

How Data Labeling With Datasaur Can Save Your Team Time

September 2022 | Case Study



01 Introduction

Manually labeling data can be a massive time sink. Whether through working in spreadsheets or building your own solution, annotating data and building accurate training models can be incredibly time inefficient. Manual data labeling devours important time that could otherwise be spent focusing on all the other things that go into building a great product. Enter Datasaur. Datasaur is proven to save valuable time in data labeling and management, which frees up your time, energy, and bandwidth. Labeling automation, advanced workforce management options, and automatic project creation and exportation options all drive down project times so that your team can spend less time annotating text and more time building models.

02 The Time Challenges of Data Labeling

Labeling data without a specialized tool drains time. Spreadsheets are completely manual. Managing and annotating bulk data in spreadsheets can be laborious, tedious, and prone to error. Developing your own in-house companion to—or alternative to—spreadsheets takes a lot of engineering time. This is not only for development, but also for build, ideation, iteration, and maintenance.

Let's take a look at workforce management, too. Workforce management without automatic reporting systems is laborious and relies on manual timestamps for labelers. This can lead to inaccuracies, and it can also waste a lot of team time in self reporting and corresponding review and QA. Overall, manual data labeling can lead to poor efficiency in annotating training data, difficulty identifying issues and inaccuracies in data, and suboptimal time per task.

03 Why Customers Choose Datasaur

Datasaur expedites data labeling in a few ways:

1. Ease of use

Datasaur's intuitive UI lets you set up projects quickly. Features like hotkey shortcuts are designed with the power-user in mind, so that you can label quickly and accurately without taking your hands off the keyboard.

2. Labeling automation

The ability to upload prelabel sets cuts down time significantly, with options for ML-Assisted labeling through tools like Hugging Face or by importing your own custom model. Similarly, bulk labeling is available via Datasaur's search extension, which lets you search for keywords and label all the instances across all your files.

3. Workforce management

With advanced workforce management features you can track team efficiency, quickly monitor project completion, assess time per project (and see if there are projects that are running slower), track labels applied, and see conflicts resolved for the team. What's more, you can zoom out for team metrics and then zoom into each specific member to see the same efficiency metrics. Gone are the days of relying on self reporting for time and labeling.

Why does this matter? If you can track things like which labels consistently have disagreements, you can clear up confusion with your labelers and train for tricky labels for the future. Plus, with built in dashboards you can gain high-level insights for all your projects at a glance, which saves managers time.

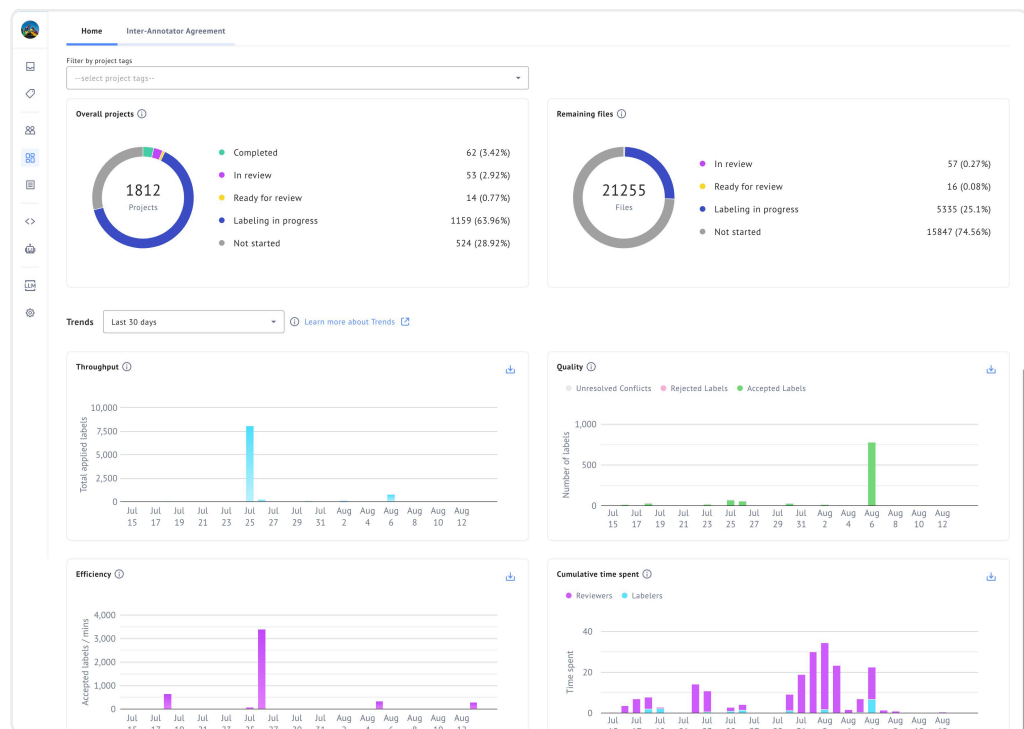


Fig. 1: Datasaur Workforce management feature

4. Automatic project creation

Automatically create projects, load templates, and export via API.



04 The Results

Datasaur streamlines data labeling and lets you get quality, labeled data fast. Customers use Datasaur as an efficient way to label data, identify issues and inaccuracies in data quality, and drive workforce scalability and management. Annotating training data doesn't have to be tedious, and the results speak for themselves. With Datasaur, customers experience:

80% automation

of data labeling

10x improvement

in project times

Features and prelabeling options let you automate 80% of data labeling, and hotkey shortcuts and other features help your team to work effectively through the remaining 20% of data labeling. All of this combines to significantly cut down time per task, freeing up vital hours to build models.

“ ...information labeling tasks has been reduced by 80% which has allowed us to optimize our workflow much more, allowing us to focus on other areas that are also priorities for US...”

— Mary L

“ Not only have they made labeling as painless as possible, they also help us manage and review our annotators' progress. Datasaur has saved us countless hours in building our own solution. My team lead never wants to go back to spreadsheets!

— Executive Sponsor in Computer Software



05 Conclusion

Datasaur lets your team build quality, trained datasets in a fraction of the time. Set up a custom demo today and we'll gather datasets from your team to show you how your unique and workflow requirements would translate in Datasaur. This lets you see how simple and time-efficient your data labeling workflow can be in practice.

Contact us at sales@datasaur.ai to schedule your demo!

About Datasaur

Datasaur is a private LLM provider and data labeling platform designed for companies to build their AI ecosystem with ease and efficiency. It assists organizations and universities in setting up custom LLMs and annotating data more efficiently and accurately through automation, quality control, and human-in-the-loop workflows. For more information, visit www.datasaur.ai.

Schedule a demo