IMT Mines Alès
École Mines-Télécom

RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

ELISA

A requirement ELIcitation Smart Agent

https://factory.euromov.eu/elisa

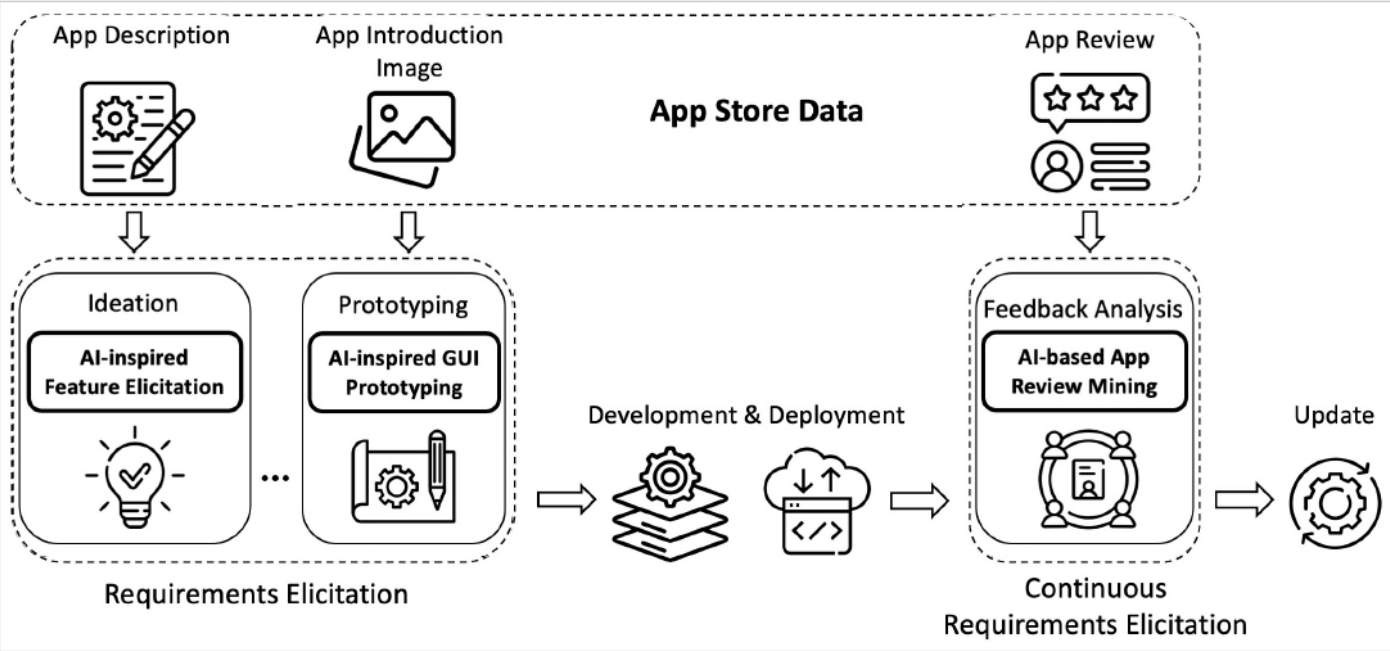
Anne-Lise Courbis, Thomas Lambolais
Euromov-DHM, Univ. Montpellier, IMT mines Ales, Ales, France
Jialiang Wei
Hamburg Univ., Germany

GDR GT IE
Toulouse, 20 novembre 2025

CONTEXT

➢ Enhancing Requirements Elicitation through App Stores Mining,

Jialiang Wei, PhD IMT mines Ales, November 2024,
Supervisor: G. Dray, P.L. Bernard*, A.-L. Courbis*, T. Lambolais**,
**Euromov-DHM, Univ. Montpellier, IMT mines Ales, Fr.*

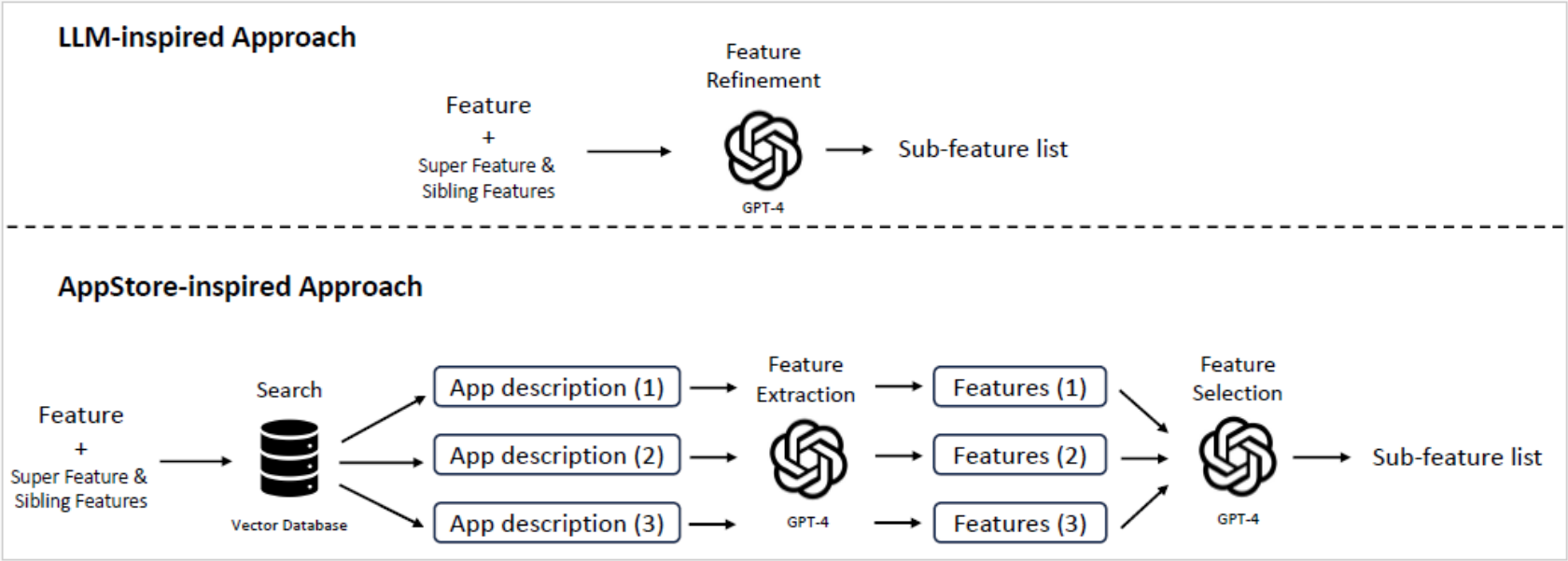


Overview of contribution

Source: J. Wei, PhD thesis, Chapter 1.

CONTEXT

- Enhancing requirements Elicitation through App Stores Mining,
  Jialiang Wei, PhD IMT mines Ales, November 2024,
  Supervisor: G. Dray, P.L. Bernard*, A.-L. Courbis*, T. Lambolais*,*
  **Euromov-DHM, Univ. Montpellier, IMT mines Ales, Fr.*



Overview of the 2 approaches for feature extraction

Source: J. Wei, PhD thesis, Chapter 3.

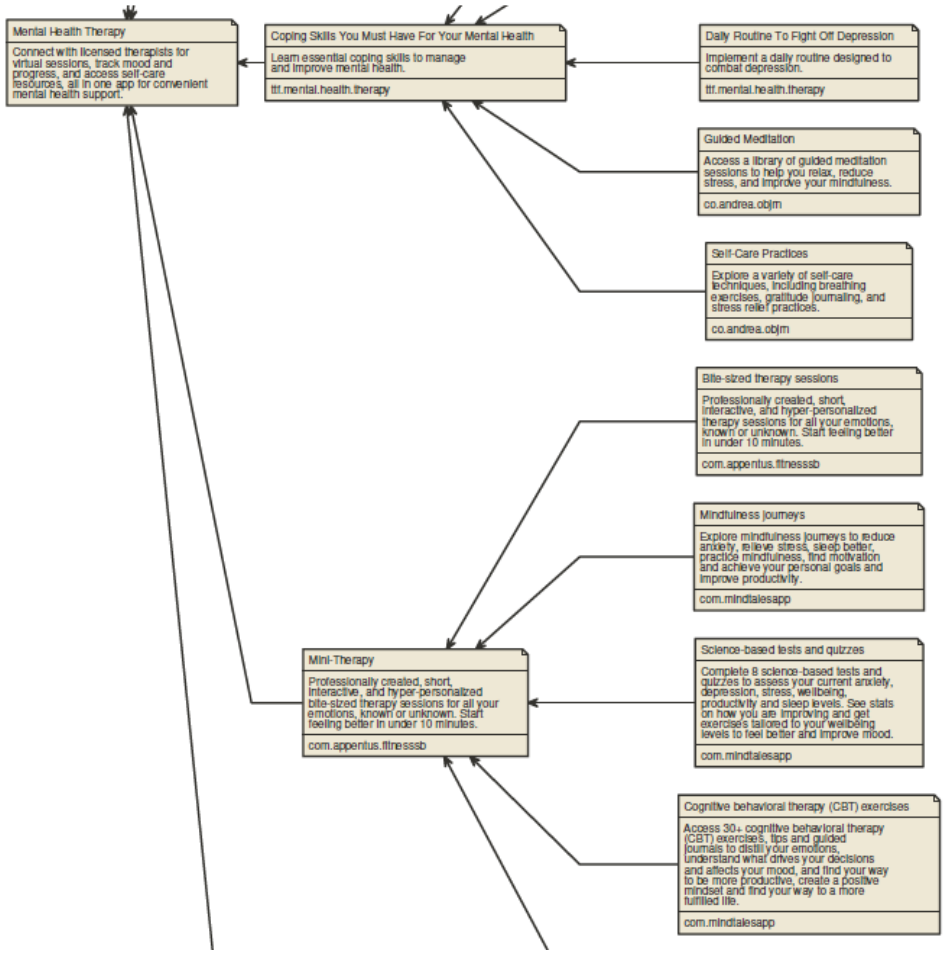
CONTEXT

- Enhancing requirements Elicitation through App Stores Mining,
  Jialiang Wei, PhD IMT mines Ales, November 2024,
  Supervisor: G. Dray, PL. Bernard*, A.L. Courbis*, T. Lambolais*,*
  **Euromov-DHM, Univ. Montpellier, IMT mines Ales, Fr.*

Existing features	Novel features
Anti Smartphone Addiction	Contextual Soundscape
Criminal Alert	Driver Guardian
Interior Design	Interactive Historical Overlay
Mental Health Therapy	Laugh evaluation
Parking Space Finder	Mood-Adaptive UI
Random Chat	Predictive Subscription Management
Supermarket Checkout	Social Health Analytics
Travel Planner	Symbiotic Music Creation
Virtual Fashion Assistant	Synesthetic Sensory Augmentation
Voice Translation	Thought reading

Root feature used for evaluation

Source: J. Wei, PhD thesis, Chapter 3.



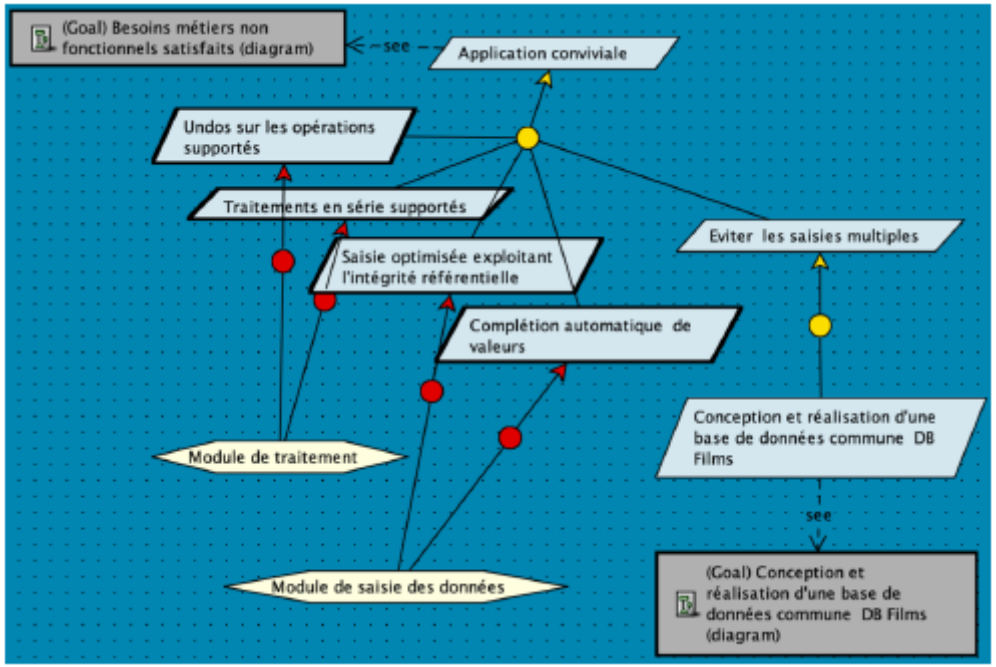
Feature tree generated by App Store mining

Source: J. Wei, PhD thesis, Chapter 3.

CONTEXT

➢ From KAOS diagram to textual description

R&D Mission 2025, IMT mines Ales. Students: A. Dadani, L. Vasseur
Supervisor: R. Darimont, C. Ponsard*, A.L. Courbis**, T. Lambolais**,*
** Respect-IT, Ottignies-Louvain-la-neuve, Belgium.*
***Euromov-DHM, Univ. Montpellier, IMT mines Ales, Ales, Fr.*



Le diagramme XML représenté ci-dessus est un diagramme KAOS pour une application conviviale. Le but global du diagramme est de définir les exigences pour une application qui vise à éviter les saisies multiples et à concevoir et réaliser une base de données commune pour les films.

Les nœuds importants dans ce diagramme sont les suivants :

** La goal "Application conviviale" qui est le nœud parent de plusieurs autres goals.*
** La goal "Éviter les saisies multiples", qui est elle-même un sous-but de la goal "Application conviviale".*
** Les requirement "Travailler avec des données en masse", "Undos sur les opérations supportés" et "Saisie des données", qui sont toutes associées à la goal "Application conviviale".*

…/…

Kaos Diagram and its corresponding generated textual description

Source: A. Dadani, L. Vasseur, rapport de Mission R&D.

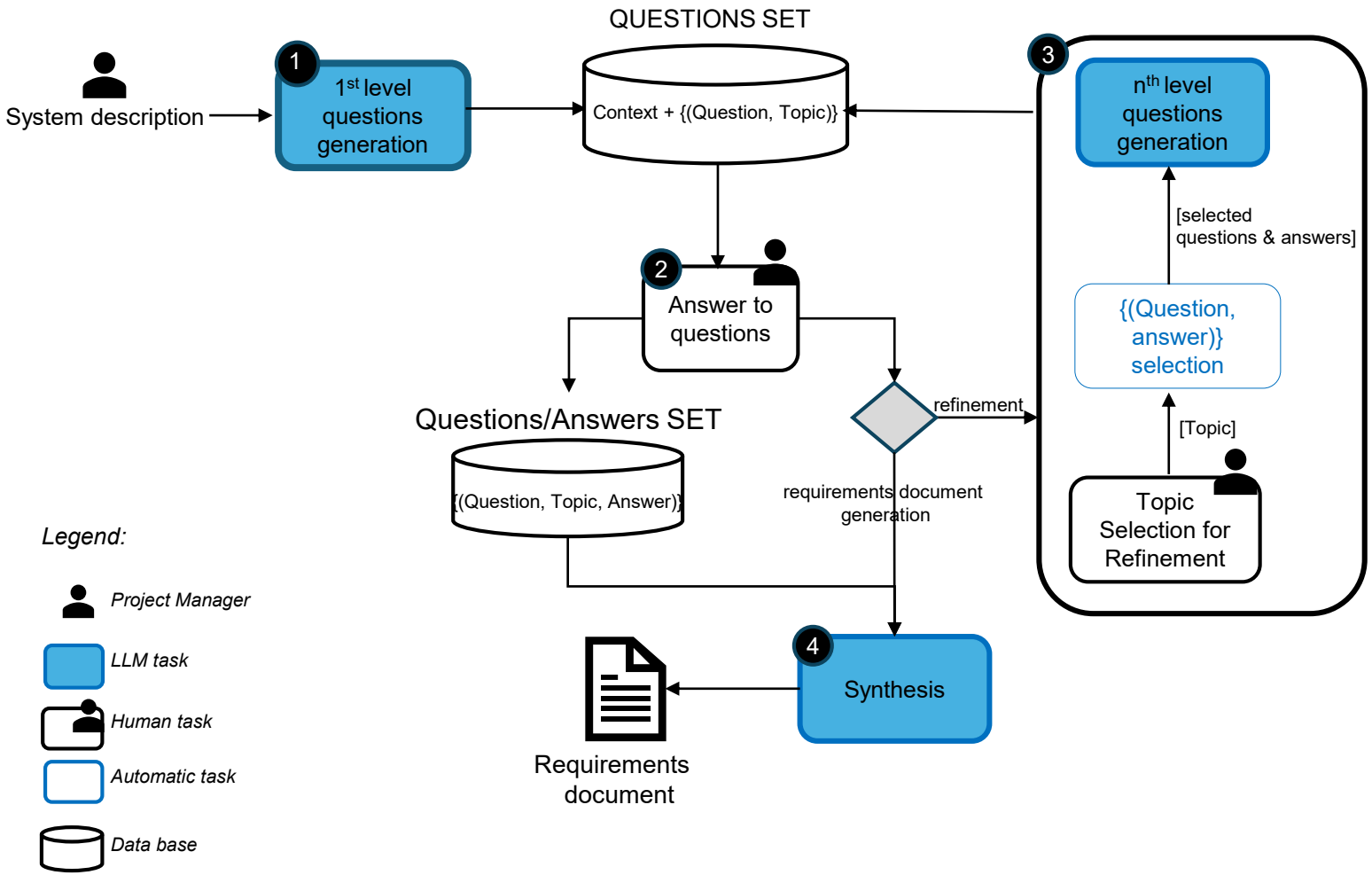
MOTIVATION

Are LLMs capable of conducting interviews to collect requirements ?

Hypotheses:

- Interview method:
  - One stakeholder (P.M. Project Manager) who responds truthly, in earnest
  - Input: a simple description of the system to be developed
  - Iterative process
    - Generation of a questions set
    - Refinement approach w.r.t. answers provided by the P.M.
- LLMs:
  - LLMs are not trained for the requirements elicitation task
  - Prompts are the same for all chosen LLMs
  - No domain knowledge related to the System-to-be provided to the LLMs
  - No requirements elicitation method provided to the LLMs

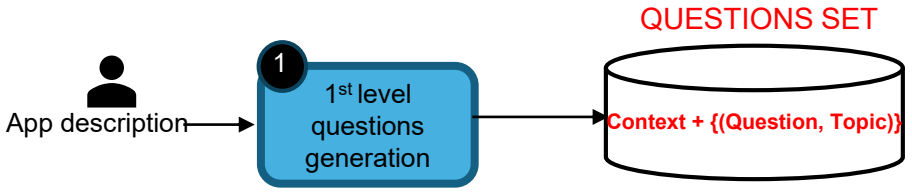
ELISA WORKFLOW - Overview



➢ LLMs selected for experimentations

LLM	Context Window	Nb of parameters
Mistral NeMo	128k tokens	12 billions
GPT-5	400k tokens	–
Qwen3-Max	262k tokens	1 trillion

ELISA WORKFLOW – Question Generation



- Prompt (main request):

> *Your are an expert in requirements Engineering. Generate a list of questions for a Sytem domain provided after <<<. … Evaluate the relevance of each question and explain the scoring. Under the context key, give your understanding about the system.…*

- Question Set attributes:
  - Context: textual description
    - Goals of the Questions generated by the LLM
    - Rationale for the Relevance of the Questions
  - Question: content of the question
  - Topic: Keyword associated with the generated question

ELISA WORKFLOW - Illustration of Questions Generation & Refinement

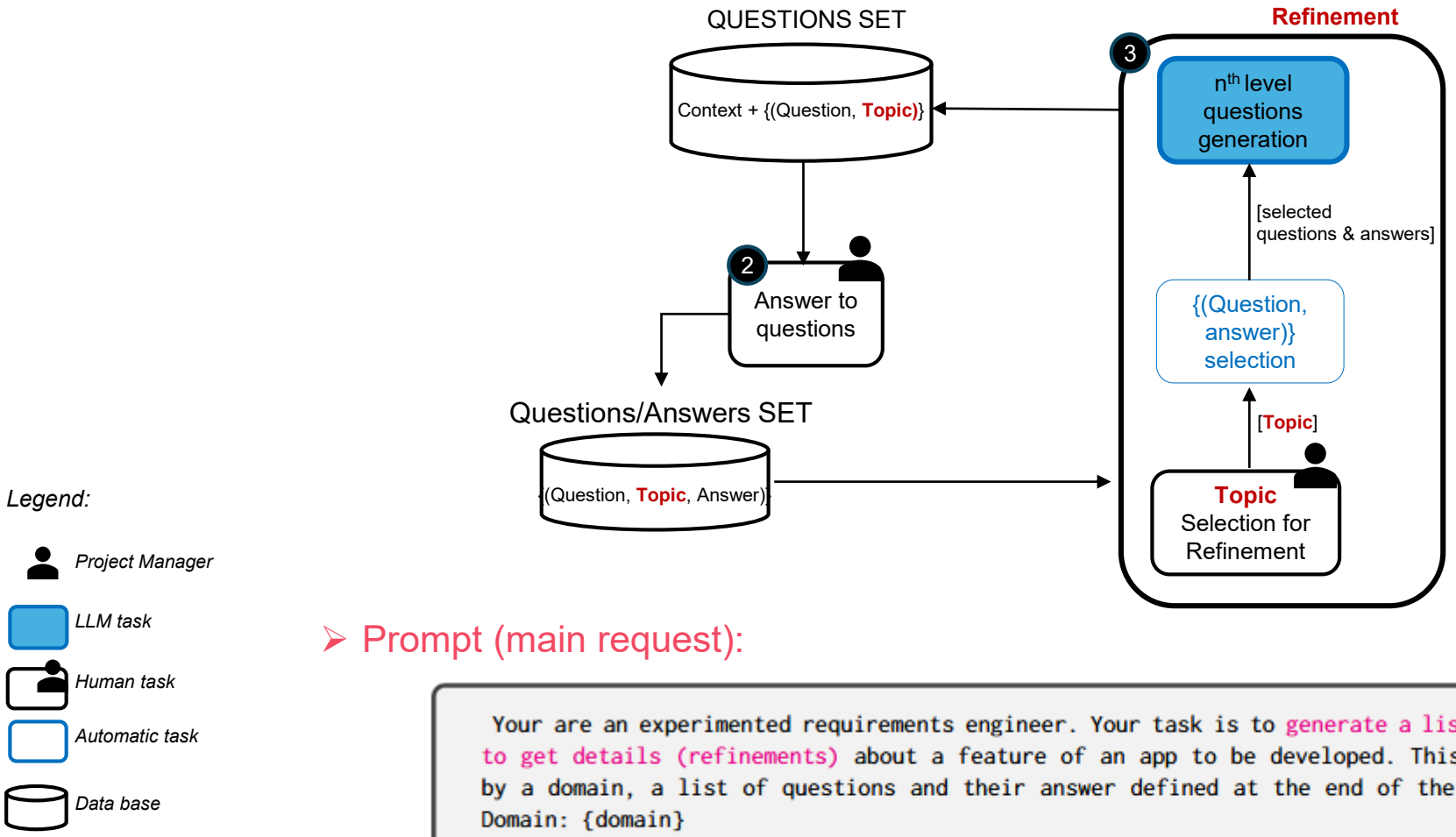
- Telerehab: a telerehabilitation system for COPD patients

Context: The system in question is a Tele Rehabilitation platform designed to monitor and support Chronic Obstructive Pulmonary Disease (COPD) patients. The platform should allow healthcare professionals to remotely monitor patients' health data, provide personalized exercise routines, and facilitate communication between patients and healthcare providers. The system should be user-friendly, secure, and reliable to ensure patient compliance and effective remote management of COPD patients.

Questions	Examples of answers	Topics	Relevance
What are the key user roles in the system?	patients, healthcare providers, administrators	Usability	4
What are the key health data parameters that the system should monitor and track for COPD patients?	SpO2 levels, heart rate, lung function tests, medication adherence, etc.	Functionality	5
What security measures should the system implement to protect patient data?	Data encryption, secure authentication, access controls	Security	5
What are the technical skills and requirements of the users, and how should the system accommodate them?	Users may have limited computer literacy, so the system should be intuitive and easy to use.	Usability	3
What are the system's availability and reliability requirements?	The system should be available 99.9% of the time, with minimal downtime.	Reliability	4

Level-1 Question Generation by Mistral-Nemo (excerpt from 8 questions)

ELISA WORKFLOW – Question Refinement



➢ Prompt (main request):

```
Your are an experimented requirements engineer. Your task is to generate a list of questions in order
to get details (refinements) about a feature of an app to be developed. This feature is determined
by a domain, a list of questions and their answer defined at the end of the prompt after «<. <<<
Domain: {domain}
QuestionAnswer : {questionAnswer} >>>
```

ELISA WORKFLOW - Illustration of Refinement

- Topic : Functionnality

Context: The system is a TeleRehab platform for COPD patients, focusing on providing appropriate exercises and routines, facilitating communication with healthcare providers, and monitoring key health data parameters. To gather detailed requirements, we need to refine the given answers and understand the underlying concepts better. The number of questions generated aims to cover all necessary aspects without being overly exhaustive.

- Functionnality for a telerehabilitation system for COPD patients

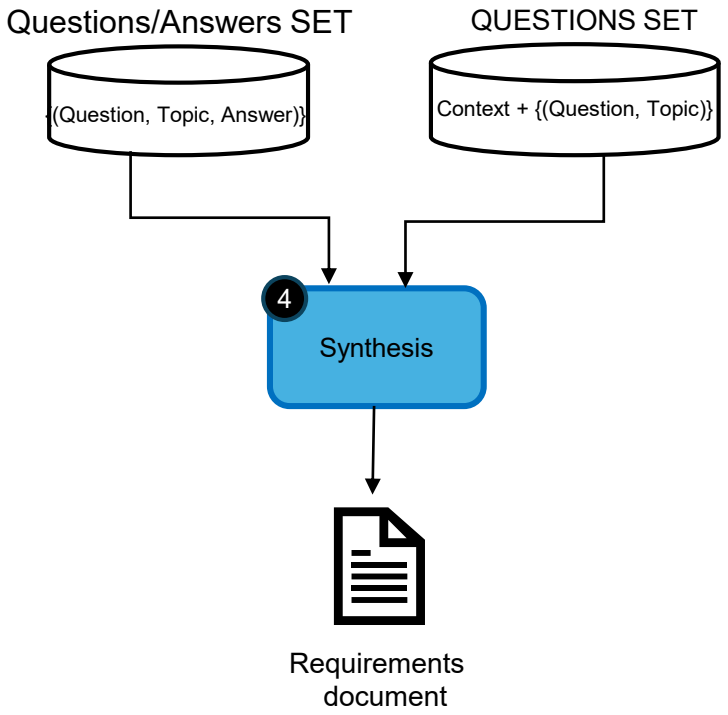
Questions	Examples of answers	Topics	Relevance
What specific physical activities should the system provide for COPD patients under the 'endurance' category?	Walking, Cycling, Swimming	Exercise Type	4
What medical data should the system consider as a precondition for patients to perform 'Walking' as an endurance activity?	FEV1 level, Blood Pressure	Exercice precondition	3
Who should initiate the communication between the patient and the care manager in the TeleRehab system?	Patient, Care Manager	Communication Channel	4

Level-2 Questions generated by Refinement on Topic « Functionnality » by Mistral-Nemo (excerpt from 5 questions)

level-1 Topics	level-2 Topics
Patient Monitoring	Medical Symptom Monitoring Data Collection Frequency Manual Symptom Reporting
Alerting System	Predefined range Expected Progress
Data Security	Data Encryption Data Storage Regulatory Compliance Role base Access Control Dashboard Personalization
Rehabilitation program	Medical Activities Nutrition Activities Physical Activities Adherence tracking Expert role

Generated and RefinedTopics by Qwen3-max

ELISA WORKFLOW – Requirements Document Generation



- Requirements Summary (generated)
- List of concepts associated with Q/A (generated)
- List of generated questions (extracted from data base)
- List of non answered questions (extracted from data base)

ELISA WORKFLOW – Requirements Document Generation Illustration

Requirements Summary (excerpt) generated by Qwen3-max

The telerehabilitation system is designed for monitoring COPD patients and supporting non-emergency rehabilitation under the supervision of healthcare professionals. It must comply with GDPR (Articles 9 and 32), French medical data regulations, and obtain Hébergeur de Données de Santé (HDS) certification.

All patient data must be stored in the EU (AWS eu-west-3, Paris), encrypted at rest using AES-256 with keys managed via AWS KMS, and transmitted over TLS 1.3. Data must be anonymized after 10 years or deleted immediately upon patient request per GDPR Article 17.

The system supports three activity types: Medical (e.g., spirometry, SpO2, heart rate, respiratory rate, blood pressure via connected devices, and weekly mMRC dyspnea scoring), Physical (patient-chosen activities like walking or cycling with evolving intensity), and Nutrition (weight tracking, hunger/satiety questionnaires).

Data collection frequency is activity-specific and adjustable based on patient condition.

Adherence is tracked via questionnaire responses, sensor data, and user actions, with gamification through trophies for completing tasks.Patients can manually report symptoms (mood, anxiety, fatigue, sleep quality, shortness of breath) via a 'Report Symptom' button using visual analog scales. The system explicitly states it is not an emergency tool; patients are notified to contact emergency services externally if needed. No real-time alerts or emergency integrations (e.g., 911) are implemented.

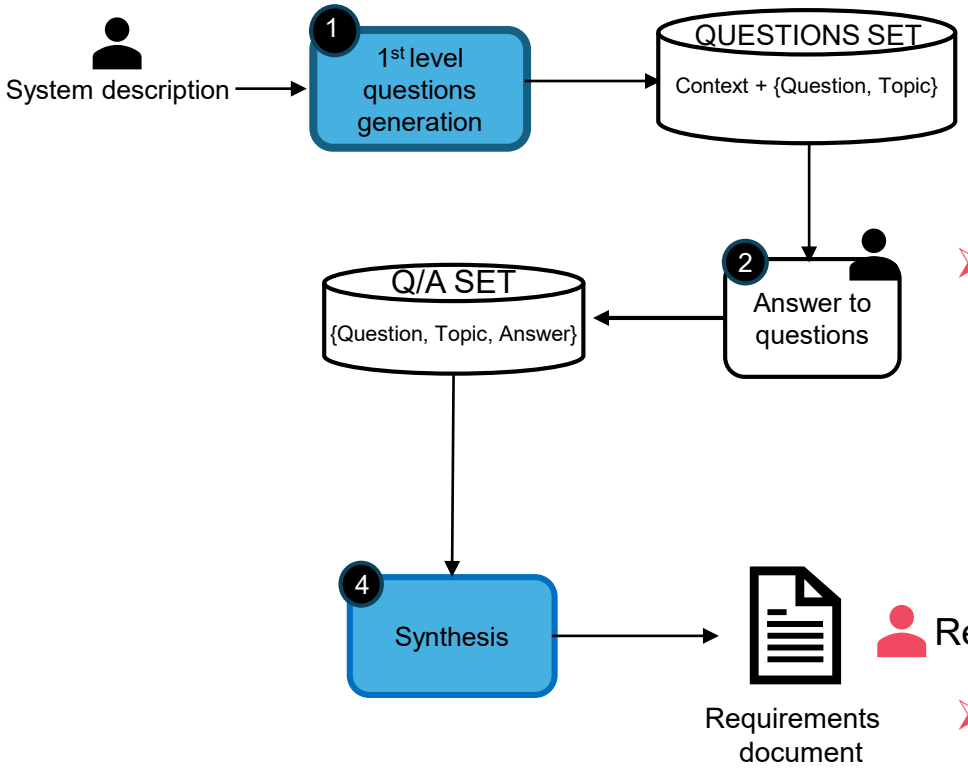
Communication occurs only between patients and a designated Care Manager via secure, end-to-end encrypted messaging (with 90-day retention, PDF/image attachments ≤10MB, and read receipts) or external phone calls (triggering the device's native dialer).

…/…

List of Concepts

- Patient
- CareManager
- Physiotherapist
- Doctor
- MedicalActivity
- PhysicalActivity
- NutritionActivity
- Questionnaire
- SymptomReport
- Dashboard
- Notification
- RehabilitationProgram
- AdherenceTracking
- Gamification
- DataBuffer
- SecureMessaging
- PatientProfile
- Goal
- Metric
- Threshold
- ProgressMilestone
- EncryptionPolicy
- DataRetentionPolicy
- AccessibilitySettings

ELISA LLMs ASSESSMENT



➢ Question Assessment

- ➢ Clarity: is questions formulated clearly and appropriately?
- ➢ Relevance: is the question relevant to the project and the topic?
- ➢ Examples: are provided examples relevant and inspiring?

Not at all	*Slightly*	*Moderately*	*Very*	*Extremely*
○	○	○	○	○
1	2	3	4	5

➢ Requirements document Assessment

- ➢ Are stakeholders identified?
- ➢ Are refinements conducted in a proper way?
- ➢ Is requirements document representative of the interview? What about its completeness?

RESULTS & ANALYSIS

➢ Projects for experiments

- BurnOut, an AI conversational agent to detect risks of burnout.
- Telerehab, a rehabilitation system for patients with chronic respiratory conditions
- GoldenAge, a system of health monitoring of seniors by gerontologists in a controlled environment
- Drug Pathway, a drug management system for hospitals.

➢ Question Assessment

- Clarity: is questions formulated clearly and appropriately?
- Relevance: is the question relevant to the project and the topic?
- Examples: are provided examples relevant and inspiring?

Project	Question Clarity			Question Relevance			Examples Relevance		
	Mistral NeMo	GPT-5	Qwen3-Max	Mistral NeMo	GPT-5	Qwen3-Max	Mistral NeMo	GPT-5	Qwen3-Max
Burnout	5	4.98	5	4.09	4.22	4.73	5	4.95	4.4
Telerehab	5	5	4.9	4.7	4.8	4.9	4.7	5	4.9
GoldenAge	5	4.9	5	4.8	4.43	4.7	4.8	4.8	4.8
Drug Pathway	4.8	4.8	5	3.7	4.2	4.8	4.5	4.9	5
Average	4.95	4.93	4.97	4.3	4.4	4.8	4.85	4.92	4.8

RESULTS & ANALYSIS

➢ Analysis of Refinement and Topics Generation

Project	# level-1 Questions (# Q_1s)			# level-2 Questions (# Q_2s)				
	Mistral NeMo	GPT-5	Qwen3-Max	Mistral NeMo		GPT-5	Qwen3-Max	
				rr_1	#Qs		rr_1	#Qs
Burnout	7	36	12	1.8	15	–	3.6	43
Telerehab	8	32	15	1.7	19	–	2.7	40
GoldenAge	8	42	14	2.2	18	–	1.6	23
Drug Pathway	12	30	15	4	48	–	3.9	59
Average	8.7	35	14	2.9	25	–	2.9	41

Number of questions generated at first and second levels. $rr_{i,\ i\geq 1} = \frac{\#\,Q_{i+1}s}{\#\,Q_i s}$ denotes the *refinement ratio* at level $i+1$

Project	# level-1 Topics									# level-2 Topics								
	Mistral NeMo			GPT-5			Qwen3-Max			Mistral NeMo			GPT-5			Qwen3-Max		
	cr_1	F	NF	cr_1	F	NF	cr_1	F	NF	cr_2	F	NF	cr_2	F	NF	cr_2	F	NF
Burnout	3.5	1	1	1.4	13	13	1	5	7	2.5	3	3	–	–	–	1.1	28	11
Telerehab	2	2	2	2.5	10	3	1.7	5	4	1.1	8	2	–	–	–	1.9	19	12
GoldenAge	1.3	4	2	2.3	15	5	2.8	3	2	1.1	13	4	–	–	–	1.1	16	4
Drug Pathway	1	7	5	1	22	8	1.1	10	4	1.2	27	13	–	–	–	1.3	33	11
Average	1.95			1.79			1.65			1.47			–			1.35		

Number of Functional and Non Functional topics generated at first and second levels. $cr_{i=1,2} = \frac{\#Q_i s}{\#\textit{level-i topics}}$ denotes the *clustering ratio* at level i

➢ Requirements document Assessment

- ➢ Are stakeholders identified?
  - ➢ Good identification of roles and stakeholders for GoldenAge, TeleRehab and DrugPathway Projects
  - ➢ Incomplete identification for Burnout Project
- ➢ Are refinements conducted in a proper way?
  - ➢ Generally very relevant refinement
  - ➢ Some cases in Mistral-Nemo where refinement is partial (features are omitted when response is an enumeration)
- ➢ Is requirements document representative of the interview? What about its completeness?
  - ➢ The generated synthesis is well done, without hallucinations
  - ➢ Completeness requires many level of refinements
  - ➢ Completeness analysis requires requirements formalisation to be discussed by stakeholders and RE experts

CONCLUSION & FUTURE WORK

- Conclusion
  - Comparison of the 3 LLMs
    - GPT-5 asks too many level-1 questions, not very usable with the current prompts
    - Mistral-Nemo is fairly well received, but it is not always comprehensive in terms of the questions asked and topics covered.
    - Qwen-3 is the most appreciated LLM for question relevance and refinement, quality of suggested examples, PM responses integration
  - What is missing
    - Stakeholders identification and relationships with the system
    - Targeted refinement of questions to adress questions to identified stakeholders
    - Collecting knowledge about:
      - Environment of the system, assumptions and constraints
      - Limits of the System-to-be (what it is not supposed to do)
    - Requirements classification
- Future work
  - Experiments with prompts supporting a Requirements Elicitation method
  - Multi-modal requirements generation : textual, and model-based (class diagram, goal model)

LOOKING FOR ASSESSORS…

- https://factory.euromov.eu/elisa
  - Ask a user id and a password to: anne-lise.courbis@mines-ales.fr
  - Send your feed-back to: anne-lise.courbis@mines-ales.fr
- New research idea: setting a group of Requirements Experts to define prompts & IA based Requirements Elicitation System…