



**eva**

Enterprise Conversational AI  
by NTT DATA

**Boost your conversations  
with eva automated  
learning**

## INDEX

1.	WHAT'S EVA AUTOMATED LEARNING .....	1
2.	HOW DOES IT WORK? .....	2
3.	USE CASES .....	3
4.	HOW TO SETUP THE INFORMATION TO UPLOAD .....	4-6
5.	HOW TO UPLOAD THE INFORMATION .....	6-8
6.	HOW TO TEST THE INFORMATION .....	8-10
7.	RECOMMENDATIONS .....	11
8.	TECHNICAL INFORMATION .....	12



## 1. WHAT'S AUTOMATED LEARNING

---

Automated learning (AL) is a cognitive engine developed by eva team with the latest advances in NLP. Automated Learning's main function is facilitate the training of virtual assistants and be an additional reference base when the intention is not trained.

It is a question-answering software that interpret users' questions and answers them based in documents uploaded beforehand.

When a user asks something, such as "What is the capital of Madagascar",

Automated Learning interprets this question, reads all documents in the knowledge base after the answer and finally delivers "Antanarivo".

The Automated Learning documents and questions are trainable even if your bot doesn't use Clever. The training works just like Clever training. You can check the details after training and see what went wrong after a failed training.

Today is available in **spanish, english and portuguese.**

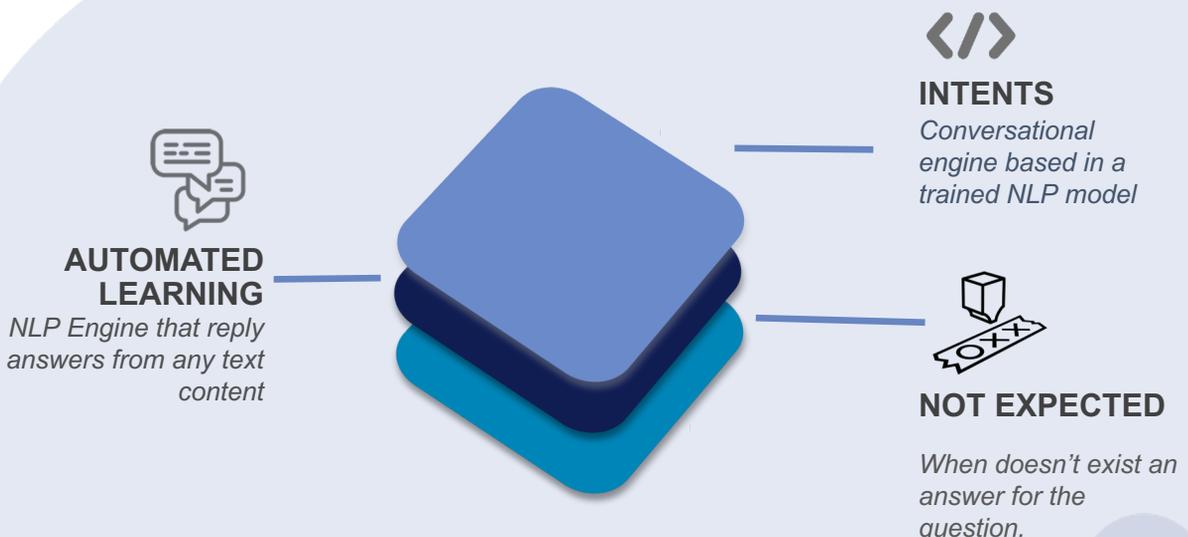
## 2. HOW DOES IT WORK?

Automated Learning finds the answer to the user's question in one or multiple documents, without the need to curate a specific dataset for it, like it is needed for the intents engine so, there is no need to execute specific model training to work with a text knowledge base.

Once the user enters the question, the cognitive engine looks for an intent that matches the question. In case of not finding it or having a confidence score below the defined threshold, the cognitive engine starts looking in the uploaded Automated Learning documents for an answer that matches the question and generates an output with a score,

indicating the likelihood of that answer being the correct one. It also has the ability to detect when the question doesn't have an answer and, in that case, fallbacks to the not expected answer.

To reduce response times for each request, AL has a recommendation component, that is going to be responsible to filter the most probable paragraphs, that contains the answer to a given question.



### 3. USE CASES

You can use AL to speed up responses related to different topics. We highly recommend use AL in this cases:

- **Long-tail and FAQs:** It's a kind of content that works very well with AL. It's mostly use to answer subjects that are least frequently asked by the users. When the question isn't trained as an intent, AL can be use as fallback.
- **About the company:** Some clients want to know more about the company. Instead of training your virtual assistant with that information, you can upload a document with the history of the company, its vision, mission, addresses and contact information.
- **User guide and handbooks:** If you have user guides for your clients related to procedures, how to do's, etc. It can be automatized by Automated Learning.
- **Technical information:** Information about architecture design, training, functionalities or any other technical information base on products.
- **As standalone service:** You can use only Automated Learning as a bot, to answer questions about registered/trained documents.  
  
Keep these points in mind if you take this use case:
  - Minimum control over the conversation
  - Only one flow to activate the service
  - More than one possible answer to the user question
  - User can choose the correct answer to disambiguate
  - Low curation time
  - It is still necessary to curate and adjust the documents to AL standards, but it takes less time than creating an intents/entities dataset

Important: For specific or security information that requires 0% error, we highly recommend training the bot through intentions

## 4. HOW TO SETUP THE INFORMATION TO UPLOAD

Before expanding a bot knowledge base with documents, it is important to follow some guidelines when building those texts. Although Automated Learning can read and extract answers from any text, some best practices make its assertiveness better. As a general guideline, try to keep your text direct and simple.

Today, Automated Learning **only accepts documents in .txt format.**

### 1. Longer paragraphs works better than shorter ones

Automated Learning works with context. So, the longer the material is, the inference is better. Paragraphs with 4 sentences works better than paragraphs with 2 sentences. If you have 2 small paragraphs, it is better to join them and make a larger, 4-sentence paragraph. Example: Instead of writing John is a nice man. John lives in New York. John likes zucchini. Write "Jonh is a nice man. He lives in New York. John likes zucchini".

### 2. Avoid bullet lists

Bullet lists are the extreme example of the lack of context that small paragraphs have. Try to make a paragraph of those lists. Instead of a list, such as:

Gazpacho

- Tomato
- Onion
- Garlic

- Stale bread
- Salt, olive oil and vinegar

Try condensing in a paragraph, such as:

*To make gazpacho, you will need tomato, onion, garlic, stale bread, salt, olive oil and vinegar.*

The answers delivered by Automated Learning are short, up to 10 words. If you have a large list, try to split in 2 paragraphs.

Example:

If you have the List A:

List A

- Items 1 to 10
- Items 11 to 20
- Items 21 to 30

Build three lists:

List A – Items 1 to 10

List B – Items 11 to 20

List C – Items 21 to 30

Then, fashion those lists as three different paragraphs.

### 3. Avoid tables and forms

---

Automated Learning does not read well tables and forms. It is better to rewrite the contents in tables and forms as paragraphs or dismiss them.

### 4. Remove special characters, emojis and images

---

Non-ASCII characters tend to turn to strings of characters that are not readable by Automated Learning. Although some special characters are common (like the dollar sign, \$), some are more obscure and can hamper Clever inference. Example, the character ♠. Emojis count as special characters and are better removed. Images also might interfere with Clever AL functioning.

### 5. Remove headers, footers, summaries, references and any other elements that are not relevant

---

Headers, footers, summaries and elements alike add unnecessary information that might make harder for the cognitive engine to deliver precise answers.

### 6. Pay special attention to punctuation

---

Errors in punctuation hamper understanding much more for a cognitive engine than for a human reader. Punctuation matters a lot. “I’m sorry; I love you” is an apology. “I’m sorry I love you” is cruel.

### 7. Avoid passive voice

---

“Gazpacho is an example of cold tomato soup” works better than “An example of cold tomato soup is gazpacho”.

### 8. Avoid subordinate clauses

---

If you have a sentence with subordinate clauses, try to dismember it in smaller sentences. “Gazpacho is a cold tomato soup from Spain that, although characterized by the tomato, didn’t took tomato until the 19th century” is a long sentence with lots of subordinate clauses.

“Gazpacho is a cold tomato soup. It is from Spain. Although characterized by the tomato content, it didn’t take tomato until the 19th century” works better.

## 9. Pay attention to the voice and tone of your documents

As the answer inference is automatic, if you extend your bot voice tone to its documents, you will guarantee a consistent bot personality.

Some tips when uploading FAQs (Frequently Asked Questions):

- Before uploading a FAQ on Automated Learning, make sure that you remove the questions and upload only the answers. Automated Learning works on FAQs better if the questions are absent.
- When writing a FAQ, think of your text as question - answers pairs. First, write them and then remove the questions.
- Think of answer to short questions. Not only most doubts can be expressed

with short questions, but Automated Learning works better when the questions are short.

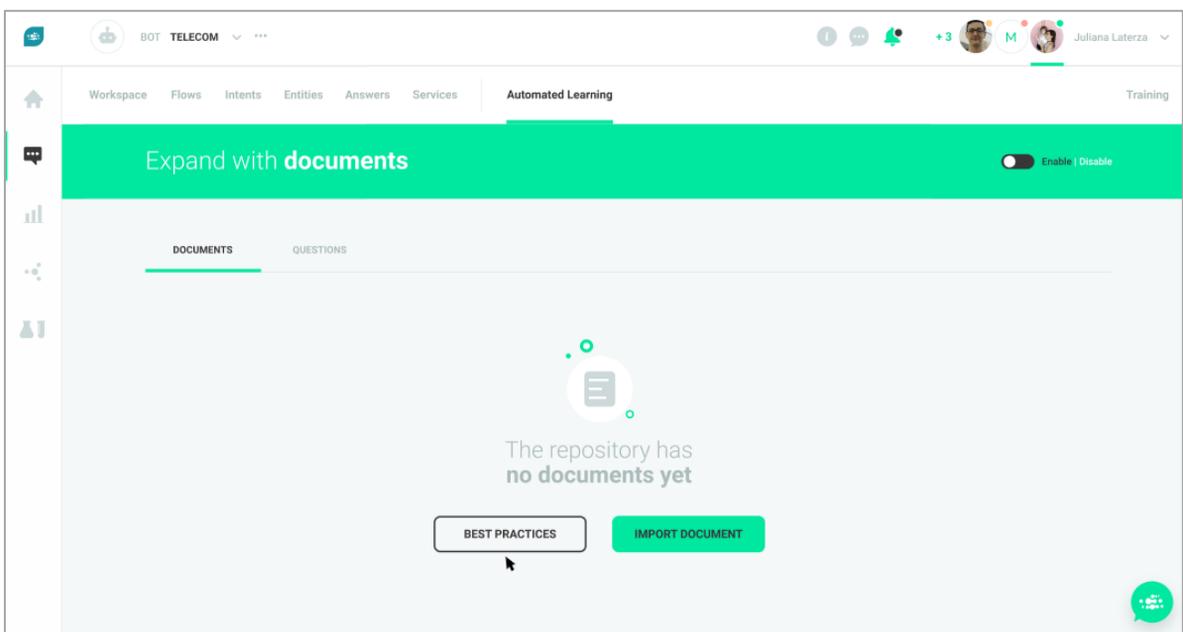
## 5. HOW TO UPLOAD THE INFORMATION

After you adapted your documents following those guidelines and converted them to .txt, you can upload them to Automated Learning.

Clicking on the Automated Learning tab in the Cockpit will lead you to Automated Learning first page.

First, you have to enable Automated Learning in eva, as it comes disabled by default.

After that, you can upload documents.



To import a document, click on “Import document”. After the first document is imported, you will be taken to the documents’ repository.

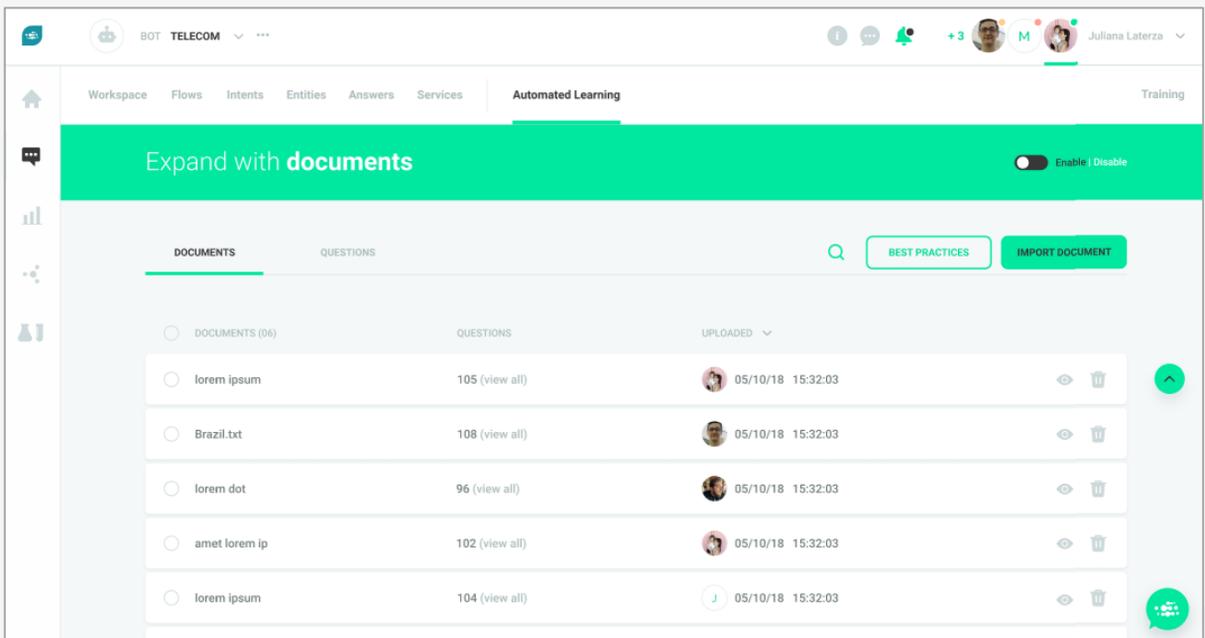
There, you can check the documents that are part of a bot knowledge base, see how many and which questions were coupled to a particular document, delete, disable or enable a document.

Important: Always train the bot after importing the document

## Documents’ repository

By clicking on “view all” besides the number of questions, you will see the questions tied to this particular document.

By clicking on “view document” (the eye icon), you will be taken to a particular document.



## Expanded document

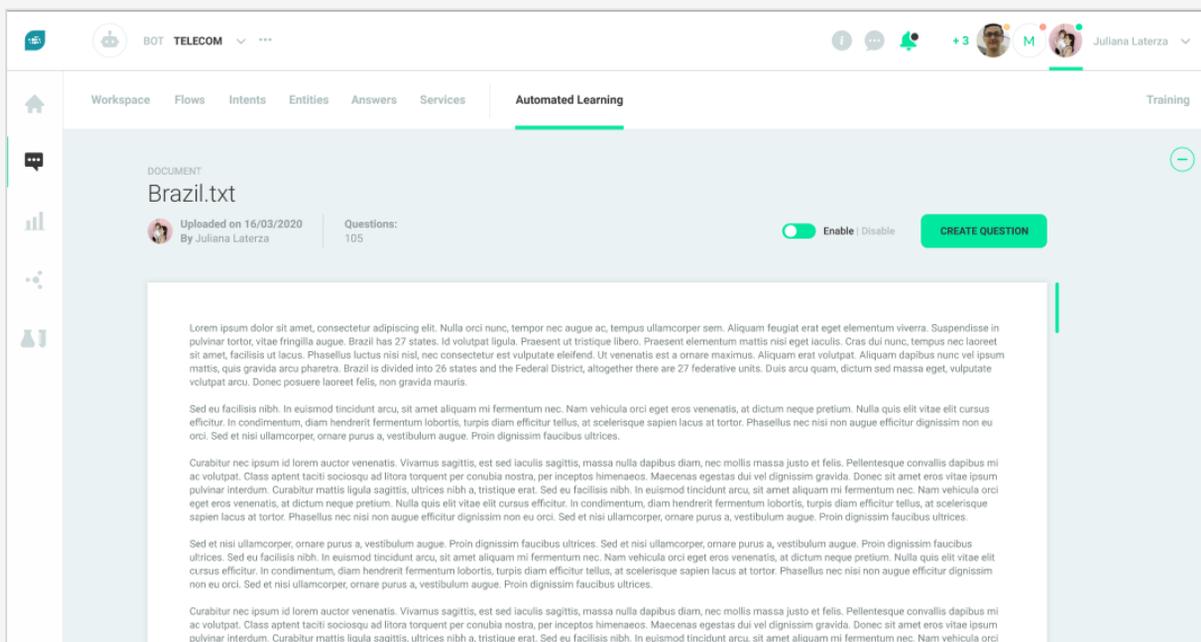
Once you are in a document, you can enable or disable it and create specific question and answer pairs to this document.

**When you disable or delete a document, all questions attached to it are also disabled or deleted.**

By clicking on “Create question”, you will be taken to the question creation modal,

where you can couple a question and its variants to a specific answer.

Although Automated Learning can answer users’ questions without prior registration, by creating questions, you can tie a user query to a specific answer and edit this answer.



## 6. HOW TO TEST THE INFORMATION

### Creating a question

After you name (and optionally describe) a question, you will have to add the question itself and a number of variants. Simply insert the question and its variants in the “add question” field. You can use tags to classify this question.

Example: in a text about soups, you want to tie the question “What is gazpacho?” to the answer “Gazpacho is a cold soup from Andalucia, Spain”.

Some people might ask this question as “What is gazpacho?” but might ask as “What is gazpacho?” or even “What gazpacho is?”. There are numerous variants to any question. Try to add the most variants so most forms are covered.

After you created the question, Automated Learning will infer the answer and deliver to you while opening a modal so you can edit this answer.

Automated Learning infer the answer after the first question is registered.

The screenshot shows the 'Automated Learning' interface. On the left, a document titled 'Brazil.txt' is displayed, uploaded on 16/03/2020 by Juliana Laterza. The document contains several paragraphs of placeholder text. A modal window titled 'Create question' is open on the right, featuring a 'Name' field, a 'Description' field, an 'Add question' field with a plus sign, and a '# Tags' field. The modal also includes an 'Enable' toggle switch and a 'NEXT' button at the bottom right.

## Inferred answer

After the answer is inferred, a modal named “Curate answer” opens with the answer (like a regular answer modal, but with the inferred answer in the text box).

You can copy any part of the text and paste on the text box or rewrite the answer as you wish. As in a regular answer, you can add buttons and/or technical text.

Example:  
For the question “What is gazpacho?”,

Clever AL found the answer “A cold soup from Andalucia, Spain”. You can edit it so it becomes “Gazpacho is a cold soup from the region of Andalucia, Spain”.

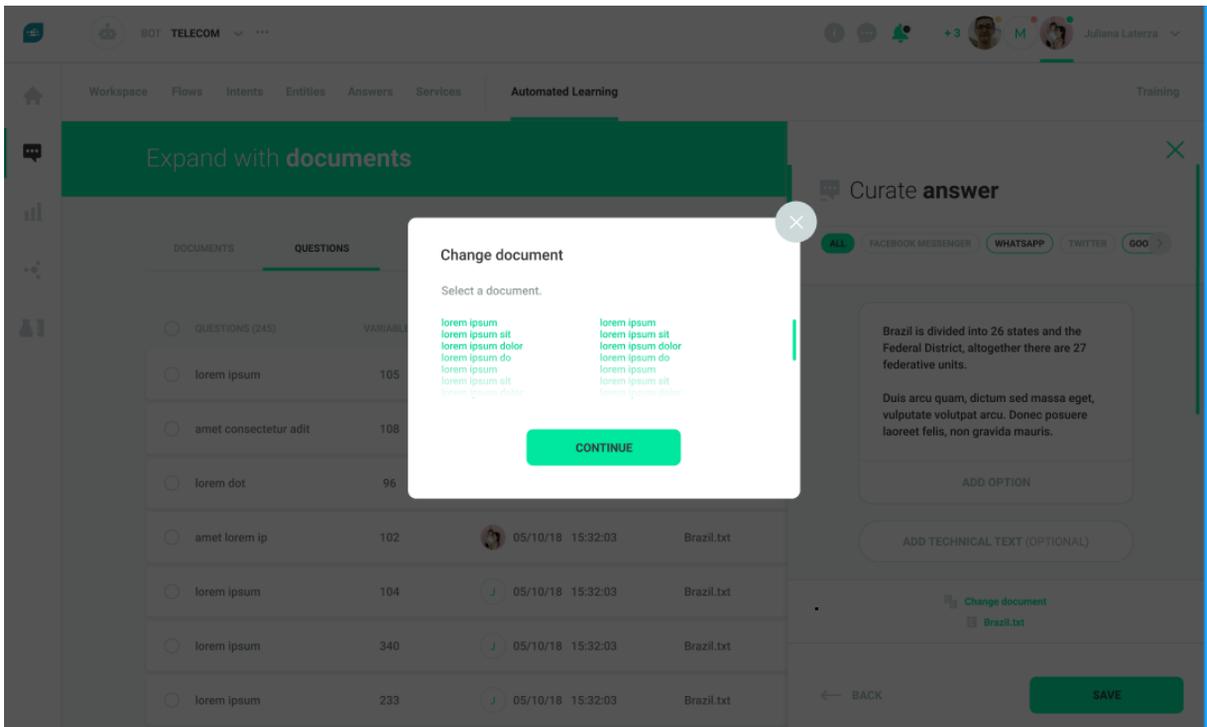
If you want to link the answer to other text, you can swap the document that the answer belongs by clicking the “change document” button. A modal with documents open so you can select the new document.

## Every question must be attached to a document.

The screenshot shows the NTT Data AI interface. At the top, there's a navigation bar with 'Workspace', 'Flows', 'Intents', 'Entities', 'Answers', 'Services', and 'Automated Learning'. The main area displays a document titled 'Brazil.txt' with a question 'Questions: 105' and an inferred answer highlighted in green. The answer is: 'Brazil is divided into 26 states and the Federal District, altogether there are 27 federative units.' A modal titled 'Curate answer' is open, showing the same answer and allowing for editing. The modal includes buttons for 'ALL', 'FACEBOOK MESSENGER', 'WHATSAPP', 'TWITTER', and 'GOO'. Below the answer, there are buttons for 'ADD OPTION' and 'ADD TECHNICAL TEXT (OPTIONAL)'. At the bottom of the modal, there is a 'Change document' button and a list of documents, including 'Brazil.txt'. A 'SAVE' button is at the bottom right of the modal.

## Document swap modal

After you click “save”, the question/answer pair will be stored in the Questions’ repository



## Questions’ repository

In this repository, you can see the question name, how many variants this question has, who last edited it and which document it belongs to.

By clicking on the filter icon, you can see filter questions by document.



## Filter icon

By clicking the pencil icon, you can edit the question.



## 7. RECOMMENDATIONS

---

Follow these tips for better use of AL:

- We recommend upload document up to 400 KB o 685 paragraph.
- Always test the information uploaded with possible questions that could be answered with the information in the document.
- Use the format .TXT to upload the information
- If part of the question is within the document, there is a better chance of giving the correct answer
- The number of paragraphs that is sent to the Q&A component depends on different variables like the **contents of the document** or the **number of paragraphs**. More paragraphs **increases the accuracy** but it also **increases the response time** of each prediction.

## 9. TECHNICAL INFORMATION

- Allows installation on cloud or on premise allowing full control of the information that clever travels in training and inference.
- It allows a deep customization of the models in the characteristics and linguistic environment of the use cases.
- All modules are implemented on DataOps architecture with horizontal scaling in inference and training, ensuring agility and security in the deploy process.
- Clever is ready to run on Nvidia GPUs



NTT DATA



**eva**

Enterprise Conversational AI  
by NTT DATA