

OVERVIEW GUIDE:

Innodata's
Full Suite of
Data-Centric
Services &
Platforms



WHAT YOU NEED TO KNOW

Frequently Asked Questions

What Services Does Innodata Provide?

- Innodata provides a comprehensive suite of data engineering services for end-to-end artificial intelligence development, including data collection services, annotation services, and AI consulting.
- Innodata's web-based, SaaS platforms are designed to reduce the cost of AI/ML projects while enabling users to develop more accurate models.

How Can Innodata Help My Company?



Data Collection Services

Innodata specializes in sourcing proprietary and public datasets across any modality, vertical, and language. With over 40+ languages supported across the globe and customized offerings to meet any industry domain need, Innodata provides the diverse image, video, text, and voice data required for successful AI and ML model development. What can't be sourced can be created with our synthetic data services.



Data Annotation Services

Combining ML-enabled platforms and subject matter experts (SMEs), Innodata delivers the highest quality ground-truth training data. Leveraging a sophisticated ML-powered workbench, our 4000+ SMEs seamlessly ingest, assess, and synthesize all data types to generate pristine text, image, video, or audio datasets.



AI Services

Innodata's custom AI implementation services merge human expertise and machine learning technology to derive business intelligence from unstructured documents. We enable end-to-end workflow automation of document-based processes to reduce waste, increase value, and deliver insights for informed decision-making.



Data Annotation Platform

Innodata's web-based self-serve data annotation platform supports the creation of expertly labeled and annotated data at scale. Users can upload highly complex documents and unstructured content to generate more informed insights, improve model accuracy, and deliver better customer experiences.



Document Intelligence Platform

With Innodata's Document Intelligence ecosystem, users can extract data from any complex document in seconds. Powered by industry-leading NLP and OCR ML algorithms, users' models will continue improving in accuracy and confidence level.

WHAT YOU NEED TO KNOW

Frequently Asked Questions

Why Choose Innodata?

- Innodata's **AI-driven platforms** deliver cutting-edge tooling for all data types and unparalleled control of collection, annotation, and automation workflows that can integrate seamlessly into your existing workflows. New time-to-value enhancers are constantly developed to augment and streamline work. Systems are production ready to provide infinite scale.
- Innodata offers **multi-pass arbitrated annotation** by experts-in-the-loop to drive better model performance.
- Innodata's staff of **4,000+ in-house SMEs** hold advanced degrees in healthcare, medicine, sciences, finance, and law and are proficient in 40+ languages. To ensure quality, SMEs complete a rigorous training program with use case-specific training modules and receive weekly audit reports, random QC spot checking, and performance evaluations based on KPIs.

What is Innodata's Customer Satisfaction Guarantee?

Innodata works with clients to set project service-level agreements (SLAs) that cover turn-around time, quality metric benchmarks, language, preferred locality, volume, and level of staff expertise to ensure success.

Does Innodata Provide Pilot Programs?

Depending on the volume and scale of a project, Innodata provides a pilot to scope and prove the ability to meet quality benchmarks.

Does Innodata Offer Special Programs for Startups?

To support qualifying startups, Innodata offers special discounts, timelines, and packages.

How Does Innodata Ensure Security and Privacy?

Innodata stays up-to-date with compliance measures and the following certifications: HIPAA | ISO/IEC 27001: 2013 | SSAE 18 SOC-1 Type-2 | DPA | Sarbanes Oxley | GDPR

How Are Projects Priced?

- Innodata offers flexible pricing models according to a Full-time Equivalent (FTE) model and a unit/volume-based model.
- Try the Data Annotation Platform free for 30 days as part of our trial program.

Data Collection & Creation

Capture, Source, & Generate High-Quality Data to Scale Model Development

Customized Natural and Synthetic Data Collection for AI Model Training

Let Innodata source, collect, and generate speech, audio, image, video, text, and document data for AI and ML model development. With 40+ languages supported across the globe and customized data collection and generation offerings to meet any industry domain need, we offer a one-stop-shop for web, internal, or external data collection and generation.



Global Footprint & Delivery



Specialized Domain Expertise



Data & Sample Requirement Flexibility



Your Data, Secured



Text and Document Data Collection and Creation Services

With Innodata's suite of text and document data collection and synthetic generation services, you can scale your AI models and ensure model flexibility with high-quality and diverse text data in multiple languages and formats. Let Innodata's global network of 4,000+ experts, including native speakers of 40+ languages, collect or create the data you need for any initiative.



Image and Video Data Collection and Curation Services

Innodata's image and video data collection and creation services allow you to ensure model flexibility and scale your AI models with high-quality and diverse data in multiple formats and languages. Let Innodata's global network of 4,000+ experts, including native speakers of 40+ languages, collect or create the image and video data you need for any initiative.



Speech and Audio Data Collection and Creation Services

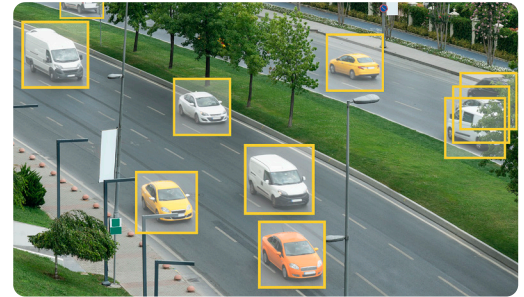
With Innodata's full suite of audio and speech data collection services, you can scale your AI models and ensure model flexibility with high-quality and diverse data in multiple languages, dialects, demographics, speaker traits, dialogue types, environments, and scenarios. Let Innodata's global network of 4,000+ experts, including native speakers of 40+ languages, capture the samples you need for any initiative.


Data Annotation

With 4,000+ Subject Matter Experts and Production-grade Platforms,
No Job is Too Big or Complex

High Quality Training Data to Accelerate AI and Machine Learning

Highest-quality annotation of text, images, audio and video data for complex models. Ideal for computer vision, sentiment analysis, entity linking, text categorization, and syntactic parsing and tagging models.



-  **High Accuracy**
World-class workflows and easy-to-use interfaces coupled with process control techniques developed over 30 years building the world's highest quality data products.
-  **Efficient, Fast Results**
World-class workflows and easy-to-use interfaces coupled with process control techniques developed over 30 years building the world's highest quality data products.
-  **Domain Experts**
World-class workflows and easy-to-use interfaces coupled with process control techniques developed over 30 years building the world's highest quality data products.
-  **Data Security**
World-class workflows and easy-to-use interfaces coupled with process control techniques developed over 30 years building the world's highest quality data products.

Data Annotation Services

From complex documents to vital images, our experts precisely tag the data you need to train your models.

- **Image & Video**
- **Text**
- **Speech & Audio**
- **Content Moderation**
- **Conversational AI**

Deep Expertise Across Diverse Data Formats & Industries

- **Contracts**
- **Legal Data**
- **Financial Reports**
- **Patent Data**
- **Scientific Data**
- **Medical Records**
- **Image & Video**
- **Audio**
- **Insurance Claims**

Applied AI

Guiding Clients From Ideation Through Implementation With AI-Powered Technologies & Advisory Services









The Applied AI Method

The Trusted and Proven Process For Implementing Effective Transformation

Innodata deploys best-in-class third-party and proprietary data collection, annotation, and extraction AI/ML models and platforms to transform your data flows and business operations. Our experts leverage our proven methodology to help automate operations, speed up manual tasks, and shift your expert talent to creative and analytical work.



You Bring the Use Case, We Bring the Expertise

- | | |
|--|--|
|  Redesign Organizational Structure to Provide Value Realization |  Establish a Business Case to Achieve Enterprise Goals |
|  Remove Paper Processing With Intelligent Document Extraction |  Provide Self-Serve Customer Support With Conversational AI (Chatbots) |
|  Utilize Prebuilt Synthetic Data for Faster AI Model Training and Deployment |  Capture, Source, and Generate Data to Scale Model Development |
|  Identify Content of Image, Video, Sensor, Speech, Audio, and Text Data With Data Annotation |  Build Trust and Improve User Experience With AI & Human-Powered Content Moderation |

Accessible AI Transformation

Eliminate boring and mundane tasks like data gathering, data crunching, and document processing so your valuable human experts can shift to higher-level tasks that require analysis, creative thinking, judgment, and other skill sets.

By leveraging our experienced consulting team, you can start thinking differently about how to solve business challenges with our AI-based platforms and solutions.

- **Improve Customer Experience**
- **Scale Business Operations**
- **Ensure Market Flexibility**
- **Substitute Manual Workforces**
- **Increase Efficiency**
- **Make Informed Decisions**
- **Align Enterprise Ambitions**
- **Combine Technologies**

Annotation Platform

Create Training Data for Advanced AI Systems

Four Customizable Workbenches

With Innodata's Annotation Platform, you can tackle your most challenging data problems — no matter the data type. Easily configure multi-pass annotation, categories per record or document, and arbitration. Quickly allocate work, designate annotation tasks, and track progress within the platform. In addition, our backend integration with our proprietary algorithms allows for a seamless, no-code solution for annotation work and model building.

Auto Annotation

With Innodata's no-code solution, you can build a model from scratch, bring your own model, or connect to an open-source infrastructure in a highly-secure environment. In addition, with our four years of neural network training, our backend quality checks detect and eliminate bias in your data, while deep active learning continuously improves the accuracy of your output.



Document Classification

Maximize readability for annotators and enable high-level labeling of text-based documents.

**File indexing, meta-data extraction, news extraction, etc.*



Record Classification

The most efficient table-style annotation workbench on the market. Effortlessly train a model to auto-label data organized into rows and columns.

**Financial data, chatbots, sentiment analysis, etc.*



Inline Classification

Combine the most powerful NLP technologies for Named Entity Recognition (NER) and Relationship Extraction. Label words or phrases in text data, then identify relationships between them.

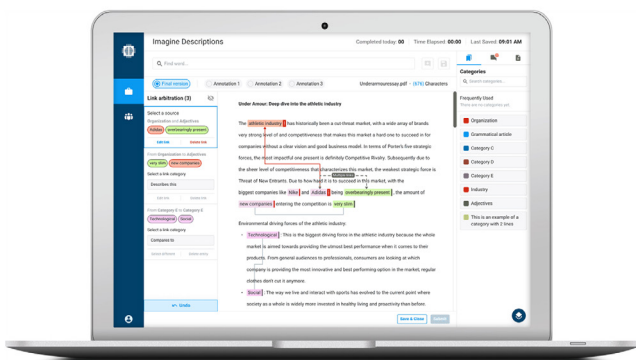
**Legal documents, financial applications, deep analysis, etc.*



Image Annotation

Provides annotators a clear view to create bounding boxes and point labels.

**Drone footage analysis, inventory stocking, traffic recognition, etc.*



API Integrated Real-Time KPIs and Data Management

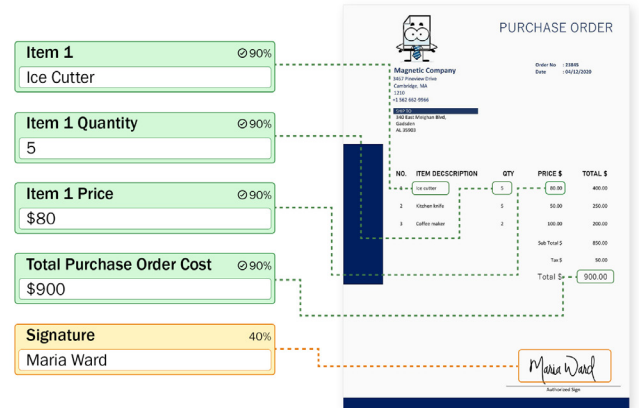
Ensure performance and quality with access to real-time metrics throughout project lifecycles, monitor agreement ratings, progress, efficiency, and more. Easily manage and maintain a queue remotely with our feature-rich API. Create and manage customizable taxonomies so your team always uses the best data possible.

Document Intelligence





Turn Your Document Data Into Intelligence With the Leading AI-Powered Data Extraction Platform

Powerful, High-Confidence Extraction for Your Document-Heavy Operations




With Innodata's Document Intelligence ecosystem, you can extract data from any complex document in seconds. Powered by industry-leading NLP and OCR ML algorithms, your models will continue improving in accuracy and confidence level.






Industry-Leading Platform Features

-  Advanced Extraction
-  Multiple Languages Supported
-  Handwriting Processing
-  Data Normalization & Language Generation

Integrate Seamlessly

-  Easy-To-Use
-  Data-Centric Approach
-  Synthetic Data Ready

Human-Supported Operations

-  Humans-In-The-Loop & SMEs
-  Internal Data Science Teams
-  In-House Advisory

Advanced Extraction Capabilities

Layout Analysis With Computer Vision and NLP
Table and Form Parsing
Reading Order Detection

- Image Cleaning
- OCR, Language Detection, and PDF Text Extraction
- Dehyphenation
- Data Point Extraction, Normalization, and Linking
- Data Point Generation / Re-Writing
- Clause and Paragraph Tagging