Real-World Use Cases for Human-Annotated Data
Investment in AI in 2016 was in the range of $26 billion to $39 billion. Yet, even with the huge advances made in AI solutions in the last decade, and the growing number of them on the market and in our lives, this basic truth holds: AI is only as good as the machine learning data that trained it.

To build a successful solution, you need the right data, and a lot of it. You can have the most appropriate algorithm, but if you train your machine on bad data, then it will learn the wrong lessons, come to the wrong conclusions, and not work as you (or your customers) expect. Your success is almost entirely reliant on your data.

As enterprises race to stand out from the competition and do more with the same resources, structured, human-annotated data can help optimize their machine learning efforts to achieve their business goals.

Leaders in Machine Learning are adopting the following use cases to improve the end user experience for web search, eCommerce sites and social media platforms:

**eCommerce**
- Auto-fill Evaluation
- Content Moderation
- Defect Testing
- Entity Tagging
- Field Testing

**Site Search**
- Query Relevance
- Side by Side Evaluation
- UX testing

**Web Search**
- Ads Relevance
- Adult Search
- Caption Evaluation
- Categorization
- Document Quality
- Field Testing
- Fresh/News Evaluation
- Map Verification
- Media Search
- Multimedia Search
- Query Relevance
- Related Search
- Side by Side Evaluation
- Social Search
- Spam Junk Page Evaluation
- Spelling
- Whole Page Measurement

**Social**
- Personalization
- Post Interest Tagging
- Query Classification
- Query Relevance
- Refined Search Evaluation
- UX Testing
- Video Categorization
Data with a human touch

High-quality data for machine learning, enhanced by human interaction

Your machine learning models need training data to make sure that your algorithms are continually optimized. But not just any data—high-quality, human-annotated data. Humans are simply better than computers at managing subjectivity, understanding intent, and coping with ambiguity.

Data Collection

Scale your data collection efforts across multiple file formats including text, image, video, speech, and audio.

Data Annotation

Annotated data enables richer and more valuable machine learning-based products. Appen's curated crowd allows you to get the high-quality data you need to develop better products for your customers.
Appen at a Glance

Appen is a global leader in the development of high-quality, human-annotated datasets for machine learning and artificial intelligence. Appen brings over 20 years of experience capturing and enriching a wide variety of data types including speech, text, image and video. With deep expertise in more than 180 languages and access to a global crowd of over 1,000,000 skilled contractors, Appen partners with technology, automotive and eCommerce companies—as well as governments worldwide—to help them develop, enhance and use products that rely on natural languages and machine learning.

Coverage for over

180 languages & dialects

Access to a global crowd of over 1,000,000

20+ years of experience providing high-quality, human-annotated data for machine learning for leading solutions in speech, search, eCommerce and more

"We needed to offer consumers more personalization on our site. Partnering with Appen allowed us to collect high-quality training data for our machine learning model, which refined our algorithm much more quickly and consistently than we’ve been able to in the past."

- Major international software provider