßenbau u. Verkehr Sc	Straßenbau	114
+	ZS Langenhorn	1812361	20.04.2021	07.04.2021	Erneuerung bit. Fugen auf Brücken	24582	Bordesholm	Brückensanierung	Landesbetrieb Straßenbau u. Verkehr Sc	Brückensanierung Straßenbau Fugen-/Ri	92
+	ZS Langenhorn	1805944	20.04.2021	09.03.2021	Deckenerneuerung Hollbeck	23570	Lübeck	Asphaltdecken außerorts	Hansestadt Lübeck - FB 5 Planen und Ba	Straßenbau Graben- u. Bankettarbeiten	149
+	ZS Langenhorn	1808537	20.04.2021	18.03.2021	Grundinstandsetzung Uferwand Lagerp	25813	Husum	Wasserbau	Landesbetrieb für Küstenschutz, Nationa	Spundwände Ufersicherung Pfahlgründu	27
+	ZS Langenhorn	1809363	20.04.2021	22.03.2021	Küstenschutz-Regiemaßnahme 2021 Lis	25813	Husum		Landesbetrieb für Küstenschutz, Nationa	Lieferleistungen	27
+	ZS Langenhorn	1811249	20.04.2021	30.03.2021	Forstwegebau im Kreis Ostholstein	23689	Pansdorf	Rad- u. Wirtschaftswege	Landwirtschaftskammer Schleswig-Holst	Wirtschaftswegebau	140
+	ZS Langenhorn	1811866	21.04.2021	01.04.2021	Schadstellensan. Reparaturzug B-, L-, K-	24941	Flensburg	Sonderbauweisen Asphalt	Landesbetrieb Straßenbau u. Verkehr Sc	Oberflächenbehandlung - Asphalt	32
+	ZS Langenhorn	1811876	21.04.2021	02.04.2021	UI Landesstraßen	24941	Flensburg	Sonderbauweisen Asphalt	Landesbetrieb Straßenbau u. Verkehr Sc	Oberflächenbehandlung - Asphalt	32

Creation of a support tool for tender selection and competition analysis (2/7)

Key steps

1 A large and complex database

- Complex database of all past tender-results and future tenders with a high granularity
- Enriched with other data sources (duration, text extraction, geographical information [distances], etc)
- Missin[illegible]mation about bidding behaviours

Example of input data

Regex extraction

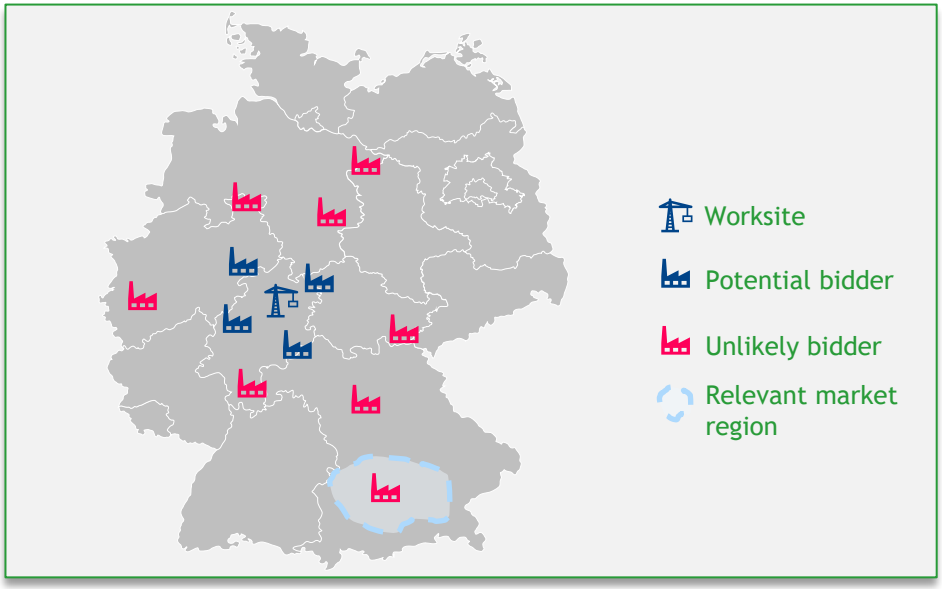
Creation of a support tool for tender selection and competition analysis (3/7)

Key steps

1 Enriched database

2 Pre-selection of bidders for each tender

- Definition of a relevant market region for each competitor unit based on past bids
- Pre-selection of a list of potential bidders for each coming tender



Creation of a support tool for tender selection and competition analysis (4/7)

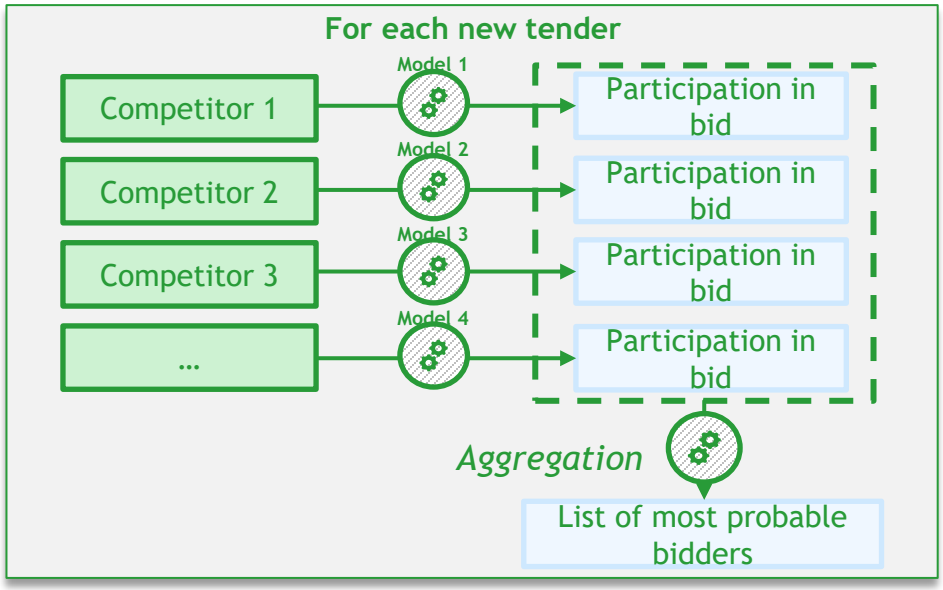
Key steps

1 Enriched database

2 Pre-selection of bidders

3 AI modelling of competitor participation

- For each tender, prediction of a list of most probable bidders
- Training and optimization of AI models per competitor(~ 1,500 models)



Creation of a support tool for tender selection and competition analysis (5/7)

Key steps

1 Enriched database

2 Pre-selection of bidders

3 AI modelling of competitor participation

- For each tender, prediction of a list of most probable bidders
- Training and optimization of AI models per competitor(~ 1,500 models)

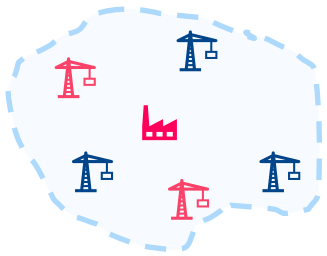
Features

- Distance to worksite
- Type of work
- Tags
- Materials
- Estimated amount
- Current load
- Rolling revenue

Enrichment

Selection bias (only positive samples)

- Use of close tenders to correct the bias



Pipeline

Extraction + Enrichment

Feature selection

Scaling + PCA

Oversampling

Grid Search + Random Forest

Creation of a support tool for tender selection and competition analysis (6/7)

Key steps

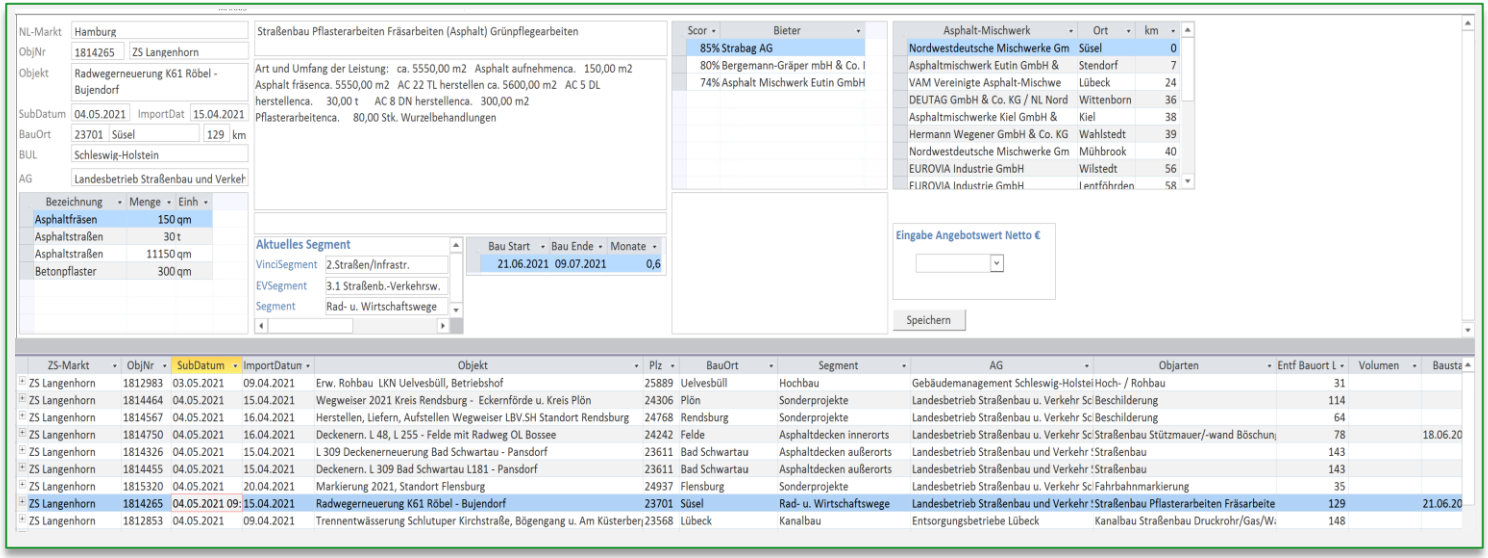
1 Enriched database

2 Pre-selection of bidders

3 AI models

4 Decision support tool

- Simple and interactive user interface
- Interpretability tool for the end-user to understand the prediction of the models
- List of targeted KPIs and features to help the decision-taking



Creation of a support tool for tender selection and competition analysis (7/7)

Key steps

1. **Enriched database**
2. **Pre-selection of bidders**
3. **AI models**
4. **Decision support tool**

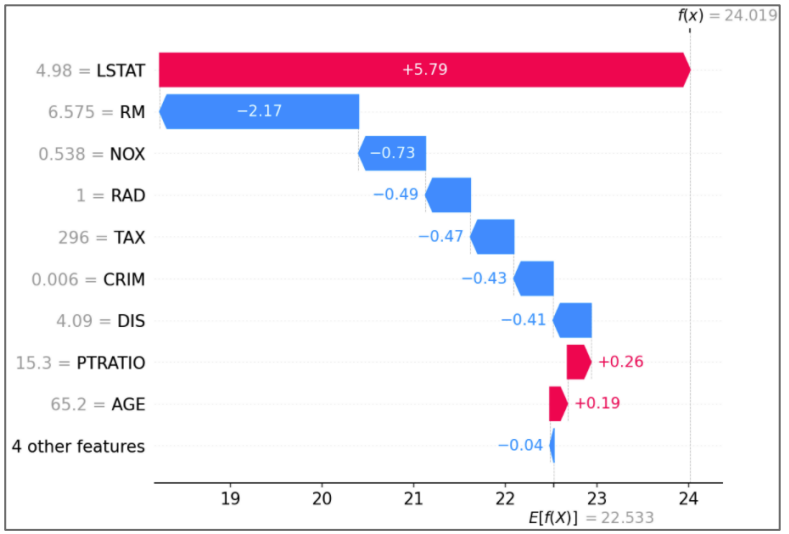
- Simple and interactive user interface
- Interpretability tool for the end-user to understand the prediction of the models
- List of t… …eted KPIs and features to help the

Why interpretability ?

Interpretability can be a key requirement in applied Machine Learning project

- **Validation** (detect model's bias)
- **Adoption** (the end user need to trust the tool)
- **Regulation** (the authorities require decisions to be explainable)

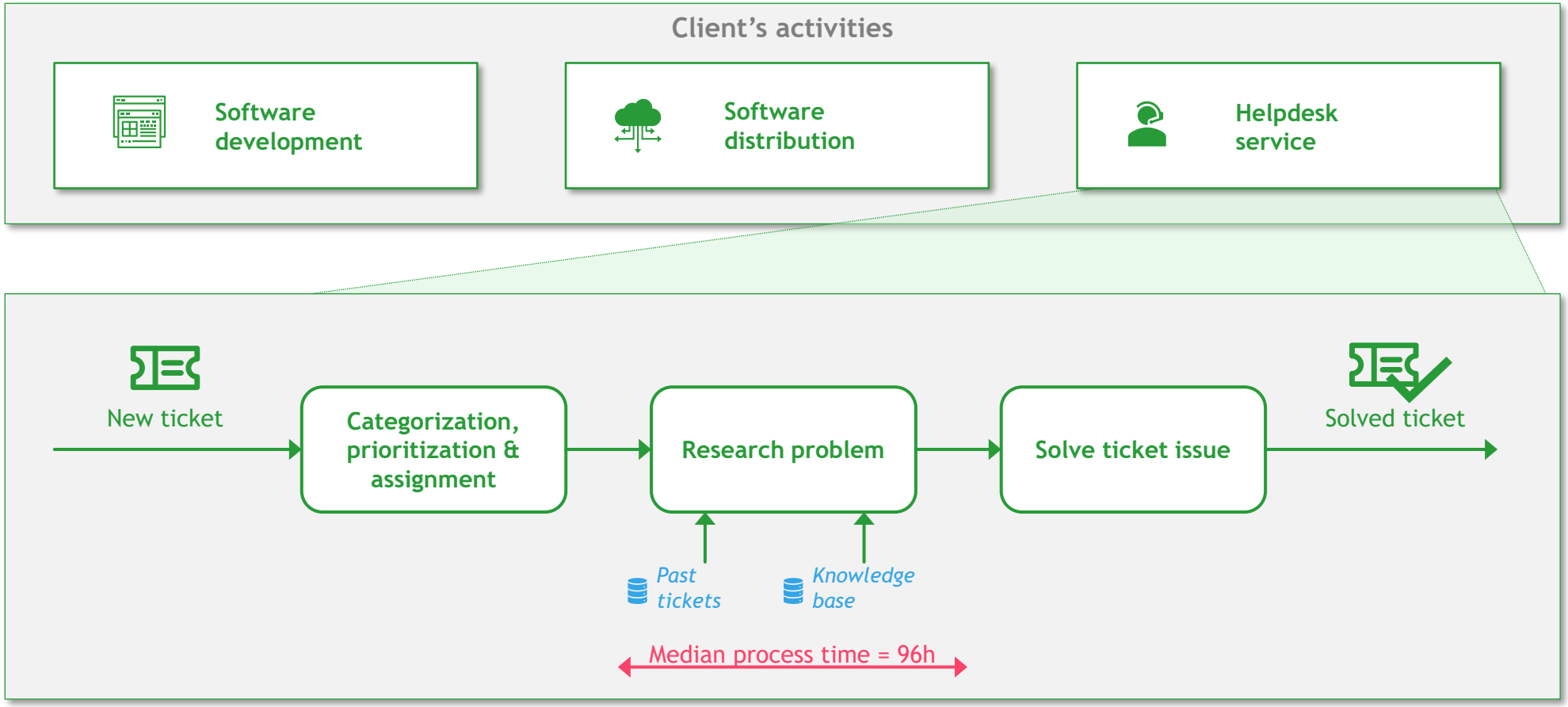
Illustration of Shapley values for a prediction (from library documentation)



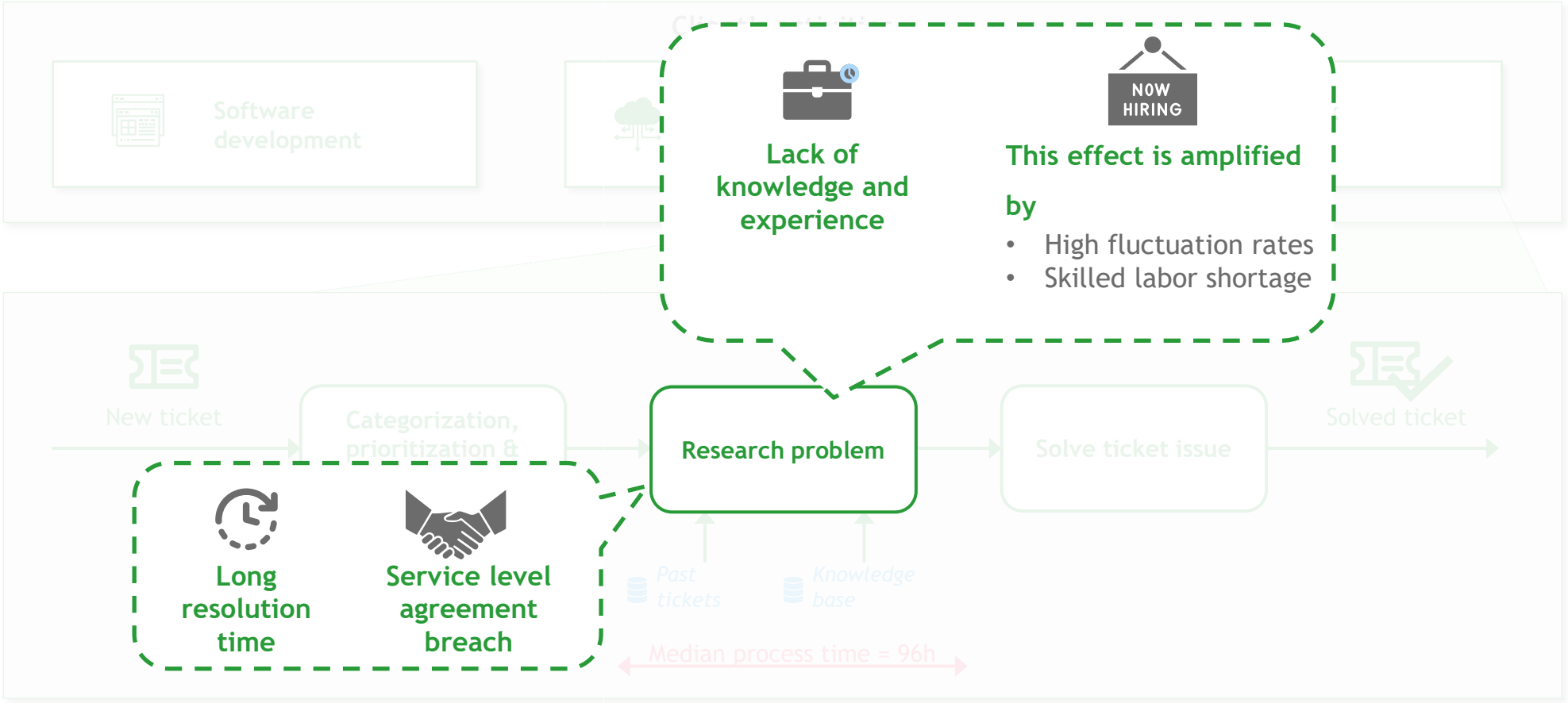
AGENDA

1. About Eleven
2. Digital Innovation projects
3. Competitive market intelligence
4. Natural Language Processing for IT ticket resolution
5. Computer Vision for satellite images quality inspection
6. Closing remarks

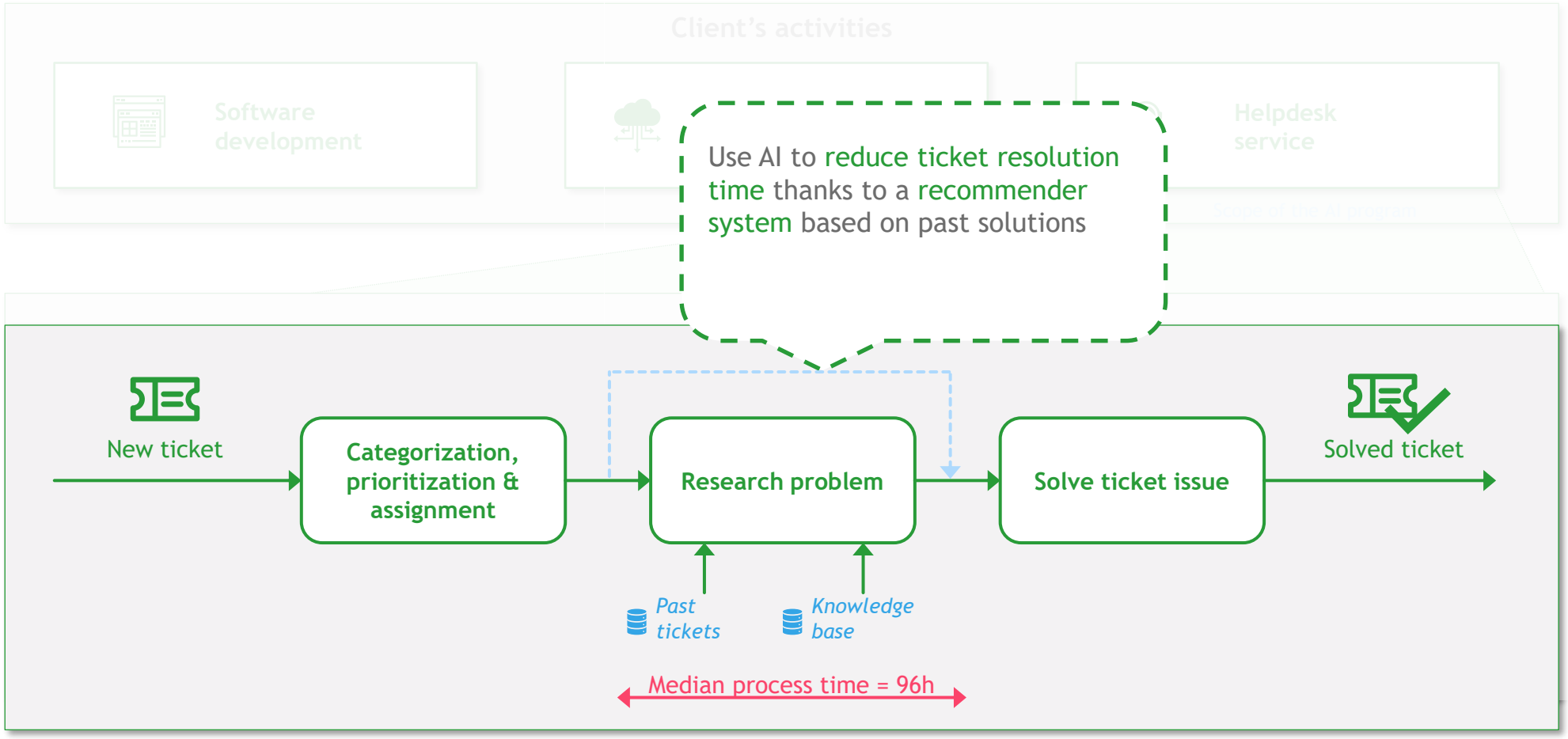
The client offers software for municipalities and supports their customers with their helpdesk services



Researching the solution to a ticket currently takes a long time because the staff lacks experience and past tickets cannot be searched for solutions

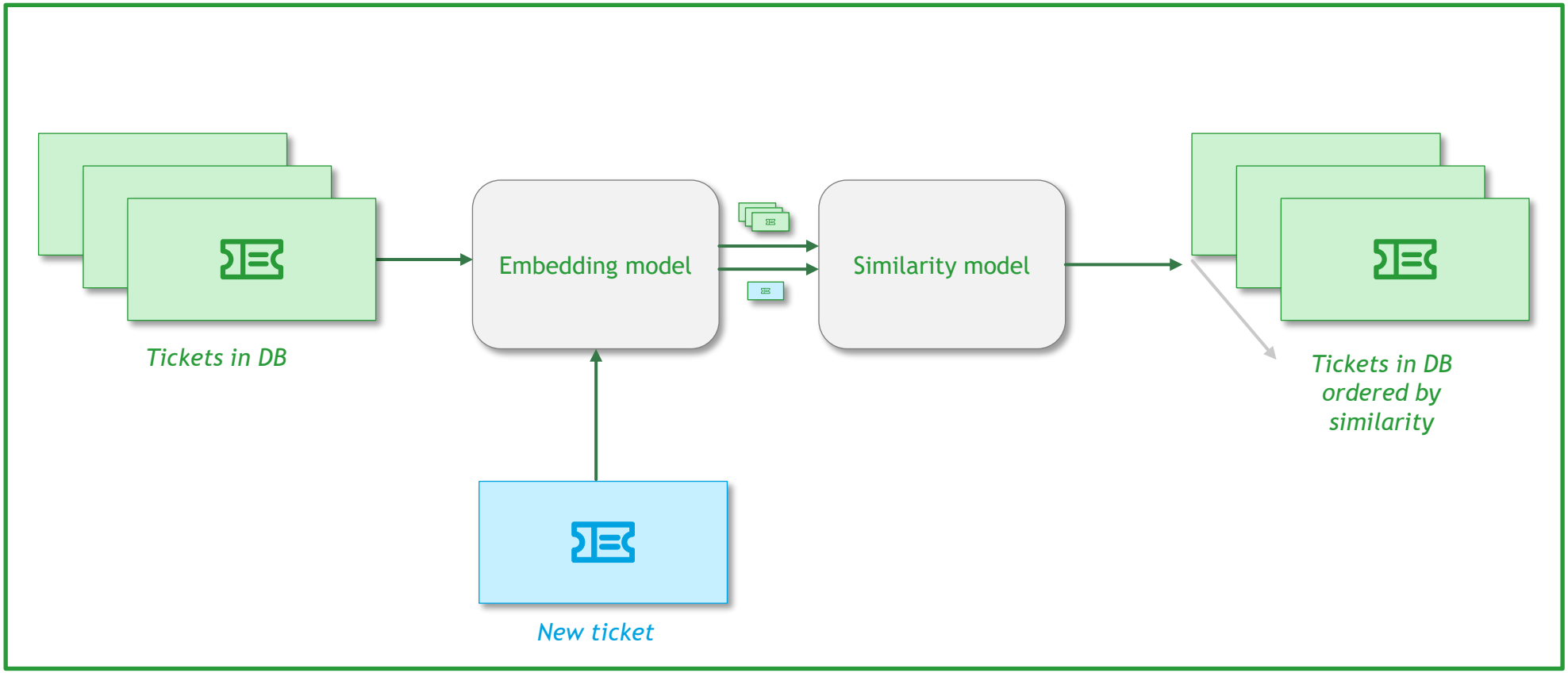


The solution aims at recommending solutions based on similar tickets in the historical database



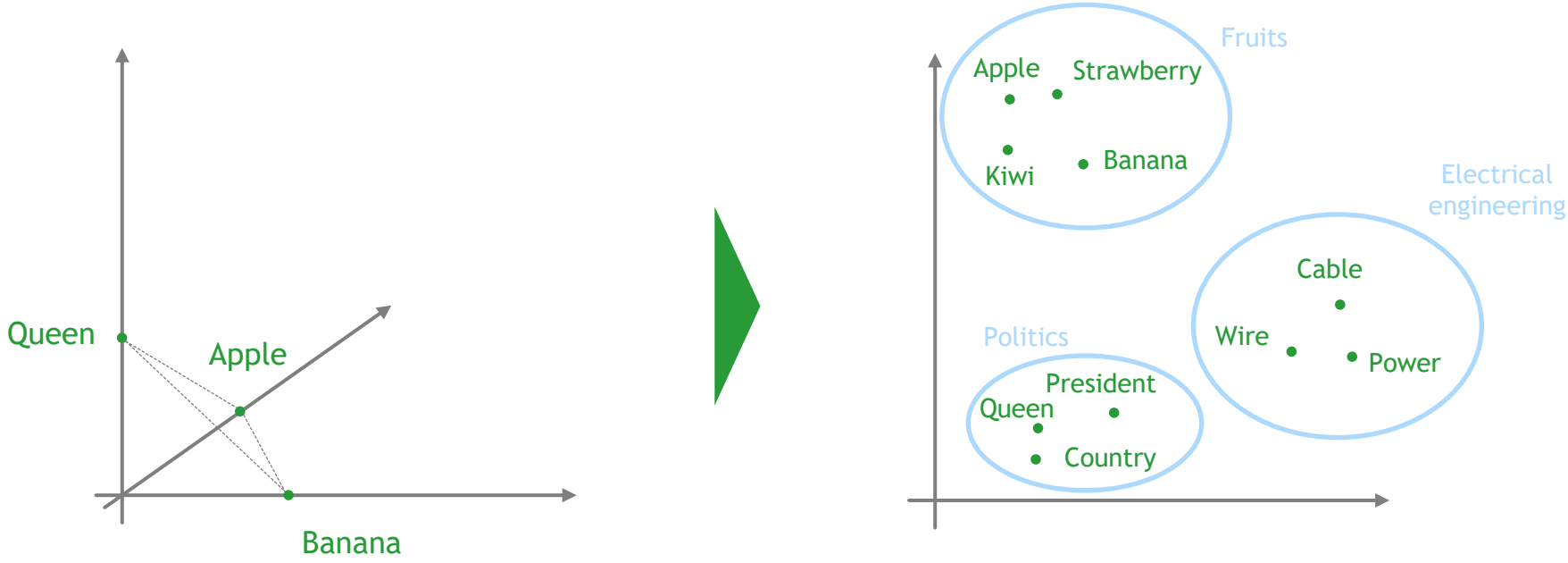
The recommender system is built around an embedding model and a similarity model which allow finding related tickets

Implemented Architecture



Text embedding is one of the big ideas of NLP, and consists in finding a vector representation of text such that similar words result in close vectors

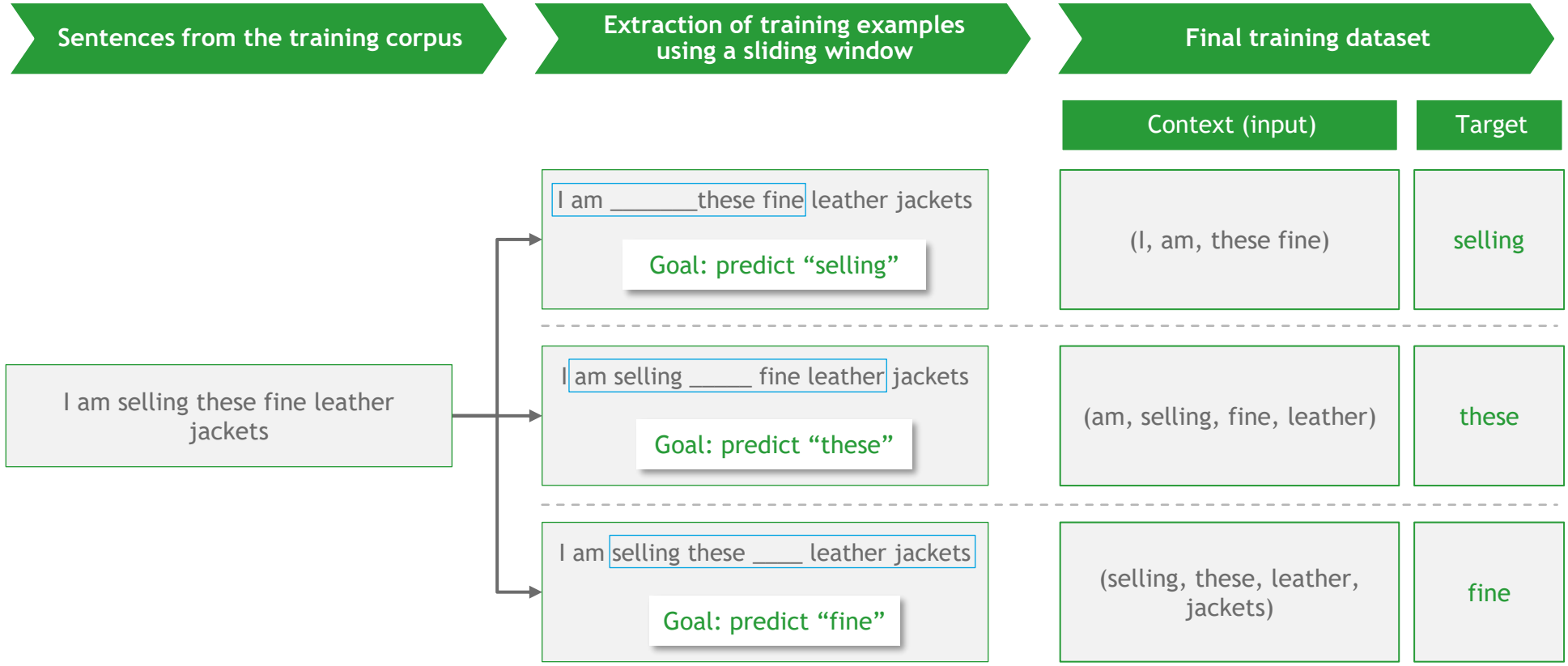
WORD EMBEDDINGS HAVE LOWER DIMENSION AND ENCODE WORD MEANING



Words that are similar are close in the embedding space

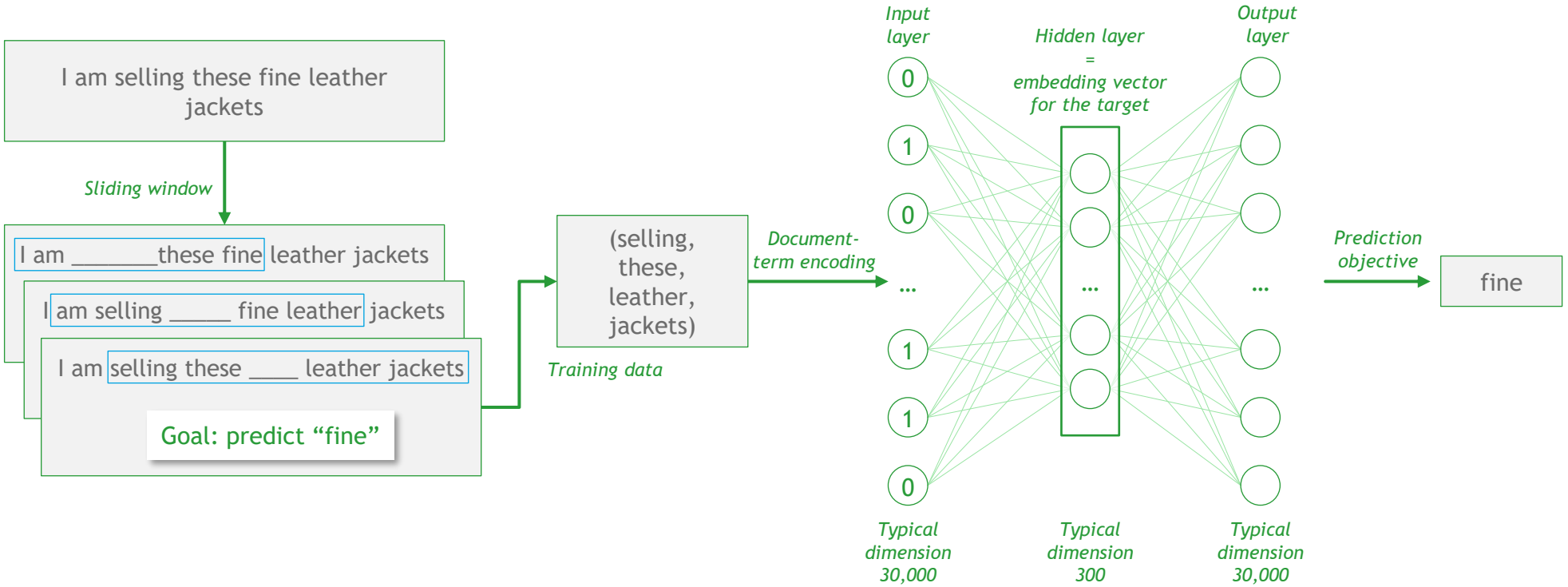
Most Word embedding models are trained to predict words from their context

Example: continuous bag of words (CBOW) approach to building the training data



Embedding models are mostly based on neural network language models (NNLM)

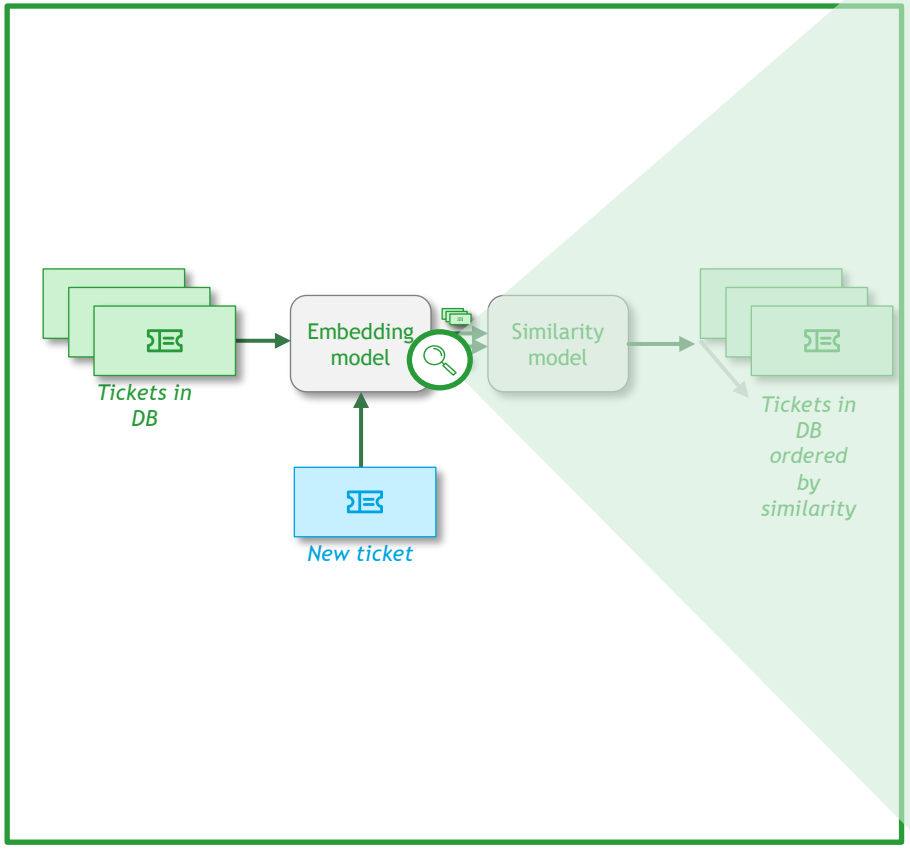
Example: word2vec is a neural network with one hidden layer trained using the CBOW approach



Word2vec embeddings will be similar for words occurring in similar contexts

The Embedding models differentiate in terms of Context, word ordering and out-of-vocabulary

Implemented Architecture



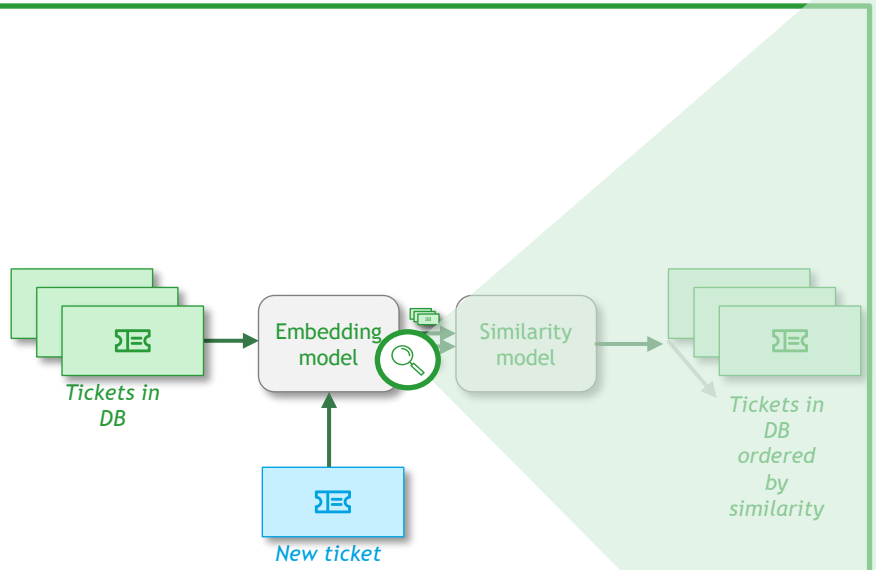
Embedding Comparison

- Embedding methods aim at representing words as a numerical vectors (i.e., list of numbers)
- Words that are semantically similar (e.g. synonyms) will be represented as vectors that are numerically close in the embedding space
- Several methods exist to compute word or document embeddings

	CONTEXT DEPENDENT	WORD ORDER	EMBEDDINGS	OUT-OF-VOCABULARY
Word2vec	✗	✗	Word	✗
fastText	✗	(✓)	Character-grams	✓
BERT	✓	✓	Sentence	✓

The Embedding models differentiate in terms of Context, word ordering and out-of-vocabulary

Implemented Architecture

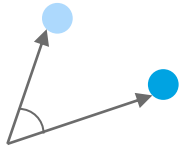


Distance Calculation Comparison

COSINE

- Cosine of the angle between two vectors
  $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$
- Frequently used for text analyses when data is represented by word counts and magnitude is not important

$$D(x, y) = \cos(\theta) = \frac{x \cdot y}{\|x\| \, \|y\|}$$



- Fast - even in high dimensional data

- Uneven lengths of documents
- Magnitude of vectors is not considered

WORD MOVERS DISTANCE

- Text is represented as a vector space
- Uses **earth mover's distance** to measure the **minimum distance** required for a word in one text to reach a word in another text in the semantic space

- **Semantic** relations among words are considered

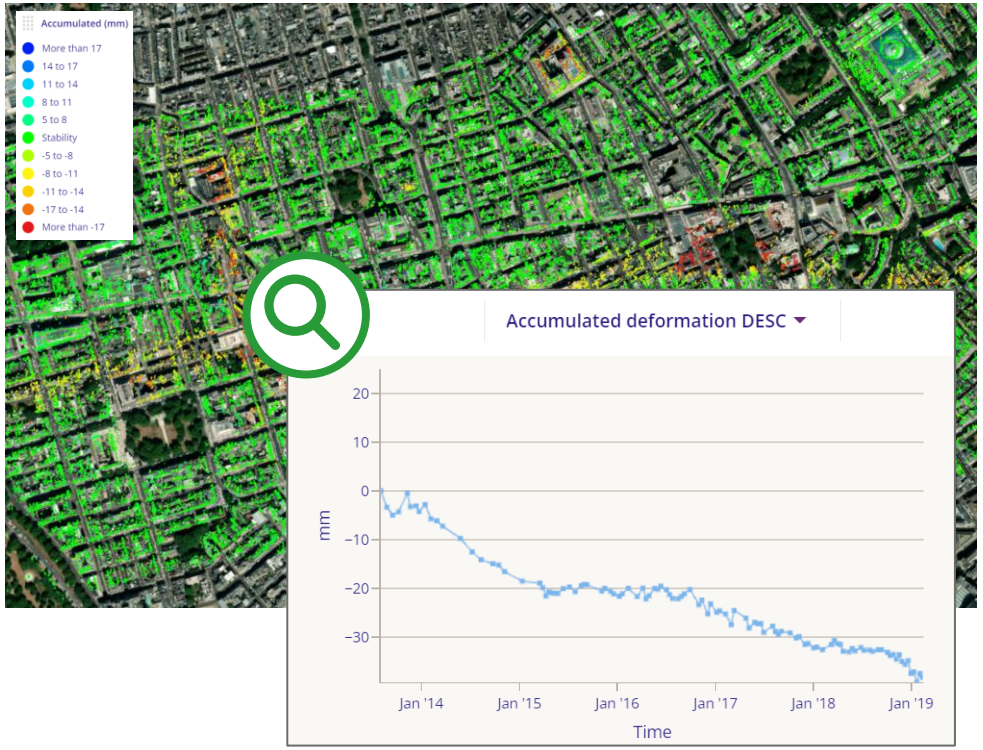
- Computational Expensive
  → Solution: Relaxed-WMD?

AGENDA

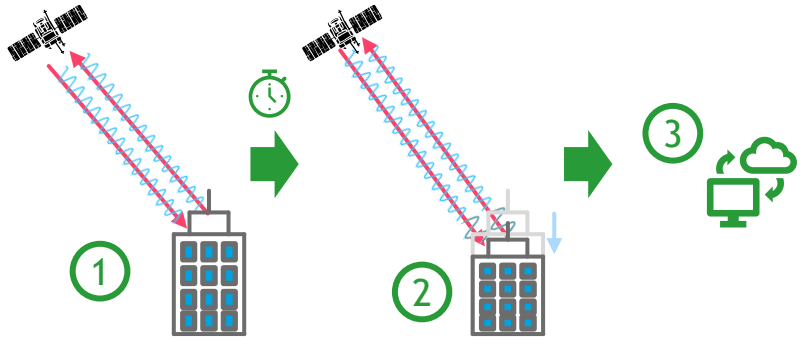
1. About Eleven
2. Digital Innovation projects
3. Competitive market intelligence
4. Natural Language Processing for IT ticket resolution
5. Computer Vision for satellite images quality inspection
6. Closing remarks

InSAR is a technology that allows to monitor the surface deformation through time from satellites

Illustration of the InSAR outputs



Interferometric Synthetic Aperture Radar (InSAR)



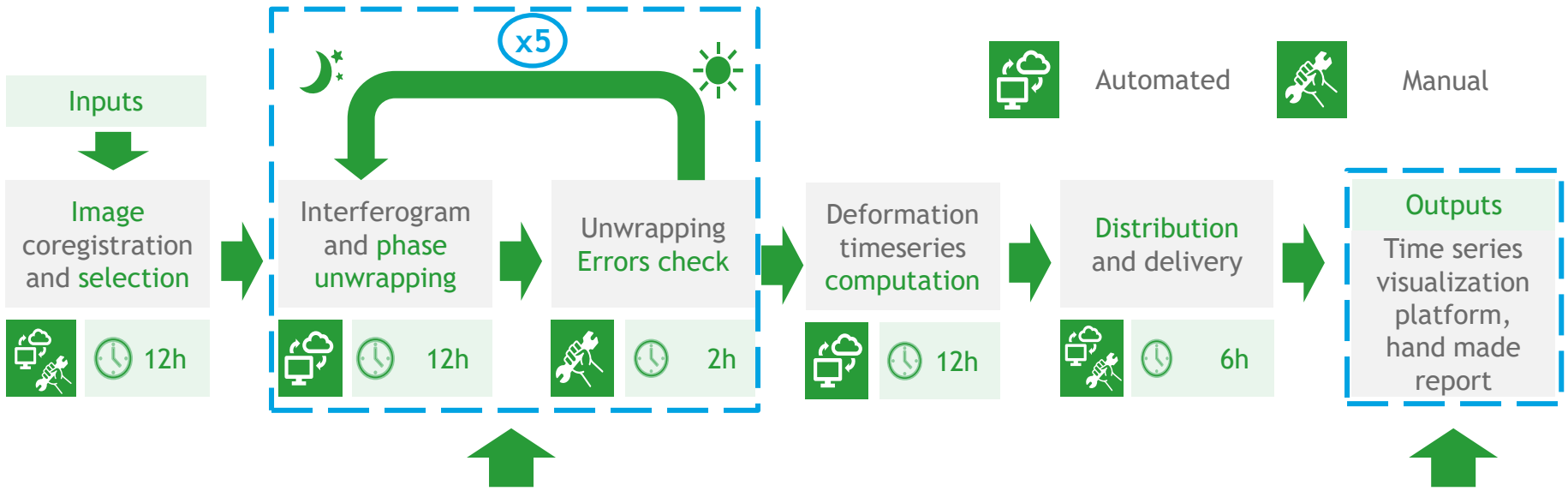
1. The satellite sends a radio wave and measures its travel time
2. On an image taken by a consecutive satellite pass, the travel time will have changed due to changes in the surface
3. Changes in surface are measured thanks to complex algorithms able to interpret phase differences

Remarks

- Since it does not rely on optical signals, InSAR is not affected by clouds or luminosity
- InSAR reaches millimetric precision in its estimate of elevation
- Due to technical limitations, InSAR is only suited for small deformations (up to a few centimeters)

The product's production pipeline faces two challenges regarding the quality control and the final delivered value

Production pipeline



Processing pipeline automation
(Unwrapping Errors Detection)

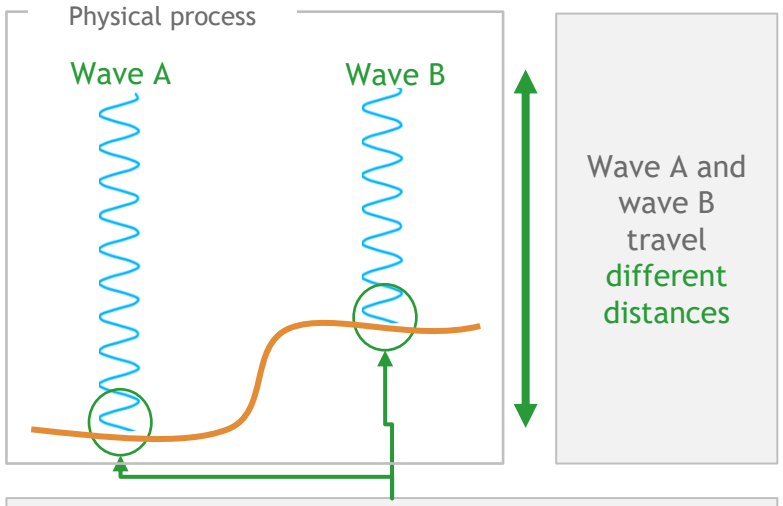
The processing pipeline is almost automated, but one time-consuming manual step remains

Advanced Analytics Tools

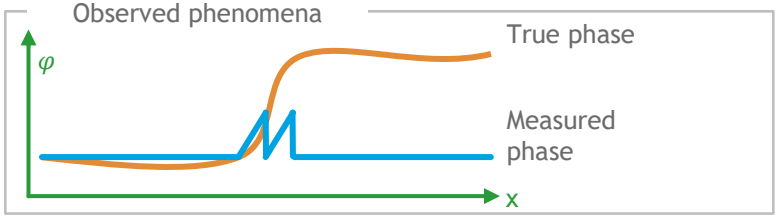
Both clients and the internal team proved interested in advanced tools to analyze the InSAR data

Unwrapping errors occur when the phase decomposition algorithm fails and creates regular discontinuities on the image

Unwrapping illustration



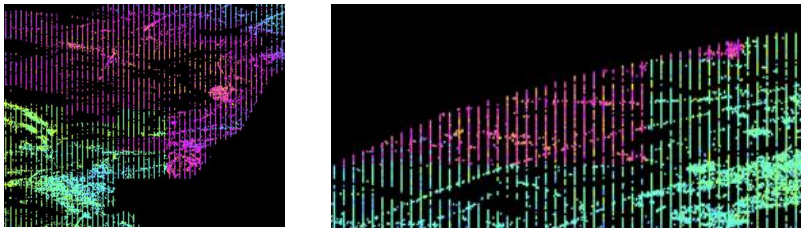
But since the phase periodic, the measured phase is the same!



Unwrapping examples

Unwrapping consists in finding where the periodicity created jumps, and removing them from the measurement

This process is error prone, but errors seem easily identifiable



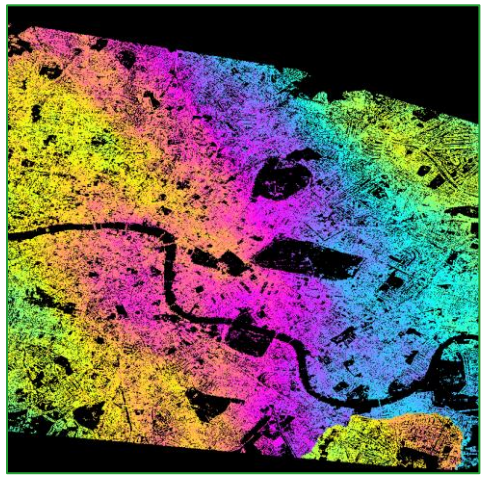
Errors detection can be framed as an image classification problem, for which one can leverage Deep Learning models, with some limitations…

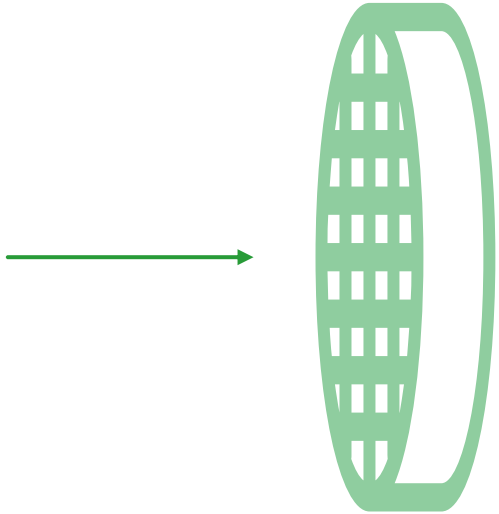
Process illustration

Original image

Convolution model

Prediction

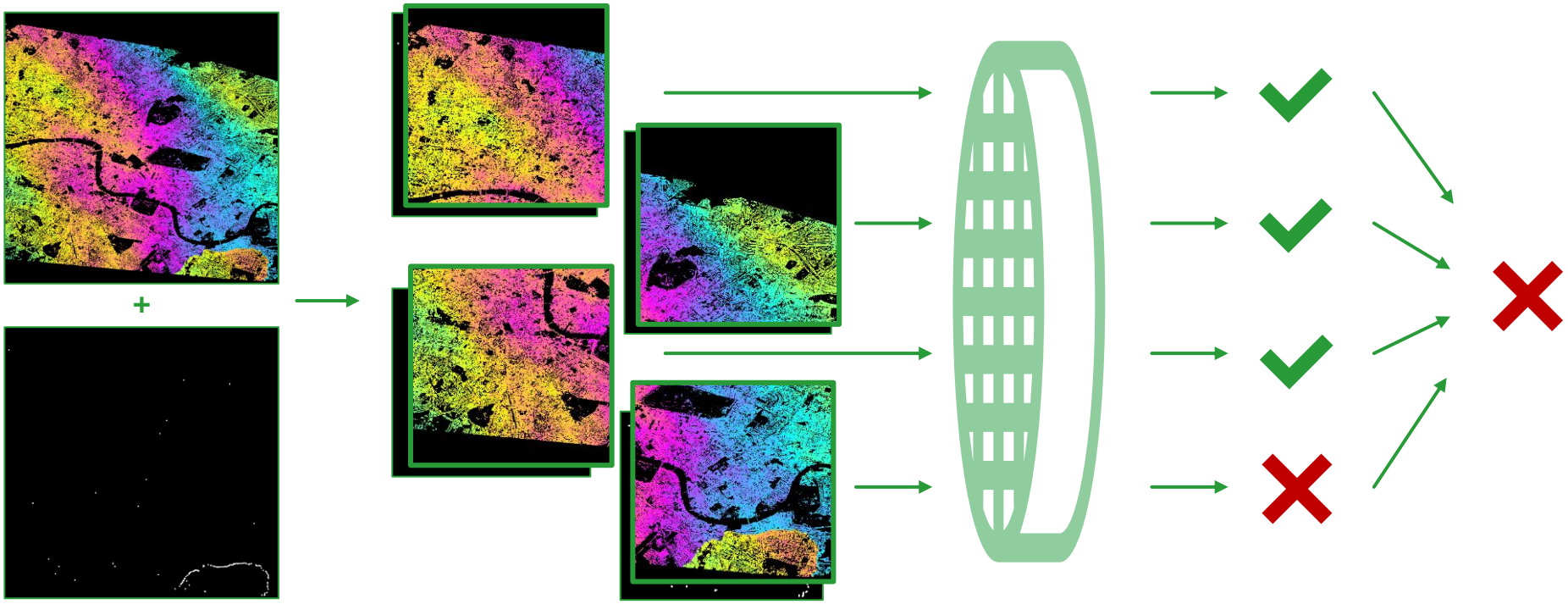




- This approach does not leverage the natural structure of the pattern to identify
- The image cannot be processed at full resolution
- No help is provided to help the inspector identify the error in case of doubt
- The original dataset might not be large or diverse enough

…which can be tackled by enriching raw images with edges and classifying each image patch with the deep learning model prior to aggregation

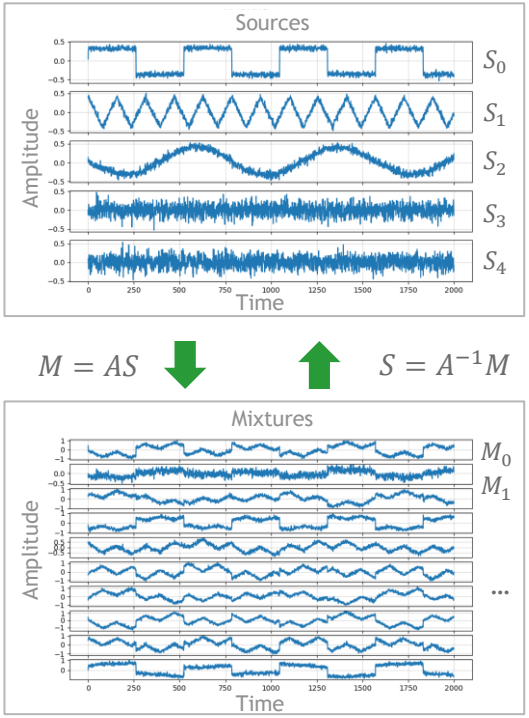
Process illustration



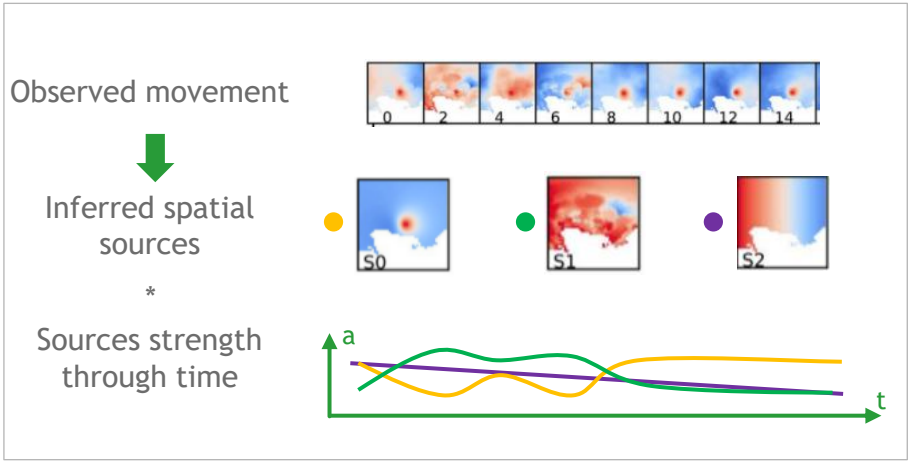
Source identification was implements through independent component analysis (ICA)

Illustration

ICA 101



ICA ON INSAR DATA



Overview

- ICA for InSAR data was implemented adapting a recent **tool from a public research team**
- ICA **decomposes the movement** of the points into different sources, each a spatial pattern
- Relevant results were obtained quickly when using this tool

AGENDA

1. About Eleven
2. Digital Innovation projects
3. Competitive market intelligence
4. Natural Language Processing for IT ticket resolution
5. Computer Vision for satellite images quality inspection
6. Closing remarks

On a data project, a standard week for an eleven consultant would include alternation between data science research, field expert interviews, training session for clients and meeting with management

	Monday	Tuesday	Wednesday	Thursday	Friday
AM	General training on an AI topic provided to client	Coding session	Coding session	Coding session	eleven training: Data Petit Dej
	Preparation of the weekly sprint	Preparation presentation of the progress	Client Workshop	Prepare potential training	eleven training: Strategy, Business, etc.
	Lunch				
PM	eleven standup	Team Footing	eleven standup	Prepare update	Coding session
	Sprint kick-off with client	Client standup	Client standup	Sprint update with client	Code training provided to clients
	Coding session	Coding session	Potential monthly steering committee	Coding session	1-1 mentor
			Project planification		

general session with client | eleven internal session | individual session with client | individual work

Merci pour votre attention !

Suivez-nous :