dathena

Classification Review

"A good human plus a machine is the best combination"

Garry Kasparov, Russian chess grandmaster and former world chess champion

Introduction

Today, many organizations employ manual processes to identify, classify, secure, and protect the data under their custody. Unfortunately, these processes are no longer a viable option as data volumes swell. To overcome this challenge, Dathena has developed solutions that modernize and enhance data security. Its use of revolutionary AI-based technologies solves the overpowering problem of data security management.

These technologies offer state-of-the-art security at a fraction of the cost of other solutions or manual approaches. In fact, for each hour worked, AI can analyze as much documents as a human would in 10'000 hours, with an accuracy that is 3 times higher in average.

While AI is an automated entity that operates independently once trained, it is essential to involve humans in the AI's learning loop. This ensures the highest quality, performance and explainability. To do so, Dathena has developed its Classification Review module.

Classification Review is a process where clients validate the results of Dathena's classification mechanism and correct wrong predictions. However, Classification Review doesn't only ensure a better quality, it also allows tailoring the solution to your organization by aligning on the asset classification policies.

Setting up Classification Review

At Dathena, we understand that the most valuable commodity of a company is time. As such, we aim to provide a streamlined module that can be integrated easily to your own processes. More importantly, we aim to provide a simple user experience that will allow your key users to efficiently review their assigned documents.

SECTION 1.1

Aligning on the Asset List

Before starting the Classification Review, it is important for Dathena and your organization to be aligned on the leveraged asset list to define the business category (e.g. Finance, Client Data, Technology …) and confidentiality level (e.g. Internal, Confidential, …) of your documents.[1] [2]

In this case, there are 3 options:

1 **Dathena** can leverage your own asset list to classify documents

2 **Dathena** can work with you to define a customized list leveraging your own asset list, as well as Dathena's key categories

3 **Dathena** can provide you with a default list of categories tailored to your industry if your company does not have an asset list

Aligning on the Classification Review Process

Once Dathena and your organization are aligned on the classification dimensions of your data (e.g. business category, confidentiality level, …), the next step is to align on the goals of this exercise. Clearly defining the process ensures a smooth review tailored to your own needs.

More specifically, three components should be defined before kicking off the Classification Review exercise:

1 **The first step** is defining the expected improvement in accuracy after classification review. Reviewing more documents ensures a better performance, but necessitates more time, while reviewing less documents will be faster but not as significant.

2 **The next step** is to define which departments should take part in the review exercise. You can either choose to involve one (or a few) departments at a time, or organize a company-wide Classification Review campaign with all departments.

3 **The final step** is to define which folders and data sources should be targeted by user review to focus on business-critical data.

SECTION 1.3

Aligning on the Users Involvement

Dathena has experience helping organizations navigate the Classification Review exercise as best as possible. As such, Dathena has developed a recommendation system to help you find the right balance between the quality premium offered by the Classification Review and the time reviewers can afford on the exercise.

1 **Dathena's** Active Learning module is able to determine the optimal number of documents to review to achieve your expected improvement in accuracy (around 20 documents per business category)

2 **Dathena** can also recommend the most appropriate set of reviewers, which is the smallest set of high-risk users with access to recent business-critical data (around 10 to 15 reviewers per campaign)

🕐 **On average, a document takes 3 minutes to be reviewed.**

Sampling the Right Documents

To choose the most significant documents to review, we leverage 2 key principles:

1 **Diversity:** ML models are very good at recognizing patterns; thus, an ML model will perform better if it can see the most diverse set of documents possible.

2 **Uncertainty:** ML models do not only predict the label of documents, they can also determine how certain they are in their predictions. This allows selecting the documents which the model is most uncertain about, usually coming from underrepresented classes.

In parallel of these 2 principles, Dathena has developed an access-based assignment of documents. This means that reviewers only receive documents they have been granted access to across the organization, instead of documents from a particular category.
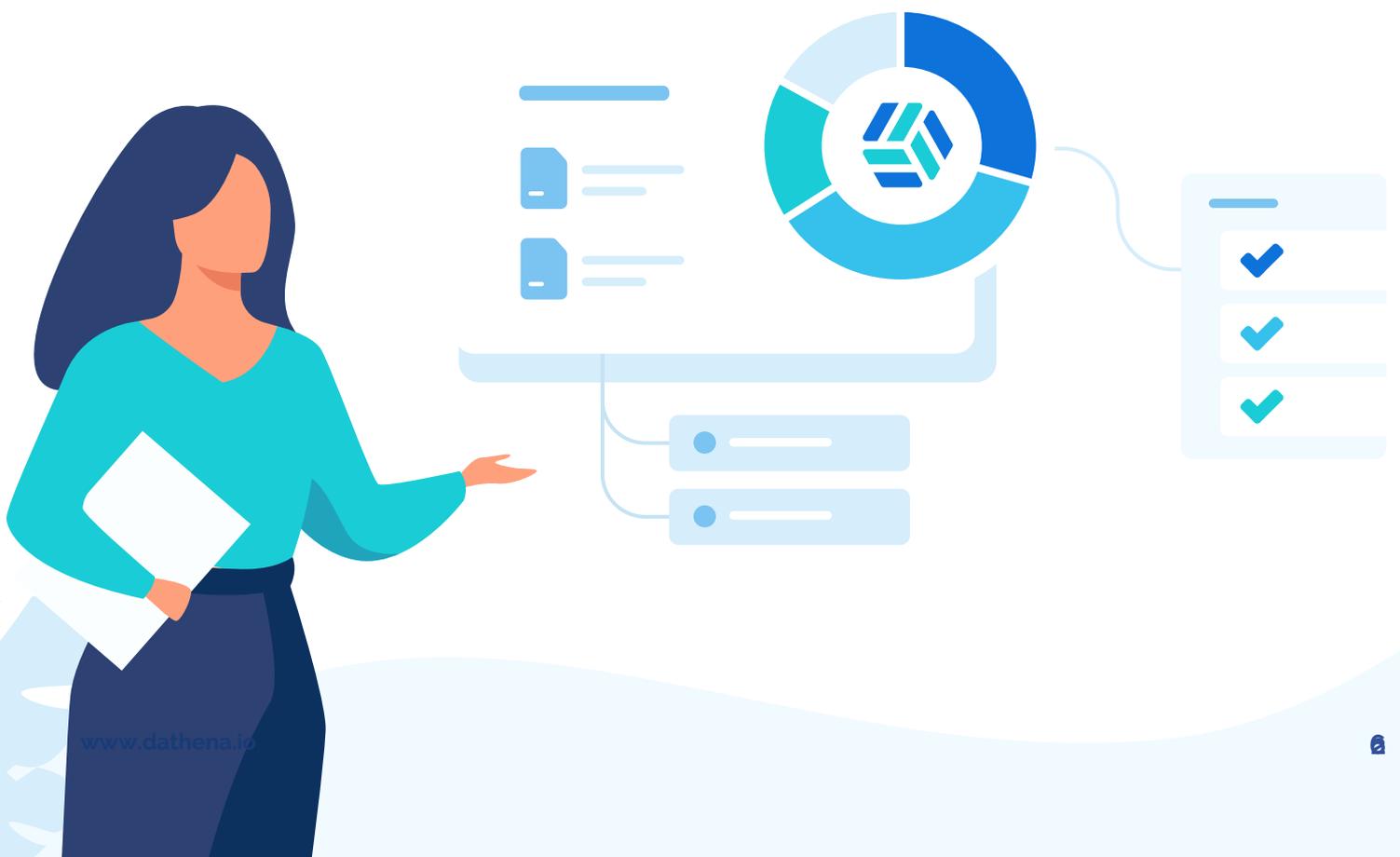
Classification Review Interface

Dathena offers a user interface seamlessly integrated to the process to facilitate the review exercise, including a preview of the documents to review. This includes both an interface for the Data Owners¹ (or Project Managers) to set up review campaigns and monitor their progress, but also for the Key Users² who will review the documents.

To ensure that all reviewed documents are classified correctly, Data Owners are in charge of validating the proposed classification – or challenging it if they feel like another label would be more appropriate.

---

1. Key Users are high-risk profile individuals with access to a large volume of the most sensitive data.

2. Data Owners are individuals who are responsible for overseeing and protecting a data domain.

Retraining the Classification Models

Retraining classification models after Classification Review leads to a significant increase in performance, driving the accuracy by up to 15%. Below are real examples showcasing the performance increase after model retraining:

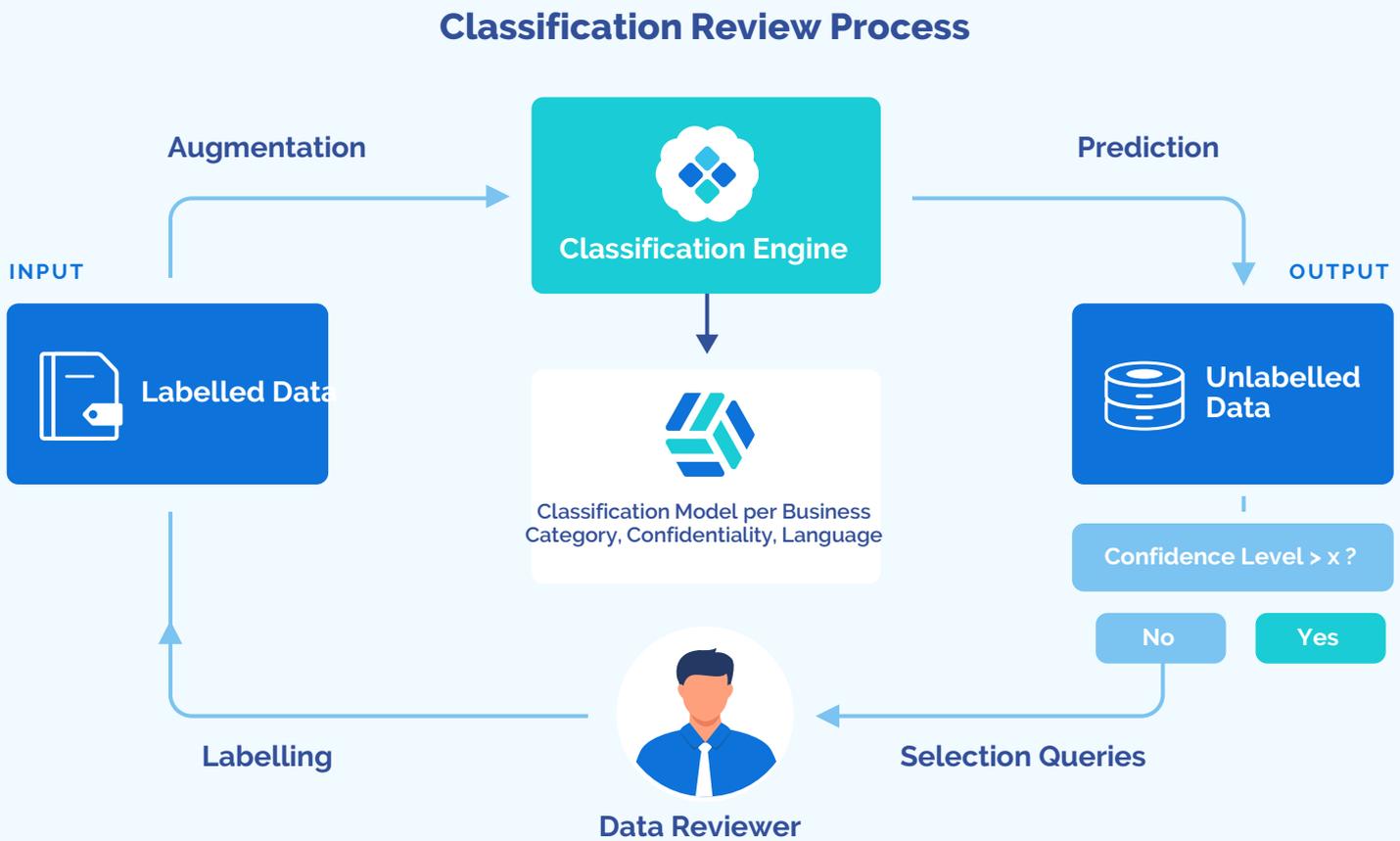|	Client 1	Client 2	Client 3	Client 4
Number of Reviewers	13	6	10	14
Total Number of Reviewed Documents	315	190	360	1200
Business Category Performance Increase	+18%	X	+16%	+15%
Confidentiality Performance Increase	+33%	+19%	+11%	X

Table 1: Performance Increase per Classification Review Campaign

Performing the Classification Review

After its setup, the Classification Review exercise follows a simple 4-step iterative process:

Classification Review Process

INPUT

Augmentation →

Classification Engine

Prediction →

OUTPUT

Labelled Data

Classification Model per Business Category, Confidentiality, Language

Unlabelled Data

Confidence Level > x ?

No   Yes

Labelling

Data Reviewer

Selection Queries

---

1 Once your data is classified, Dathena uses Active Learning to sample the most significant set of documents with the highest impact on the classification model's quality.

2 The identified Key Users receive the documents to review, and either validate Dathena's predictions or correct them.

3 These reviewed labels are then propagated to the most similar documents to improve the effectiveness of the review.

4 Finally, Dathena's classification model retrains itself by leveraging the new information to yield optimal results.

🕐 **This exercise is straight-forward and usually lasts around 2 weeks.**

It is recommended to organize a Classification Review campaign every 9 months to a year to ensure a consistent accuracy of the models.

Conclusion

Through its Classification Review module, Dathena leverages AI to allow Data Owners and Key Users to validate the results of its documents classification. More specifically, Classification Review aims to:

1	Validate the accuracy of Dathena's classification module	2	Improve the performance of the classifier	3	Align Dathena's predictions with your own internal policies

This offers an unprecedented performance as it allows you to improve the quality of the classification mechanism over time as it evolves with your own data. Moreover, it offers the possibility to tailor the mechanism to your own asset list by aligning any prediction with your own internal policies.