

pipeline
d'extraction
pour la
simplification
des données

```
graph LR; A[pipeline d'extraction pour la simplification des données] --> B[Trouver des données de publications scientifiques présentant des résultats probabiliste p.ex études médecines estimation tu taux de décès en fonction de l'excès de la vitesse, tabagisme alliant facteur,]; B --> C[Parvenir à extraire les données de la publication]; C --> D[optionnel(Vérifier la fiabilité des données avec des test de validité p.ex shapiro etc)]; D --> E[Représenter en suivant une méthodologie Gigerenzer par exemple "facts boxs".]; D --> F[Utiliser un LLM pour expliquer les données.];
```

Trouver des données de publications scientifiques présentant des résultats probabiliste p.ex études médecines estimation tu taux de décès en fonction de l'excès de la vitesse, tabagisme alliant facteur,

Parvenir à extraire les données de la publication

optionnel(Vérifier la fiabilité des données avec des test de validité p.ex shapiro etc)

Représenter en suivant une méthodologie Gigerenzer par exemple "facts boxs".

Utiliser un LLM pour expliquer les données.

Quelles approches permettent
de faciliter la compréhension
des risques probabilistes et
statistiques pour les personnes
ordinaires ?



Facts box
de
Gigerenzer

Représentation
graphiques

Assistant
conversationnel

User stories

Quelles
méthodes de risk
literacy existent
pour clarifier des
données?

Graphique
simple

Open data:
smoking and risk factors
: <https://ourworldindata.org/smoking>
Risk literacy
: <https://www.hardingcenter.de/en/transfer-and-impact/fact-boxes>
Flight statistics
: https://www.panish.law/aviation_accident_statistics.html

Source pour la validation des
processus
<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0321153>

lien pour des données sur
la littératie du risque
<https://www.research-collection.ethz.ch/entities/researchdata/3e5bc4d2-7db6-4605-b7a4-da11bbe1210b>

Source for graphic warning on
cigarette packages, go to page 10
for the
table: <https://pmc.ncbi.nlm.nih.gov/articles/PMC3377985/pdf/nihms378783.pdf>

Source for the effects in BEHAVIOR
based on graphic warning on
cigarette
packages: <https://pmc.ncbi.nlm.nih.gov/articles/PMC5458743/pdf/nihms858881.pdf>

DataBase

FactsBox
Generator

AI Chatbot
(LLM)

User Stories
Generator

Visual Graph
Generator



User Input (txt)



Intent & Context Analyzer



Fetch FactsBox (based on intent)



FactsBox Interpreter
(based on intent)



Prompt Builder + LLM



Chatbot Answer (Txt)

User Stories Generator

Pipeline pour
mesurer l'efficacité
d'une technique de
"risk literacy" -> En
gros est-ce que la
technique utilisée
fonctionne.

Est-ce que des
test sur la
literacy
existent déjà?

si oui peut
on les
utiliser
(open data?)

Si non
définir un
protocole
de test.

idée globale des test
présenter des données(les
mêmes, mais sous une
manière simplifiée VS sous
une forme scientifique
pure opaque.

comparer les
résultats->analy
ses statistiques
ou IA p.ex arbre
de décision.



LLM generated
scenarios (Txt)

Openai API

User Stories in PNG





