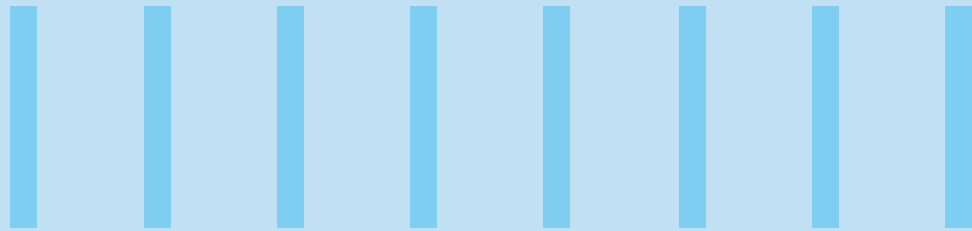




# DOCUMENT COMPREHENSION AND BENEFITS OF USING IT IN BFSI INDUSTRY



The first Chatbot was developed way back in 1960s and since then they have evolved a lot in functionality. These days, it has become almost necessary for Enterprises to include a Chatbot in their Company websites to serve two purposes - to answer Customer Queries 24/7 and to reduce the workload on its Customer Care Executives. Chatbots have been evolving continuously, and in the last couple of years, there has been significant progress. However, a key element in the current era is how to make Chatbots understand knowledge

in the way humans do. Over the last few years, Artificial Intelligence (AI) based Chatbot development is on the rise as it helps mimic human behavior and answer the queries in natural language.

There is a continuous endeavor by different organizations around the World to introduce new features, and Document Comprehension is one such revolutionary feature. The traditional approach is to configure all that is needed for Chatbots to understand the end user's queries. With significant

advancements in the underlying technology and deep learning algorithms, Document Comprehension is bringing a great change in how queries are answered by allowing Chatbots to directly comprehend from the underlying document/s. This feature will help Chatbots to understand how humans comprehend things, and thereby help them evolve into an even more powerful tool. In this article, we will delve into certain aspects of it along with the benefits of using it in the Banking, Financial Services & Insurance (BFSI) industry.



## Introduction to Document Comprehension

The main difference between how Humans and Chatbots understand things is in the way we comprehend knowledge. Humans can read and understand an artefact and based on that understanding; answer queries related to it. However, Chatbots can only answer depending on use case configuration and based on how their training was done.

Document Comprehension is an

advanced feature that helps Chatbots comprehend like humans do, and accordingly answer an end user's query by directly comprehending information from a document. It has the capability to understand details from both structured and unstructured documents, and answer an end user's queries by providing top paragraphs under a document source. The main difference between intent/FAQ-based

Chatbots and this feature is that in the former; intents, responses, and utterances are to be configured or uploaded. However, with Document Comprehension, it's about either ingesting document sources or pointing to respective knowledge bases. It is an evolving feature and is very helpful in terms of providing better responses for domain-based queries.

## Generic vs. Domain-based Models for Document Comprehension

In layman's terms, the difference between generic and domain models is similar to asking an English literate individual to comprehend Banking terminology vs. asking a Banking Professional. Even though the former can read the terminology, one will need specialization to comprehend its meaning. Hence the latter who has gained mastery in that domain by virtue

of obtaining the relevant degree/training/job is better equipped to handle it appropriately.

Similarly, when searching for something online, how one types in a query determines the answers that are displayed. Too few keywords will be insufficient, and at the same time, lengthier statements could be confusing for the system to

understand. Domain keywords will add up to the complexity in showing appropriate results as they are specific to a domain, and its meaning is derived accordingly. This is the case with generic and domain-based Document Comprehension models too. To get a better understanding of this, let's take a quick look at these models, and examine the contrasting results they generate.

### Generic Model

Generic question and answering models trained using open datasets like Stanford Question Answering Dataset SQuAD, NewsQA, have had great success in answering fact-based questions or questions with short answers. However, when these generic models are inferred to domain-specific documents, they aren't very accurate (it either does not answer the correct span or answers it incompletely).

### Limitations

- Incorrect or incomplete answers by generic models on domain specific queries
- Doesn't capture domain keywords and abbreviations
- Domain documents are not like Wiki pages on which the generic model is trained
- Distribution of question types for generic and domain queries is different

### Domain-based Model

Domain-based question and answering models are trained on domain documents. They learn domain vocabulary and linguistics while getting trained in domain data. Domain trained models also

understand domain specific keywords, abbreviations, and their relationships better. It makes these question-and-answering models more accurate on domain specialized documents.

Taking the example of an employee's query on the Termination Policy, let's see how the same query produces different results using Generic and Domain-based Models.



## Response from the Generic model

### Query

Who do I need to consult before the termination process is carried out?

### Expected Answer

1. Relevant Delivery Manager  
2. Relevant Engagement Manager  
3. Unit HR Manager  
HR will support the line managers in the process.

### Generic Model Response

Unit HR Manager

## Response from the Domain model

### Query

Who do I need to consult before the termination process is carried out?

### Expected Answer

1. Relevant Delivery Manager  
2. Relevant Engagement Manager  
3. Unit HR Manager  
HR will support the line managers in the process.

### Generic Model Response

Unit HR Manager

### Domain Model Response

The following entities must be consulted:  
1. Relevant Delivery Manager  
2. Relevant Engagement Manager  
3. Unit HR Manager  
HR will support the line managers in the process.  
An Exit Checklist must be completed on or before the employees' final day of employment.

*This illustration clearly shows how the Domain-based Model can comprehend better when compared to a Generic Model.*

## BFSI industry and future Chatbot Adoption

Market research indicates that BFSI will be in the top 3 domains that are expected to see significant Enterprise-wide Chatbot adoption by 2025. The top 3 domains as per one of such source is listed below.

Retail

Banking

Insurance

**Source:** <https://www.artificial-solutions.com/chatbots>

Another research piece by Mordor Intelligence lists Retail, BFSI and Healthcare as top 3 domains by 2025.

## Benefits of Document Comprehension in BFSI Industry

We chose the BFSI industry for this publication since it's in the top 3 domains as indicated above, and compared to other domains, new regulations come into force periodically. Unless these regulations are followed by the respective sectors, the penalty will be huge. To add to that, there are National-level regulations and State-level regulations that must be followed as well. This leads to constant changes in the underlying sources, and it will become a nightmare if only normal intent/FAQ Chatbots are developed and deployed.

Owing to this, even human agents need to continuously train themselves by looking at necessary artifacts to give relevant and correct answers to a consumer's query.

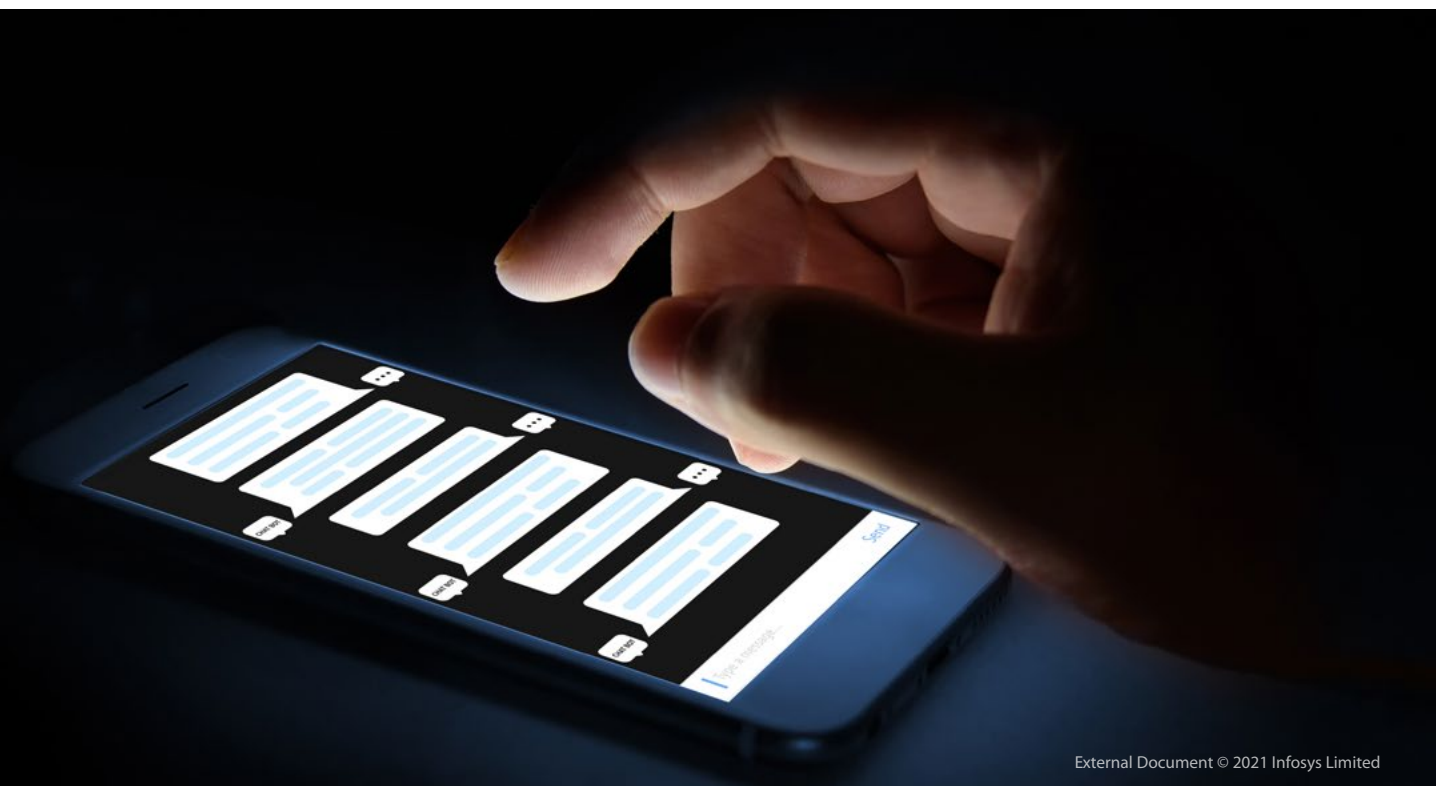
This puts enormous strain on the human workforce and on underlying systems which also need to be upgraded now and then.

Let's say, COVID-19 related expenses are covered under Health Insurance. The respective underlying systems of related companies need to be updated so that a claim under this category is processed appropriately. The workforce also needs to be educated so that they can counsel or help end-users if the need arises. So does the end-user communication in company websites, brochures, marketing emailers, etc.

Suppose there is a FAQ hosted on a company's website about a list of medical

conditions/pre-existing illnesses covered. Post update is done in the system w.r.t. COVID-19, and these details are made available. Similarly, after a while, another disease might have to be added, and the same set of activities will have to be triggered again to update communication.

If an FAQ-based Chatbot was deployed, then it would mean the update of underlying artifacts, retraining, and re-testing the configuration to show the details of COVID-19 too. However, if the chatbot had the Document Comprehension feature then, the same result could be achieved by pointing to the relevant source to comprehend the latest facts accordingly.



We think that this Document Comprehension feature could help various industries as mentioned below.

**Travel** – Revised travel guidelines in the

wake of Government regulations, natural calamities, etc.

**Retail** – Details on new product launch or new features supported, etc.

**Telecom** – Details on new consumer plans, product upgrades, etc.

**Education** – Latest details on scholarship launch, course fees, etc.

## Domain Models for BFSI Industry

The main challenge for Document Comprehension is that it needs a vast set of trustworthy data to train domain-based models. This would in turn bring accuracy to the results that are generated. The question is - how do we get such data to build better domain models? One option is to look at open access articles published in reputed forums. This then becomes a knowledge source to point to and build the model.

Below is such a list for reference:

- <https://www.mdpi.com/2227-7072/8/2>
- [https://www.mdpi.com/search?q=&authors=&journal=jrfm&article\\_type=&search=Search&section=&special\\_issue=&volume=8&issue=2&number=&page=](https://www.mdpi.com/search?q=&authors=&journal=jrfm&article_type=&search=Search&section=&special_issue=&volume=8&issue=2&number=&page=)
- <https://www.journals.elsevier.com/journal-of-banking-and-finance/open-access-articles>
- <https://onlinelibrary.wiley.com/action/doSearch?AllField=free+access&SeriesKey=15406261&startPage=&ConceptID=15941>
- <https://onlinelibrary.wiley.com/action/doSearch?ConceptID=15941&Ppub=&field1=AllField&text1=free+access&startPage=&ConceptID=87>
- <https://www.journals.elsevier.com/journal-of-international-accounting-auditing-and-taxation/most-downloaded-articles>
- <https://www.journals.elsevier.com/insurance-mathematics-and-economics/most-downloaded-articles>

## Market Players

Most market players have cracked the code to develop Intent/FAQ-based Chatbots, but only a few like Google (Knowledge Connector<sup>Beta</sup> and Document AI), Microsoft (Machine Comprehension), and Nia Chatbot Platform (Document Comprehension) are bringing advancements in various degrees with their offerings in this space.

Google's Dialogflow has developed a Product 'Knowledge Connector' which is in its Beta Stage, and has a solution named 'Document AI'. Knowledge Connector can parse information from knowledge

bases via Application Programming Interface (API), while Document AI supports the extraction of information from both structured & unstructured data, handwriting recognition via Optical Character Recognition (OCR), Form Parser & Invoice Parser via a single API.

Via its Reasoning Network, Microsoft's Machine comprehension can scan documents and try to respond like how a human would.

Currently, the Nia Chatbot Platform has the Document Comprehension feature

based on the Bidirectional Encoder Representations from Transformers (BERT) technique. On-premise deployment is supported which would allow organizations to bring on-premise documents in scope without migrating them to either Cloud or the vendor environment. Based on this feature, the Nia team recently built a solution that can scan over 47,000 documents, comprehend them, and show relevant top paragraphs out of the document.

## Summary

Document Comprehension's capabilities are still evolving, but it remains a great solution to answer an end user's queries based on domain sources. It is a huge opportunity to significantly improve a chatbot's learning capabilities, will help augment search capabilities to a great extent, and is a step towards Natural language understanding (NLU) based search. This in turn is hugely beneficial in the enterprise space, as augmenting NLU based search with Document Comprehension will bring in powerful search capabilities to chatbots.





## References

- <https://www.microsoft.com/en-us/ai/ai-lab-machine-reading>
- <https://cloud.google.com/dialogflow/docs/how/knowledge-bases>
- <https://cloud.google.com/dialogflow/docs/knowledge-connectors>
- <https://cloud.google.com/solutions/document-ai>
- <https://www.artificial-solutions.com/chatbots>
- <https://www.mordorintelligence.com/industry-reports/chatbot-market>

## Authors

### Kavitha Sundararajan

Kavitha Sundararajan is a Senior Consultant working with Emerging Technologies Team and handles multiple activities like interactions with various stakeholders, developing Marketing Strategies, building Product Roadmaps, and working in Large Transformation Deals.

### Amit Kumar

Amit is a Senior Architect working in the Conversation UI Research Team. He specializes in Artificial Intelligence, NLP, Machine Learning, Deep Learning, Cognitive, Mobile, Social, Wearable, and Converged Digital Experiences. Amit leads a team in the Innovation & Product R&D space with a passion for continually creating new business value in the entrepreneurial setting. He is involved in conceptualizing and generating Intellectual Property, and building product lines in the areas of Artificial Intelligence.

For more information, contact [askus@infosys.com](mailto:askus@infosys.com)



© 2021 Infosys Limited, Bengaluru, India. All Rights Reserved. Infosys believes the information in this document is accurate as of its publication date; such information is subject to change without notice. Infosys acknowledges the proprietary rights of other companies to the trademarks, product names and such other intellectual property rights mentioned in this document. Except as expressly permitted, neither this documentation nor any part of it may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, printing, photocopying, recording or otherwise, without the prior permission of Infosys Limited and/ or any named intellectual property rights holders under this document.