

್ಷ



Empowering Users for Secure Data Release and Misinformation Detection: Insights from KURAMi

Marco Viviani

University of Milano-Bicocca

Department of Informatics, Systems, and Communication (DISCo)

April 30th, 2025





Motivation and Context

KURAMi, Marco Viviani, April 30, 2025 – Universitat Politècnica de València





Challenges

- Users may face **various risks** in releasing and accessing content (structured, semistructured, unstructured) in online environments.
- **Content release**: Uncontrolled release of personal/sensitive data (privacy).
 - How to protect privacy?
 - How to avoid microtargeting?

- Content access: Access to "incomplete"/fake information.
 - How to identify the utility of information protected from a privacy perspective?
 - How to avoid misinformation access?





Putting the User at the Center

• In the **trade-off** between releasing personal/sensitive data and accessing useful/reliable information, **users must play a central role**.



- Provide users with automated and effective approaches promoting user autonomy.
- Easily interpretable results without the decision-making process being left only to algorithms.







- KURAMi: Knowledge-based, explainable User empowerment in Releasing private data and Assessing Misinformation in online environments.
- **PRIN 2022**: Research project funded by the Italian EU -Next Generation EU, Mission 4, Component 2, CUP D53D23008480001 and Italian MUR.





Ministero dell'Università e della Ricerca







KURAMi and Privacy: Some Tasks

- Various tasks are involved in KURAMi.
- In today's **seminar**:
 - Privacy Awareness
 - Document sanitization \rightarrow Data Marketplaces.
 - Query sanitization \rightarrow • Generative IR.
 - Misinformation Awareness
 - RAG-based Health IR



KURAMI-ENRICHED ONLINE ENVIRONMENT





Document Sanitization

Cassani, L., Livraga, G., & Viviani, M. (2024, September). Assessing document sanitization for controlled information release and retrieval in data marketplaces. In International Conference of the Cross-Language Evaluation Forum for European Languages (CLEF 2024) (pp. 88-99). Cham: Springer Nature Switzerland.





Preamble: Text Analysis and Privacy

- Step 1: Identifying sensitive entities using Named Entity Recognition (NER) and other NLP methods.
 - Common entities include: personal data like names, addresses, emails, phone numbers, etc., including sensitive data like health conditions, political or religious affiliations, financial information, and other metadata like timestamps, geolocation, etc.
- Step 2: Assigning risk scores to entities based on:
 - Entity sensitivity: Certain entities (e.g., health conditions) are inherently more sensitive than others.
 - Uniqueness: Evaluates how rare and identifiable an entity is.
 - Exposure: The probability of exposure due to attacks or misuse.
- Step 3: Aggregating risk scores by aggregating the risk scores of individual entities, often using weighted sums, averages, or maximum-based aggregation.





The Context: Data Marketplaces

- **Data Marketplaces** (DMs) are specialized virtual spaces that allow the exchange of various kinds of data that can range from highly specific and niche data to more general and broadly applicable information.
 - **Data owners** offer them for a fee on a DM.
 - **Registered users** can explore the platform to retrieve the data they need and, should they find data of interest, proceed with the purchase.
 - **DMs generate revenue** usually through commissions from processed transactions.
- In **marketplaces** for **physical items**, products can be presented with accurate descriptions and photographs, and are subject to return and warranty policies.





Open Issues and Possible Solutions

- **Digital information** presents different characteristics by its nature.
 - **Data stored** within DM platforms must be protected so that they are only visible to users who have purchased them.
 - These platforms must also equip **potential buyers** with the tools needed to determine whether the data they find is indeed useful for them, without exposing the entire content before the sale is concluded.
- Modern DMs also include unstructured data.
 - The objective of providing an accurate description remains the same.
 - Need for tailored strategies (*blurring* for images, *key frames* for videos).
 - What about textual documents? \rightarrow **Text sanitization**.





Text Sanitization

- **ALERT**: The sanitized text should:
 - **Protect** the content not meant for disclosure.
 - Be sufficiently representative of the original text → Sufficiently match the buyer's information needs.
- A twofold objective:
 - Various sanitization techniques applied to textual documents within the DM context → Masking and/or summarization.
 - Assessing retrieval effectiveness of sanitized documents to verify that data sanitization, while concealing confidential content, compromises neither retrieval effectiveness nor data saleability.





The Proposed Architecture





Document Masking and Summarization

• **Document masking**: selectively masking parts of the document (in terms of tokens) as deemed necessary by the owner.

ReGAInS

Dipartimento di

matica. Sistemistica

- E.g., if the owner does not wish to include the word 'cat' in a masked document, the original document can be sanitized by masking all occurrences of the token 'cat'.
- We can apply or not Coreference Resolution (CR) → <u>Next slide</u>.

- **Document summarization**: generating a summary of a document.
 - Keeping just the most important sentences in the summary, i.e., **extractive summarization**.
 - Rephrasing the original documents in a shorter version, i.e., abstractive summarization.
 - Extractive summarization preserves the original document's representativeness by including original sentences in the summarized document.





Coreference Resolution

• **Coreference Resolution (CR)** is the task of finding all linguistic expressions (called mentions) in a given text that refer to the same real-world entity.

CR

• The mouse and the elephant are two animals, belonging to the class of mammals. The former has an average weight of 20 g, while the latter can weigh up to 6,000 kg. In addition, the latter, unlike the former, has a proboscis. • The mouse and the elephant are two animals, belonging to the class of mammals. The mouse has an average weight of 20 g, while the elephant can weigh up to 6,000 kg. In addition, the elephant, unlike the mouse, has a proboscis.





Document Masking and CR

• [MASK] and [MASK] are two animals, belonging to the class of mammals. The former has an average weight of 20 g, while the latter can weigh up to 6,000 kg. In addition, the latter, unlike the former, has a proboscis.

The latter \rightarrow



 [MASK] and [MASK] are two animals, belonging to the class of mammals. [MASK] has an average weight of 20 g, while [MASK] can weigh up to 6,000 kg. In addition, [MASK], unlike [MASK], has a proboscis.

(?)







Confidentiality Risk Assessment

- We intend the **confidentiality risk** as the possibility of demasking tokens that have been obfuscated by the data owner.
- Assessed by means of a **demasking resistance measure**:

$$dr(d) = 1 - \frac{n_{inf}}{n_{max}}$$

- n_{inf} : the number of inferred tokens from the sanitized document.
- n_{max} : the total number of obfuscated tokens in the sanitized document.





The «Online News» Scenario

- Showing only the title or the first portion of an article may not be the best choice for a customer interested in purchasing the article itself.
- **Data**: a subset of the articles from the Washington Post collected as part of TREC.
 - The collection includes 595,037 articles, stored in a JSON Lines format file, collected around 50 different topics.
 - A qrels.txt file is also provided for performance evaluation in IR.
 - Only documents with a length of less than 512 tokens (a limit imposed by BERT) were considered for evaluation \rightarrow 3,776 articles.





Implementing the Solution

• Summarization techniques:

- Luhn,
- KLSummarizer,
- Latent Semantic Analysis (LSA),
- LexRank,
- SBertSummarizer.
- Masking assumption: tokens to be obfuscated are entities in the original documents.
 - Those extracted by means of Named-Entity Recognition (NER).

- **Demasking**: performed using the DistilRoBERTa model.
 - LLMs can be employed to infer masked tokens.
- Retrieval models:
 - TF-IDF, BM25, DLH, DPH, InL2, MDL2.
- Metrics:
 - Average demasking resistance $\rightarrow adr(D)$,
 - Mean Average Precision \rightarrow MAP,
 - Normalized Discounted Cumulative Gain \rightarrow nDCG.





Some Results: Masking Alone

 Table 1. Evaluation metrics considering document masking for sanitization.

Model	MAP_{bl}	$nDCG_{bl}$	Masked	Demasked	adr(D)	MAP	nDCG
TF-IDF	0.234	0.411	41,816	8,867	0.788	0.211	0.386
BM25	0.234	0.411	41,816	8,867	0.788	0.212	0.386
DLH	0.226	0.403	41,816	8,867	0.788	0.204	0.38
DPH	0.249	0.423	41,816	8,867	0.788	0.220	0.393
InL2	0.238	0.413	41,816	8,867	0.788	0.216	0.389
MDL2	0.201	0.373	41,816	8,867	0.788	0.183	0.357





Some Results: Summarization + Masking

Table 2. Evaluation metrics considering summarization and masking for sanitization(average over 8 summary lengths; IR model: DPH).

Model	$adr(D)_{as}$	MAP_{as}	nDCG _{as}
Luhn	0.884	0.194	0.357
KLSummarizer	0.918	0.174	0.339
Latent Semantic Analysis (LSA)	0.885	0.205	0.374
LexRank	0.891	0.195	0.357
TextRank	0.877	0.204	0.371
SBertSummarizer	0.899	0.184	0.351

Average scores obtained over distinct document summary lengths (i.e., 10%, 20%, ..., 80%) for the considered evaluation metrics, denoted in this case as $adr(D)_{as}$, MAP_{as}, and nDCG_{as}.





Some Results: Summariz. + Masking + CR + QE

Table 3. Evaluation metrics considering summarization and masking for sanitization, CR, and QE (average over 8 summary lengths; IR model: DPH; summarizer: LSA).

CR	QE	$adr(D)_{as}$	MAP_{as}	nDCG _{as}
No	No	0.885	0.205	0.374
Yes	No	0.892	0.199	0.366
No	KLQE	0.885	0.222	0.416
Yes	KLQE	0.892	0.214	0.405





Some Takeaways

- Simple token masking alone is less effective at mitigating the risk of demasking compared to the combination of token masking with text summarization.
- While improving confidentiality, this approach **negatively impacts retrieval effectiveness**.
 - A balanced approach can be achieved by incorporating Coreference Resolution during the masking process and employing Query Expansion during retrieval.
- Further research (some ideas):
 - More sophisticated summarization algorithms that inherently incorporate data confidentiality principles could be developed.
 - Conducting comprehensive testing across various marketplace scenarios and datasets could validate the applicability and resilience of our approach.





Query Sanitization

Herranz-Celotti, L., Guembe, B., Livraga, G., & Viviani, M. (2025, April). Can Generative Al Adequately Protect Queries? Analyzing the Trade-off Between Privacy Awareness and Retrieval Effectiveness. In Proceedings of the 47th European Conference on Information Retrieval (ECIR 2025), Lucca, Italy, 2025





The Context: Generative IR

- Users increasingly interact with generative AI tools and Information Retrieval Systems → Generative Information Retrieval.
- Queries have longer and more elaborate prompts → Likely containing more detailed and personal/sensitive information.
- Useful for users, yet risky for privacy → Several approaches for query protection, both security/privacy and IR research fields.





Background

- Query protection can entail protecting:
 - User identity (break link between identity and query).
 - User intent (blur query content).
 - Focus on query intent protection \rightarrow <u>Queries are modified before being sent to the IRS</u>.
- **Protection** is typically **enforced on the user side** (no intervention from not fully trusted IRSs).
 - Query modification:
 - Generalization: lexical resources (e.g., WordNet) or heuristics to remove private/sensitive intent details.
 - Noise addition (e.g., DP-based text generation).
 - Query multiplication: confuse intent with dummy queries.





The Investigated Open Issues

- **Question 1**: Can generative LLMs mimic query protection approaches?
- **Question 2**: Can generative LLMs offer a viable solution for balancing user privacy with retrieval effectiveness?
- **Approach**: Study the feasibility of applying query protection through generative LLMs.
 - Six different prompts.
 - Inspired by SOTA confusion-based methods (query modification and multiplication).





Prompt Engineering

- Let us consider the following **Original Query (OQ)**.
- Let us consider the following **Preamble (P)**.

OQ. gay clubs in New York

PB. You are a privacy expert tasked with protecting the user's privacy when performing the following task.





Prompt 1 (P1): Simple Rewriting

OQ. gay clubs in New York

PB. You are a privacy expert tasked with protecting the user's privacy when performing the following task.

P1. [PB] Rewrite this query: [OQ]

E1. Social venues in New York for LGBTQ+ community





Prompt 2 (P2): Generalization

OQ. gay clubs in New York

PB. You are a privacy expert tasked with protecting the user's privacy when performing the following task.

P2. [PB] Rewrite this query by applying generalization: [OQ]

E2. Entertainment venues for diverse communities in urban areas





Prompt 3 (P3): Differential Privacy*

OQ. gay clubs in New York

PB. You are a privacy expert tasked with protecting the user's privacy when performing the following task.

P3. [PB] Rewrite this query by applying Differential Privacy: [OQ]

E3. Gay clubs in New York, Los Angeles, and Chicago

*With a caveat





Prompt 4 (P4): Dummy Queries

OQ. gay clubs in New York

PB. You are a privacy expert tasked with protecting the user's privacy when performing the following task.

P4. [PB] Generate [k] dummy, random queries, given this query: [OQ]

E4 (k = 3). Art galleries to visit in urban settings | Cultural festivals happening in the summer | Best coffee shops with outdoor seating





Prompt 5 (P5): Dummy Queries + Semantics

OQ. gay clubs in New York

PB. You are a privacy expert tasked with protecting the user's privacy when performing the following task.

P5. [PB] Generate [k] dummy queries, which are semantically related to this query: [OQ]

E5 (k = 3). LGBTQ+ events happening in New York City | Nightlife options for the LGBTQ+ community in urban areas | Social gatherings for LGBTQ+ individuals in major cities





Prompt 6 (P6): Dummy Queries + Generalization*

OQ. gay clubs in New York

PB. You are a privacy expert tasked with protecting the user's privacy when performing the following task.

P6. [PB] Generate [k] dummy queries, which generalize this query: [OQ]
E6 (k = 3). LGBTQ+ nightlife options in major cities | Social venues for diverse communities in urban areas | Inclusive entertainment spots in metropolitan regions

*With a caveat





Implementing the Solution

- **Goal**: compare prompt-driven LLM methods with SOTA baselines.
 - Lexicon-based (WordNet)
 - Differential Privacy-based
- Different retrieval models:
 - Sparse (BM25).
 - Dense (MonoT5).

• Datasets:

- NFCorpus (medical IR).
- TREC-COVID (pandemic-related research).
- Touché (controversial topics).
- Metrics:
 - Retrieval effectiveness (MAP, nDCG).
 - Query syntactic (Jaccard index) and semantic (cosine similarity among BERT embeddings) similarity.





Best Results for Sparse Retrieval

PMC MAP DCG ₁₀ DCG ₀₀ CS _{B4} JI MAP DCG ₁₀ DCG ₁₀ CS _{B4} JI MAP DCG ₁₀ CS _{B4} JI NONE 0.032 0.201 0.033 0.123 0.111 0.615 0.200 0.000 <th></th>																
$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$			1	NFCorpus				TI	REC-Covid					$Touch \acute{e}$		
$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$	$\mathbf{Q}\mathbf{M}$	MAP	$nDCG_{10}$	$nDCG_{100}$	$\mathbf{CS}_{\mathbf{B}}\downarrow$	$\mathbf{JI}\!\!\downarrow$	$ \mathbf{MAP} $	$nDCG_{10}$	$nDCG_{100}$	$\mathbf{CS_B}{\downarrow}$	$\mathbf{JI}{\downarrow}$	MAP	$nDCG_{10}$	$nDCG_{100}$	$\mathbf{CS}_{\mathbf{B}}\downarrow$	JI↓
$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$	NONE	0.149	0.322	0.273	-	-	0.198	0.626	0.474	-	-	0.225	0.343	0.455	-	-
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $	WordNet	0.057	0.120	0.114	0.687	0.201	0.033	0.123	0.111	0.615	0.209	0.019	0.027	0.065	0.643	0.153
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $	DP \mathbf{CMP}_1	0.000	0.002	0.001	0.416	0.000	0.000	0.000	0.000	0.347	0.000	0.000	0.000	0.000	0.269	0.000
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $	$\mathbf{DP} \ \mathbf{CMP}_5$	0.001	0.003	0.004	0.430	0.005	0.000	0.000	0.000	0.366	0.000	0.000	0.000	0.000	0.277	0.000
$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$	$\mathbf{DP} \ \mathbf{CMP}_{10}$	0.035	0.075	0.075	0.563	0.166	0.011	0.025	0.027	0.448	0.067	0.024	0.033	0.065	0.426	0.119
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $	$DP CMP_{50}$	0.149	0.322	0.273	1.000	0.999	0.182	0.573	0.438	0.984	0.784	0.225	0.343	0.455	1.000	1.000
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $	$\mathbf{DP} \ \mathbf{M}_1$	0.000	0.001	0.002	0.398	0.000	0.000	0.000	0.001	0.352	0.000	0.000	0.000	0.000	0.274	0.000
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $	$\mathbf{DP} \ \mathbf{M}_5$	0.001	0.002	0.003	0.411	<u>0.002</u>	0.000	0.002	0.001	0.366	0.000	0.000	0.002	0.002	0.274	0.002
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $	$\mathbf{DP} \ \mathbf{M}_{10}$	0.019	0.047	0.048	0.497	0.106	0.012	0.025	0.035	0.420	0.035	0.008	0.004	0.032	0.387	0.069
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $	$\mathbf{DP} \ \mathbf{M}_{50}$	0.149	0.322	0.273	1.000	0.999	0.183	0.576	0.440	0.985	0.784	0.225	0.343	0.455	1.000	1.000
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $	$\mathbf{DP} \ \mathbf{V}_1$	0.000	0.001	0.001	0.404	0.000	0.000	0.000	0.000	0.346	0.000	0.000	0.000	0.000	0.272	0.003
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $	$\mathbf{DP} \ \mathbf{V}_5$	0.002	0.004	0.007	0.412	0.002	0.000	0.001	0.002	0.348	0.002	0.002	0.003	0.005	0.289	0.002
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	$\mathbf{DP} \ \mathbf{V}_{10}$	0.017	0.044	0.046	0.490	0.067	0.020	0.014	0.043	0.412	0.046	0.018	0.020	0.049	0.373	0.051
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	$\mathbf{DP} \ \mathbf{V}_{50}$	0.094	0.209	0.190	0.814	0.466	0.098	0.321	0.276	0.761	0.350	0.131	0.206	0.301	0.791	0.471
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	\mathbf{GPT}_{P1}	0.092	0.202	0.188	0.696	0.291	0.098	0.337	0.265	0.893	0.424	0.145	0.240	0.344	0.879	0.427
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	\mathbf{GPT}_{P2}	0.049	0.114	0.129	0.683	0.127	0.047	0.205	0.156	0.794	0.316	0.044	0.068	0.138	0.769	0.198
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	\mathbf{GPT}_{P3}	0.093	0.194	0.188	0.597	0.196	0.078	0.320	0.223	0.747	0.289	0.110	0.174	0.282	0.800	0.296
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	$\mathbf{GPT}_{P4}^{k=1}$	0.122	0.250	0.240	0.571	0.371	0.152	0.402	0.358	0.693	0.541	0.179	0.251	0.368	0.724	0.480
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	$\mathbf{GPT}_{P4}^{k=3}$	0.146	0.299	0.271	0.548	0.205	0.121	0.333	0.291	0.677	0.304	0.191	0.283	0.394	0.731	0.253
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	$\mathbf{GPT}_{P4}^{k=5}$	0.158	0.327	0.288	0.536	0.146	0.170	0.425	0.365	0.757	0.261	0.226	0.325	0.452	0.766	0.193
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	$\mathbf{GPT}_{P5}^{k=1}$	0.144	0.307	0.274	0.759	0.424	0.167	0.519	0.406	0.871	0.566	0.200	0.300	0.419	0.890	0.516
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	$\mathbf{GPT}_{P5}^{k=3}$	0.157	0.319	0.290	0.625	0.196	0.166	0.476	0.391	0.786	0.299	0.220	0.328	0.444	0.799	0.242
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	$\mathbf{GPT}_{P5}^{k=5}$	0.163	0.333	0.299	0.610	0.155	0.173	0.473	0.397	0.781	0.230	0.225	0.339	0.452	0.774	0.183
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	$\mathbf{GPT}_{P6}^{k=1}$	0.148	0.305	0.280	0.798	0.451	0.156	0.547	0.391	0.865	0.588	0.185	0.292	0.402	0.894	0.522
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	$\mathbf{GPT}_{P6}^{k=3}$	0.143	0.293	0.272	0.682	0.214	0.131	0.411	0.338	0.796	0.286	0.205	0.316	0.430	0.817	0.247
	$\mathbf{GPT}_{P6}^{k=5}$	0.141	0.291	0.272	0.651	0.167	0.148	0.416	0.353	0.793	0.239	0.206	0.328	0.435	0.788	0.180





Some Takeaways

- Lexicon-based SOTA achieves privacy protection in spite of retrieval effectiveness.
- **DP methods** achieve reasonable effectiveness with epsilon values too high (~50).
- LLM-based query multiplication seems to balance protection and retrieval effectiveness → "Query expansion" effect?
- **LLM-based methods** tend to perform better in sparse retrieval \rightarrow To be investigated.
- By observing the **queries generated**, it seems that the LLM perceives the protection mechanism as a blurring of the query with more general terms (unless explicitly instructed otherwise).





RAG-Based Health Information Retrieval

Upadhyay, R., & Viviani, M. (2025). Enhancing Health Information Retrieval with RAG by prioritizing topical relevance and factual accuracy. Discover Computing, 28(1), 27.





What is Retrieval Augmented Generation?

- The idea behind RAG techniques is to make use of knowledge "outside" the model to provide a "local" context (in-context) that can supplement the model with appropriate knowledge without changing its parameters.
- These are basically prompting techniques that supplement the user's input with **contextual knowledge retrieved** by accessing external sources of information through a search engine.





The Proposed Solution

- Integrating generative LLMs with a reputed, external knowledge, such as the curated scientific repository of PubMed Central (PMC), a strategy designed to increase both the topical relevance and factual accuracy of the retrieved documents
- The proposed solution is characterized by three key stages:
 - User query-based passage retrieval from PMC
 - GenText generation through LLMs
 - Calculating topicality and factual accuracy, and final document ranking





The Proposed Solution: Pipeline

Step 1: User Query-Based Passage Retrieval from PMC



Step 2: GenText Generation through LLMs



Step 3: Calculating Topicality and Factual Accuracy, and Final Document Ranking







GenText generation through LLMs

LLM prompt

Query: can 5g antennas cause covid 19

Context: People around me told me not to get vaccinated against COVID-19 and reason 12 5G antennas are linked to the COVID-19 pandemic. At the same time there was no statistically significant difference in the average values of their answers regarding these reasons (Reference: 10316077). Interference can have a significant impact on 5G networks particularly in the context of Internet of Things IoT devices. (Reference: 10144169) These measures ensure that user privacy is protected and 5G networks can be trusted to handle massive data securely. The main causes and consequences of these challenges are summarized in Table 10 (Reference: 10255561). The need to deal with the explosion of multimedia services has been considered in the 6G network which will provide greater QoS while also guaranteeing QoE (Reference: 10347022). The importance of this was well proven in pandemic conditions of Covid-19 2729 So that in most organizations employees used different communication networks to do their work and after that the scope of communication networks in organizations has always grown. (Reference: 10399785) Newly emerging variants of SARS-CoV-2 continue to pose a significant threat to global public health by causing COVID-19 epidemics (Reference: 10288941). 4 GHz transmit signal filters and amplifies the received signal and downconverts it to IF and then digitizes the signal according to the programmed parameter settings on the laptop (Reference: 9953371). These tags incorporate antennas that can collect power efficiency regarding radio frequency queries from the RFID transceiver. Software is a good platform and crosses successful evolution for the sensor in Nanotechnology and bio-industries. In the medical field the Healthcare system is usually used to monitor the condition of patients (Reference: 10258751). The controller of the level crossing must be equipped with antennas able to receive signals from the sensors in the train in a safe distance in the paper we suggested possible technologies of communication. The realworld application can be a mixed solution (Reference: 10384084). The presence of various manufacturers and the diverse applications of sensors in disaster scenarios contribute to the heterogeneity of these sensors hence hindering the integration and sharing of information 107108. Some disasters may cause sequent disasters. For example, seismic activity or inundations can cause floods (Reference: 10490738).

Write a paragraph answering the query based on the context provided above constituted by ONLY 64 words, with references for each sentence with (Reference:...).

Do not use extra knowledge.

Based on the context provided, there is a misconception linking 5G antennas to the COVID-19 pandemic (Reference: 10316077). However, this connection has no statistically significant evidence to support it (Reference: 10316077). Instead, it's important to note that 5G networks play a crucial role in ensuring secure data handling and enhancing user privacy (Reference: 10255561). Moreover, SARS-CoV-2 variants remain the main cause of COVID-19 outbreaks (Reference: 10288941).





Topicality, Factual Accuracy, and Final Ranking

• The **topicality score** *T* is computed for each document *d* in the document collection, and it is derived using the BM25 retrieval model against the considered query *q*

T(d,q) = BM25(d,q)

• The **factual accuracy score** *F* integrates two distinct metrics to assess the adherence of documents to be retrieved to scientific evidence within GenText *G*

 $F(d,G) = \alpha \cdot stance(d,G) + (1-\alpha) \cdot \cos(d,G)$

• The **final document ranking** is obtained by performing a linear combination of topicality and factual accuracy scores in order to obtain the Retrieval Status Value (RSV)

 $RSV(d,q,G) = \beta \cdot T(d,q) + (1-\beta) \cdot F(d,g)$





Some Results

CLEF eHealth 2020 dataset

TREC HM 2020 dataset

Model	CAM _{MAP}		Embeddings						
Top-5 Documents									
BM25	0.0431	0.1045	-						
DigiLab	0.0433	0.1109	-						
CITIUS	0.0455	0.1119	-						
WISE	0.0611	0.1198	BioBERT						
WISE _{NLI}	0.0883	0.1823	BioBERT						
GPT _{RAG}	0.1045	0.2098	BioBERT						
Llama _{RAG}	0.1079	0.2146	BioBERT						
Falcon _{RAG}	0.0994	0.2011	BioBERT						
	Top-10 D	ocuments							
BM25	0.0784	0.1923	-						
DigiLab	0.0823	0.1992	-						
CITIUS	0.0843	0.1999	-						
WISE	0.1102	0.211	BioBERT						
WISE _{NLI}	0.1302	0.2321	BioBERT						
GPT _{RAG}	0.1502	0.2655	BioBERT						
Llama _{RAG}	0.1532	0.2702	BioBERT						
Falcon _{RAG}	0.1495	0.2568	BioBERT						

Model	CAM _{MAP}		Embeddings						
Top-5 Documents									
BM25	0.0631	0.1435	_						
DigiLab	0.0712	0.1543	-						
CITIUS	0.0754	0.1554	-						
WISE	0.0844	0.1608	BioBERT						
WISE _{NLI}	0.0923	0.1922	BioBERT						
GPT _{RAG}	0.1178	0.2234	BioBERT						
Llama _{RAG}	0.1222	0.2298	BioBERT						
Falcon _{RAG}	0.1123	0.2165	BioBERT						
	Top-10 D	ocuments							
BM25	0.1047	0.2052	_						
DigiLab	0.1186	0.2011	-						
CITIUS	0.1194	0.2095	-						
WISE	0.1233	0.22	BioBERT						
WISE _{NLI}	0.1341	0.2455	BioBERT						
GPT _{RAG}	0.1547	0.2712	BioBERT						
Llama _{RAG}	0.1602	0.2723	BioBERT						
Falcon _{RAG}	0.1501	0.2665	BioBERT						





A Tool for Explainability?

Can 5G antennas cause COVID-19?

Search

Search Results for "Can 5G antennas cause COVID-19"

LLM Generated Text

Based on the context provided, there is a misconception linking 5G antennas to the COVID-19 pandemic (Reference: 10316077). However, this connection has no statistically significant evidence to support it (Reference: 10316077). Instead, it's important to note that 5G networks play a crucial role in ensuring secure data handling and enhancing user privacy (Reference: 10255561). Moreover, SARS-CoV-2 variants remain the main cause of COVID-19 outbreaks (Reference: 10288941).

Reference List

- 10316077 Softić, Adaleta, Elma Omeragić, Martin Kondža, Nahida Srabović, Aida Smajlović, Esmeralda Dautović, Nataša Bubić Pajić et al.
 "Knowledge and Attitudes regarding Covid-19 Vaccination among Medical and Non-medical Students in Bosnia and Herzegovina." Acta Medica Academica 52, no. 1 (2023): 1.
- 10255561 Ullah, Yasir, Mardeni Bin Roslee, Sufian Mousa Mitani, Sajjad Ahmad Khan, and Mohamad Huzaimy Jusoh. "A survey on handover and mobility management in 5G HetNets: current state, challenges, and future directions." Sensors 23, no. 11 (2023): 5081.
- 10288941 Soto, Ismael, Raul Zamorano-Illanes, Raimundo Becerra, Pablo Palacios Játiva, Cesar A. Azurdia-Meza, Wilson Alavia, Verónica García, Muhammad Ijaz, and David Zabala-Blanco. "A new COVID-19 detection method based on CSK/QAM visible light communication and machine learning." Sensors 23, no. 3 (2023): 1533.

Search Results

The conspiracy of Covid-19 and 5G: Spatial analysis

Conspiracy theories in general carry potentially serious public health risks, especially as anti-vaccination beliefs are already found to be

Evidence for a connection between coronavirus disease and 5G

We explore the scientific evidence suggesting a possible relationship between COVID-19 and radiofrequency radiation related to wireless...

5G Doesn't Cause COVID-19, But the Rumor It Does

People's fear of 5G technology is rational. Such technology does emit radiation, even if it's at low levels. But 5G isn't all that different ...

How the 5G coronavirus conspiracy theory went from fringe to mainstream

Despite what the internet might be telling you, cellphones did not cause the Covid-19 pandemic.

Is there a connection between coronavirus and 5G?

"5G mobile networks DO NOT spread COVID-19: viruses cannot travel on radio waves/mobile networks. COVID-19 is spreading in many countries that do





Some Takeaways

- The **RAG-based solution** for HIR proves to be genuinely effective.
- In this case, **open models outperform GPT models**.
- The work was conducted within a **naive RAG framework**, without specific assumptions on tasks, text representations, or query optimization.
- Future developments will focus on task-specific adaptations, improved text representations, and advanced query optimization techniques.





Overall Takeaways

KURAMi, Marco Viviani, April 30, 2025 – Universitat Politècnica de València





Challenges and Open Issues

- Simple document protection techniques like token masking are insufficient on their own → Combining them with summarization and coreference resolution (and maybe LLMs?) leads to stronger confidentiality but can compromise retrieval effectiveness.
- Lexicon-based methods for query protection offer strong privacy at the cost of retrieval quality, while LLM-based query expansion shows promise in balancing protection and effectiveness, particularly in sparse retrieval scenarios.
- **RAG-based architectures demonstrate robust performance for Health Information Retrieval** (HIR), even in naive setups, with open models outperforming proprietary ones.

• Future work should focus on more sophisticated confidentiality-preserving methods, task-specific optimizations, and improved text and query representations.

Important Consideration for the KURAMi project: When designing privacy-preserving systems, we must also carefully address the trade-off between privacy risk and protection and the risk of misinformation generation and access.





Thank you for your attention

Gracias por la atención

KURAMi, Marco Viviani, April 30, 2025 – Universitat Politècnica de València