

FAQ – Cortellis Regulatory Assistant

Table of contents

1. What types of questions can I ask the Regulatory Assistant (General Release)?
2. Can I enter questions in multiple languages?
3. What is the advantage of using the Regulatory Assistant over general purpose AI chatbots and are there any limitations?
4. Can I choose to opt out?
5. How do I opt in?
6. Will I be able to continue using the regular search?
7. Will I need to access a different platform?
8. Is my conversation history secure and private?
9. Are you training public Large Language Models (LLM)?
10. What sources are used to generate content?
11. How are you ensuring against 'AI content hallucinations' and bad information?
12. What transparency do you offer regarding your AI-based responses?
13. How is customer data protected and kept secured?
14. How do your AI-based search tools rank and prioritize the sources listed?
15. How do you handle sensitive or offensive user questions?
16. How is Clarivate addressing the environmental impact of AI tools?

1. What types of questions can I ask the Regulatory Assistant (General Release)?

- a. You can currently ask questions to understand a requirement in a single market.
 - i. *Example:* What is the requirement for record retention in clinical trials conducted in Canada?
- b. You can also ask questions to find a specific document.
 - i. *Example:* Help me find a document containing a summary of Pharmacovigilance and risk management in Brazil.
- c. You can ask follow-up questions about the answers provided.
 - i. *Example:* "What about Thailand?", as a follow-up to a more detailed question about another country.

2. Can I enter questions in multiple languages?

- a. Yes, you can ask questions in your preferred language and will receive answers in the same language.

3. What is the advantage of using the Regulatory Assistant over general purpose AI chatbots and are there any limitations?

- a. Our domain specific Regulatory Assistant stands out from general-purpose chatbots by only leveraging Cortellis Regulatory Intelligence's high quality and trusted regulatory content including our exclusive Clarivate curated content. Developed with regulatory expertise and a user-centric approach, it is designed to understand user context deeply and provide accurate, tailored responses that meet the specific expectations and needs of regulatory users.
- b. For now, the Regulatory Assistant (General Release) helps with the types of questions above, but over time we will add more functionality based on client feedback.
- c. Depending on the type of question asked, the Regulatory Assistant will limit references accordingly.
 - i. *Example:* For questions to understand a requirement, the Regulatory Assistant will only reply with the latest valid final version of the document. Documents falling in the Product or Compliance & Inspection domains will be excluded.

4. Can I choose to opt out?

- a. Yes, you can opt-out of this functionality, as it will not be enabled by default for our initial release and will not be mandated in subsequent releases. If you or your company prefer not to have access to the Regulatory Assistant in future releases as well, please communicate this to your Clarivate Account Manager.

5. How do I opt in?

- a. The general release on September 30th will only become available if you opt-in to enable the functionality via [this web form](#). Access can be granted for all users on your account. If you are unsure whether this has already been requested for your company, please contact your Account Manager.

6. Will I be able to continue using the regular search?

- a. Yes, you can continue using the current search tools available on Cortellis Regulatory Intelligence. The Regulatory Assistant aims to enhance your search experience with a conversational search option.

7. Will I need to access a different platform?

- a. No, the Regulatory Assistant will be added as a new page within the existing platform.

8. Is my conversation history secure and private?

- a. Yes, your search history is never shared. Query data is the only piece of information shared with the AI model, but it is never shared with others, and it is stored in an anonymized database.

9. Are you training public Large Language Models (LLM)?

- a. We do not train public LLMs. We use commercially pre-trained Large Language Models as part of our information retrieval and augmentation framework. Currently, this includes the use of a Retrieval Augmented Generation (RAG) architecture among other advanced techniques. While we are using the pre-trained LLMs to support the creation of narrative content, the facts in this content are generated from our trusted regulatory sources. We test this setup rigorously to ensure integrity and alignment with the regulatory ecosystem. Testing includes validation of responses through regulatory subject matter experts who evaluate the outputs for accuracy and relevance. Additionally, we conduct extensive user testing that involve real-world regulatory scenarios to further refine accuracy and performance.
- b. We are committed to the highest standards of user privacy and security. We do not share or pass any user data to large language models (LLMs) for any purpose. We are using a Clarivate environment, so no data is sent back to the LLM providers.

User privacy and security are further enhanced by Azure AI Content Safety features.

10. What sources are used to generate content?

- a. While the LLM is a key tool to provide a fluent narrative, answers to user queries are based only on our extensive collection of curated regulatory content. This means that our AI-generated responses draw from trusted regulatory sources, such as expert curated content and agency-provided source documents, rather than broad (and potentially inaccurate or biased) internet content that common chatbots might use.
- b. Answers from the Regulatory Assistant will be generated based on content included with your user license. If your question is answered by content that you do not have access to, the Regulatory Assistant will inform you of this limitation accordingly.
- c. For sources subject to additional copyright, e.g. ISO, the Regulatory Assistant will have access to document metadata and will be able to find links to those documents. It will not be able to provide answers derived from the full text of the document.

11. How are you ensuring against 'AI content hallucinations' and bad information?

- a. We strongly believe that we have a critical responsibility to the regulatory community to mitigate AI-induced inaccuracies, and we take many steps toward achieving this goal:
- b. The prompts used in our products are crafted by expert prompt engineers who ensure that the settings of the LLM are optimized to maximize faithfulness to the source content and minimize hallucinations. Our system is designed to provide references to the source text when delivering an answer. Additionally, our systems can handle negative rejections, ensuring that it does not fabricate one when an answer cannot be provided due to insufficient information. Instead, the system presents an explanation to the user for the lack of response.
- c. The information presented to users always originates from our trustworthy, curated content. We use a combination of RAG models to ensure that the information users see is based on the vetted content your organization can access via Cortellis Regulatory Intelligence. Our tools offer full transparency regarding the content that was used to generate the response and ensure proper attribution to the specific sources used.
- d. We continuously test our solutions and the results they produce, including thorough dedicated beta programs and close collaboration with customers and subject matter experts. Our data science expertise helps increase system accuracy, fairness, robustness and interpretability in a programmatic way.

12. What transparency do you offer regarding your AI-based responses?

- a. Ensuring clarity and trust in our solutions is one of our top priorities. Our conversational search and AI-powered tools present a list of regulatory resources on which their responses are based, so that you can always explore relevant materials for further context.

13. How is customer data protected and kept secured?

- a. Data privacy and trust are top priority when designing our AI tools. We comply with data privacy regulations and adhere to the evolving global AI legislation. As part of the user query, only the content that users themselves input into the query is transmitted to the LLM. No additional data is shared during this process, ensuring the protection of sensitive information. Furthermore, we are not using any of the LLM API endpoints directly but accessing LLMs through a private setup. This ensures that data entered by users in the query stays protected and cannot be seen or accessed by any other party. For more information on our privacy and data protection program, visit: <https://clarivate.com/privacy-center/>

14. How do your AI-based search tools rank and prioritize the sources listed?

- a. The ranking and prioritization of sources by our AI-based discovery tools will vary according to the specific characteristics of the user's query and the context of each query.

The approach to ranking and prioritization is similar to the way it is traditionally done in the Cortellis Regulatory Intelligence platform. This understanding enables us to present the most relevant and valuable sources first, ensuring that the information provided matches the user's needs as closely as possible.

15. How do you handle sensitive or offensive user questions?

- a. Our tools provide the means for finding, comparing, and interacting with regulatory intelligence. But sometimes users may enter queries that are inherently biased, prejudiced, or seek sensitive information.
- b. While our tools will not normally block user queries, the foundational large language models we use for narration capabilities are trained to recognize sensitive or offensive queries and handle them appropriately.
- c. This means that such questions will typically be answered with a balanced perspective, drawing on our curated regulatory content. Please note that there is no guarantee that all sensitive or offensive content will be identified, and errors might occur. If you do encounter such content, please use the feedback mechanism to report it.

16. How is Clarivate addressing the environmental impact of AI tools?

- a. The conversation regarding carbon emissions caused by AI is part of a wider discussion around systems, data storage and sustainability which cannot be solved by any one organization. This is an important industry-wide challenge, which we are working with our vendors, customers and community to understand and address.
- b. By continuously focusing on actions and outcomes at Clarivate, we are making a positive impact on our business, our people and our planet. We are mindful of the need to reduce waste in systems and data storage and are committed to get to net zero carbon emissions before 2040. Our products and services are designed, developed and deployed following environmental and sustainable best practices including optimizing to reduce waste and pollution such as CO2 emissions. We are building a comprehensive climate transition plan that includes setting Science Based Targets (SBTs).
- c. We are working in close partnership with all our cloud systems and data storage providers (Amazon, Google and Microsoft.) Each of these companies has their own sustainability commitments and ambitions to get to net zero by 2030 or 2040, the majority of which is aimed at reducing emissions created in the first place through using green energy rather than offsetting.
- d. Our Environmental Management Statement outlines our framework to adjust existing working practices towards our Net Zero before 2040 target, which includes our usage of AI services. For more information, please see our Sustainability Report.
- e. For Cortellis Regulatory Intelligence, our Clarivate AI platform serves as a technology backbone, enabling a centralized and consistent deployment of AI capabilities across our portfolio of solutions and promoting efficient performance. We choose AI models that balance performance, cost, and sustainability. For each task, we always prioritize high-quality output with the least resource-intensive technology.
- f. Our centralized platform approach also helps eliminate system redundancies, reducing both resource consumption and emissions. Additionally, we use caching and text compression mechanisms to reduce the workload on the LLM, making LLM calls more efficient.



- g. We continue to engage with stakeholders across academic and life sciences communities to explore and implement best practices for reducing the environmental impact of AI.

For more information contact Customer Service at LSH.support@clarivate.com.

About Clarivate

Clarivate is a leading global provider of transformative intelligence. We offer enriched data, insights & analytics, workflow solutions and expert services in the areas of Academia & Government, Intellectual Property and Life Sciences & Healthcare. For more information, please visit clarivate.com.