How **Not** to Train Your Algorithm

When AI projects fail, it's most often because the machine learning algorithm never got trained properly.

Disciplines like **computer vision**, **natural language processing** and **entity resolution** require massive amounts of data. And achieving acceptable levels of model confidence demands meticulous labeling and annotation.

DIY Training Data: Is It For You?

Many organizations assume they can produce ML training data in-house. Then, they discover

1 **DIY requires specialized technology.** Software tools for tracking and managing huge numbers of images, video frames, text fragments. Tools for defining data labeling or annotation tasks, and for assigning those tasks to people. Tools for evaluating human decisions and determining the need for further review or adjudication.

2 **DIY involves human labor.** Human judgement is essential to structuring ML training data. Large volumes of data call for many people, some or all of whom may need particular skills or qualifications.

3 **DIY demands particular project skills.** Skills needed to root out model-wrecking bias in the data. Skills required to design tasks that promote efficiency, scalability and quality. Skills needed to oversee iterative and very fluid processes.

We build ML training datasets for the enterprise. So, if you *don't* have the software, people and skills to do training data in-house, we should talk.

> We have our own **Training Data Platform** and a global pool of nearly a million **data specialists**. And our **project leads** offload the entire process, leaving you with only your algorithm to worry about.

> Our **platform** qualifies and groups specialists and distributes work to them. At any moment, there are thousands of specialists working on the platform.

> It also trains specialists on tasks, evaluates their progress and monitors them as they work.

> The **AI in our platform** makes specialists faster and more accurate. In video annotation tasks it makes one specialist the equivalent of 120.

> And our **project leads** are experts at designing tasks that promote efficiency, scalability and model confidence.

Let us demonstrate why we are an obvious alternative to DIY Training Data. We make it very easy to get started.

And if you're done with assigning data labeling tasks to expensive data scientists, you can connect with us at **content.alegion.com/nomoreDIY**.