AI DRIVEN

CONTENT PRODUCTION,

CONTENT MANAGEMENT

AND CONTENT DELIVERY

n this whitepaper, we provide you an overview of our content labelling platform and its applications.

If you are new to EDIA's content labelling, this whitepaper will provide you with a quick overview of our technology and its applications.
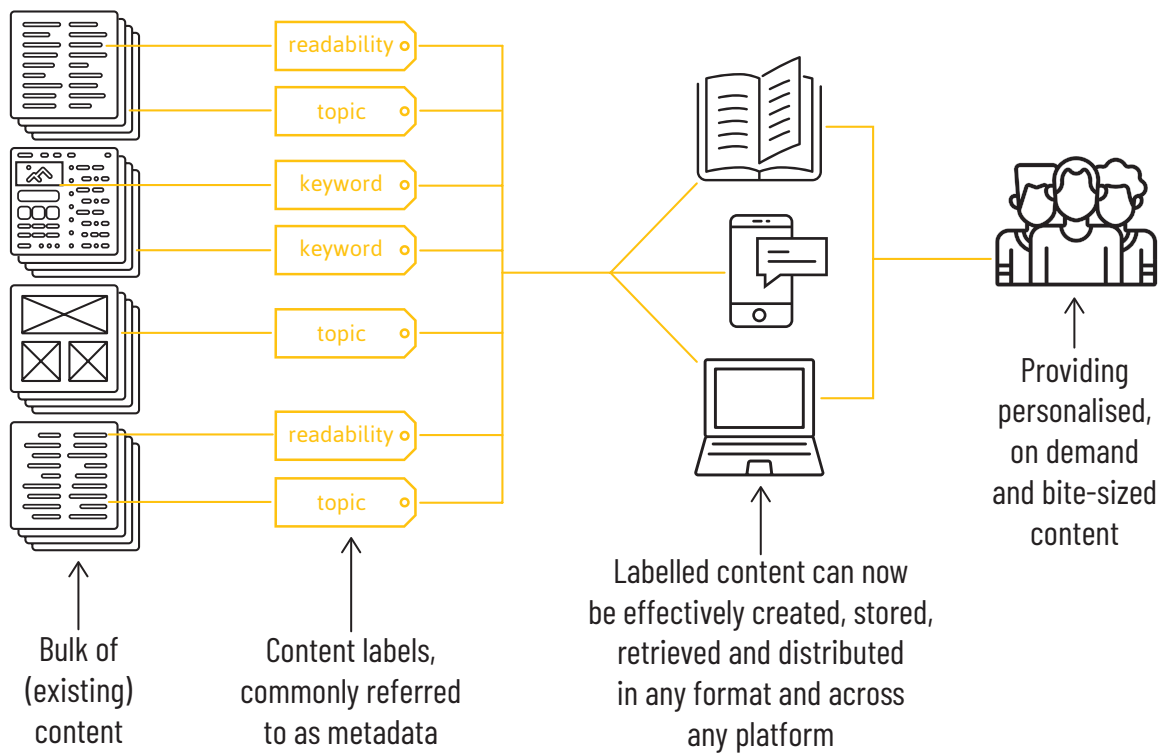
As an existing customer, this whitepaper can give you the information you need to discover new and additional capabilities of the EDIA content labelling platform.

AI-driven content labelling

Content structuring and content enrichment are core challenges for publishers and content creators as they move into an increasingly digital and data-driven world.

EDIA enables publishers and content creators to work successfully in today's data-driven world. In essence, our work is all about guiding the flow of content from inception to its target audience. Our content labelling platform aims to radically improve content production, content management and content delivery.
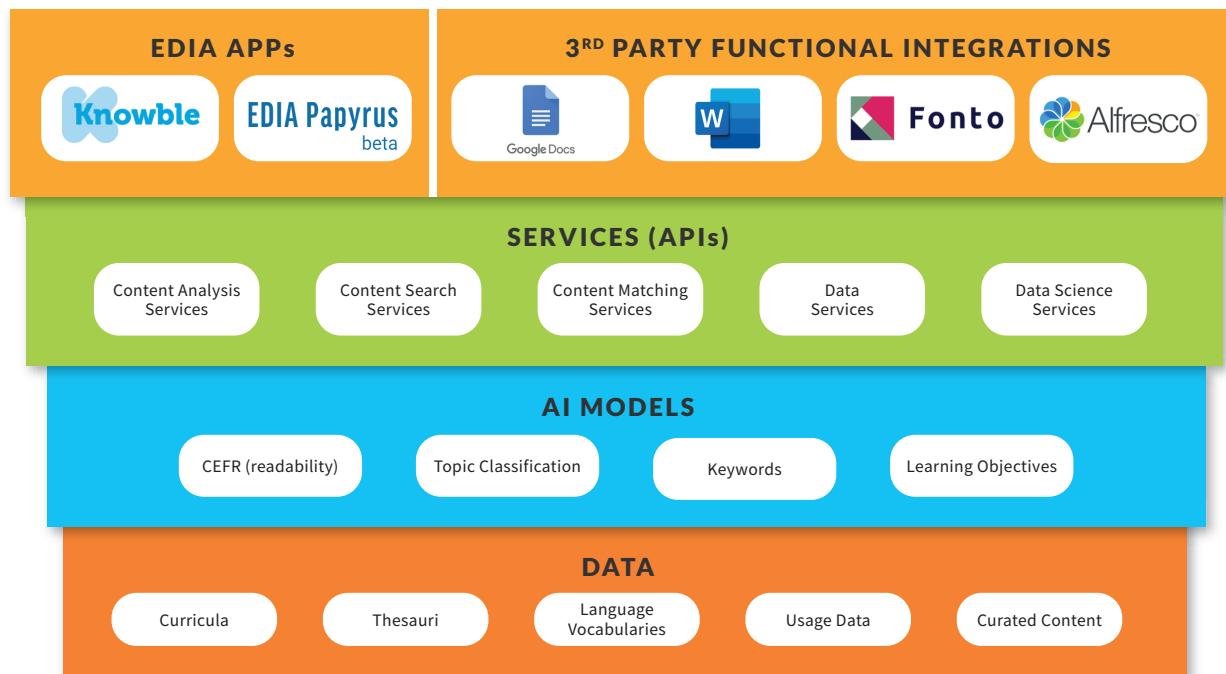
The EDIA content labelling platform includes several turnkey content labelling solutions as well as a unique set of tools that allow us to solve specific content labelling challenges in collaboration with our customers.

Bulk of (existing) content

Content labels, commonly referred to as metadata

Labelled content can now be effectively created, stored, retrieved and distributed in any format and across any platform

Providing personalised, on demand and bite-sized content

This means that on the one hand, the platform can be used as a solution for publishers that need a turnkey solution for common challenges such as keyword extraction or readability analysis. On the other hand the platform is a fully managed service that provides publishers with the ability to build, train, and deploy their own machine learning models for content labelling. The unique ability to give our customers the opportunity to create their own models with our help, besides access to existing models, is what makes our platform extremely versatile. All the tooling we have added is designed to remove the heavy lifting from each step of the machine learning process to develop high quality models both faster and easier.

The EDIA content labelling platform

The content labelling platform consists of 4 layers, each supporting a set of use cases:

EDIA APPs	3RD PARTY FUNCTIONAL INTEGRATIONS
Knowble EDIA Papyrus beta	Google Docs W Fonto Alfresco

SERVICES (APIs)

| Content Analysis Services | Content Search Services | Content Matching Services | Data Services | Data Science Services |

AI MODELS

| CEFR (readability) | Topic Classification | Keywords | Learning Objectives |

DATA

| Curricula | Thesauri | Language Vocabularies | Usage Data | Curated Content |

App & integration layer

The application layer offers functionalities and interfaces that use the platform's AI models. There are a number of stand-alone apps offered by EDIA and several readily integrated 3rd party apps.

Services layer

The service layer provides API access to various data operations and AI-model services. The API's are used by apps in the app layer, but can also be used by external parties to integrate in their own apps.

Model layer

This layer houses both the generic models that can be used by the public as well as custom machine learning models that have been created by and for specific customers.

Data & knowledge layer

In this layer, raw and unstructured data along with available (human-generated) labels are ingested, structured and managed by the platform.

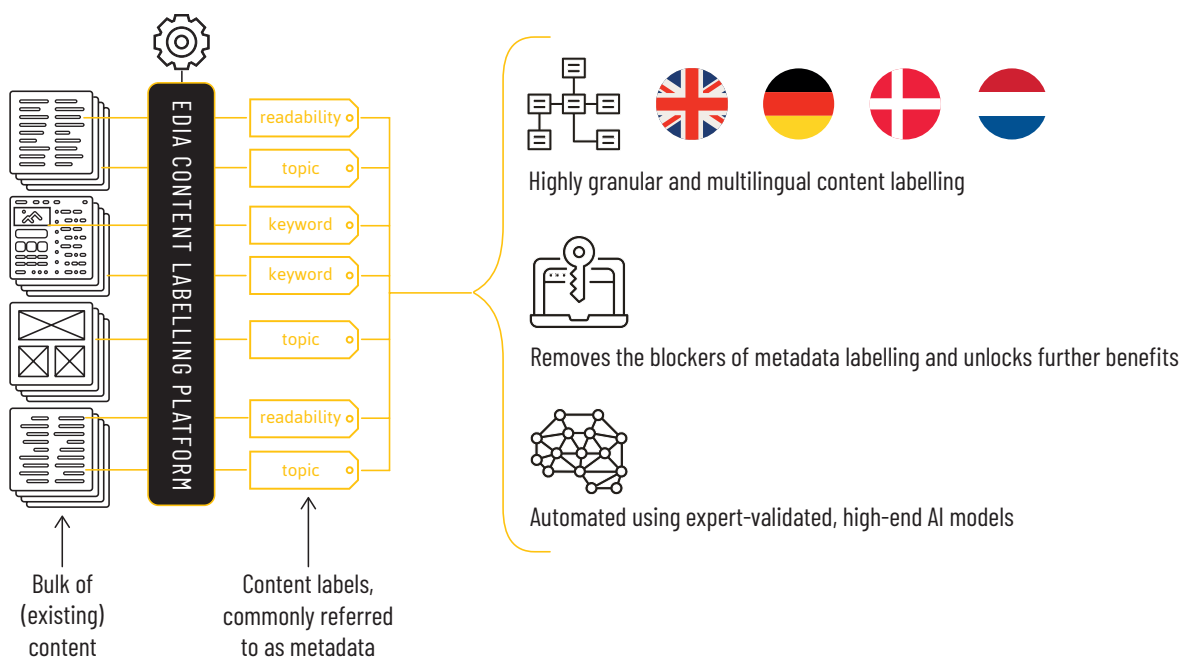# How do publishers and content creators engage with the platform

Users engage with the platform in various ways, as consumers of turnkey apps or as developers using our tools and services to produce bespoke content labelling solutions:

Use an app or create an app by yourself

EDIA offers out-of-the-box applications that make use of AI models and we have created integrations with several apps from leading solution providers. Apps in this layer support:

- **Content Production**
  For example authoring tools that support authoring/editorial processes

- **Content Management**
  For example tools that integrate with content management systems

- **Content Recommendation**
  For example tools that integrate with learning management systems

- **Data Science**
  A suite of tools that support AI model building

At the same time, customers have the ability to develop bespoke apps and make use of existing or custom API's as building blocks for their own applications.

Highly granular and multilingual content labelling

Removes the blockers of metadata labelling and unlocks further benefits

Automated using expert-validated, high-end AI models

Bulk of (existing) content

Content labels, commonly referred to as metadata

Our solution: an accurate and scalable content labelling platform

Use our APIs' or create a custom API

Machine learning models for content labelling can save a lot of time and reach a high level of accuracy. The machine learning models models on our platform can be used to process large amounts of textual content, and can be easily integrated into your own application for your own purpose. If you for example use a CMS for content and you want to extract keywords for every piece of content, you can set up an existing API to do this job for you. But also if you need a custom api using a combination of machine learning models and/or knowledge graphs, and we help you define a custom API. This can be very useful if your organisation already uses a knowledge graph that needs enrichment with additional data. Our API based services include:

Typical API based services include:

- **Content Analysis Services**
  Extract relevant educational metadata from text-based content.

- **Content Search Services**
  Enables retrieval of content given a topic, learning objective etc.

- **Content Matching Service**
  Given a user profile, this service can recommend relevant content.

- **Data Services**
  Query educational data such as curricula or thesauri.

- **Data Science Services**
  Enables labelling of data to support the creation of custom classification models.

Build your own content labelling model

Once you have a solid dataset and storage (e.g. a knowledge graph) you can start automating the data labelling using AI models. Traditional machine learning development is a complex, expensive, iterative process made even harder because there are no integrated tools for the entire machine learning workflow. You need to acquire know-how and glue together tools and workflows, which is time-consuming and error-prone. We solve this challenge by providing all of the components used for machine learning in a single toolset so models get to production faster with much less effort and at lower cost. To get you started even quicker, a number of pre-trained models are available for you to serve as a kickstarter for your specific content labelling needs.

Setup and manage data (in a knowledge graph)

Organizations often possess large amounts of content that require mapping against a domain model or taxonomy (e.g. a curriculum). Labelling of content against domain model or taxonomy is usually done by human domain experts and both taxonomies and labels are often inadequately archived and hard to maintain. We solve this challenge by providing you the tools and expertise to support content and data ingestion as well as a safe place to create and maintain such data.

When to use the platform

Our customers are typically undergoing a transition from printed content to digital content, or they may be already looking at ways to offer interactive digital content. Others might even create adaptive content for personalised delivery. For each of these stages in the content transition, the EDIA content platform will offer means to support the tasks at hand.

Stage 1.
From print to digital

A key element in the transition from print to digital is content management. A content management system combined with well-labelled content can greatly improve internal workflow as well external discoverability (SEO). The EDIA platform can play a role in this stage, as it can provide you with automated content labelling capabilities, which can be directly integrated into your content management solution. For example, using EDIA's keyword automation inside your cms, you will be able to quickly retrieve existing content for re-use in the authoring process, or make sure your content can be discovered by teachers and students through search engine optimization (SEO).

Stage 2.
From digital to interactive

Our customers use the content labelling platform when they produce and manage interactive content. Interactive content is modular and often multiple variations of content need to be created and distributed. Because of its modularity, interactive content content management is more complex and involves additional (user) data which needs structuring. Structuring data can be done using a taxonomy, which in turn can be stored in a knowledge graph. The content labelling platform offers the necessary tools to build a knowledge graph and subsequently populate it through automated content labelling.

Stage 3.
From interactive to personalised

In the third stage, interactive content becomes dynamic content and next to content management, content delivery becomes a point of focus. Dynamic content, also known as smart content, is atomic, context-free and it can change and adapt dynamically, depending on the reader. With this ability, content can be effectively targeted and create a personalised experience. Atomic content parts can only be delivered in a personalised way, when the right labels are applied, a task that is automated through the EDIA platform.

Do you have questions or is something unclear? Don't hesitate to contact us.

+31 20 716 3612   info@edia.nl   www.edia.nl