

m47.ai info@m47labs.ai

#### Speed up your labeling operations

Bring your NLP projects to production faster by creating and managing your training data, labeling teams and projects in one single platform.



#### Generate

Generate high-quality training data at scale with our custom annotation UI and fully integrated solution.



#### Track

Track progress, review performance and monitor key indicators of large annotation projects and teams.

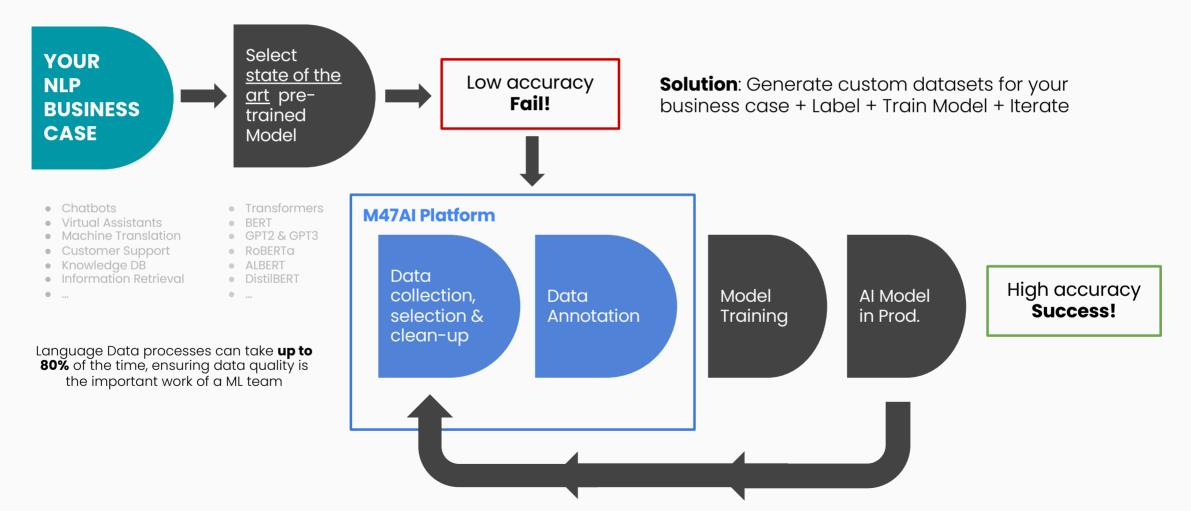


#### Automate

Reduce time and improve the quality of your datasets with Intelligent Automation

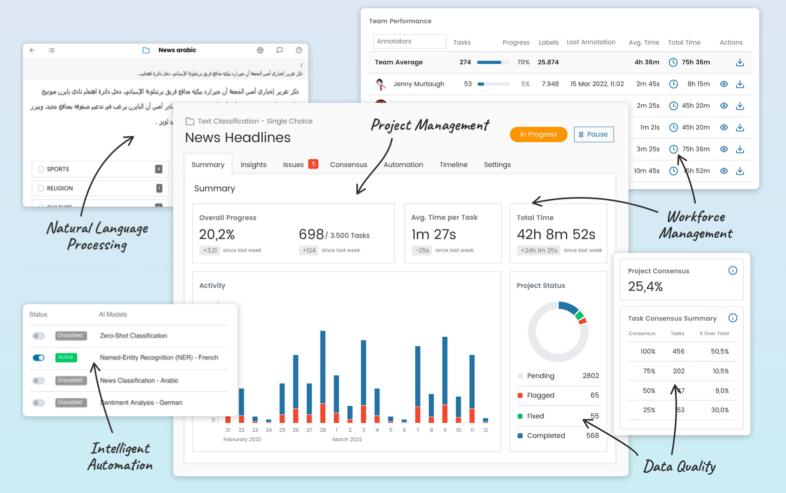
and ML features.

# The path to an **efficient** Machine Learning project



# **Our platform**

**Reduce Time and Cost** 2 Improve Model Accuracy 3 Consolidate Labeling Ops



#### Natural Language Processing

Dozens of supported NLP annotation projects that match every enterprise use case in one single tool. Multilingual. Modular and customizable UI.

(- ∷≣	🗅 News arabic 🛞 🟳 🕥
	ا دكر تقرير إخباري أمس الجمعة أن جيرارد بيكيه مدافع فريق برعلونه الإسباني، دخل دائرة العطيب.
نادي بايرن ميونيخ	ذكر تقرير إخباري أمس الجمعة أن جيرارد بيكيه مدافع فريق برشلونة الإسباني، دخل دائرة اهتمام
فوفه بمدافع جدید، ویبر ز	الألماني وأوضحت صحيفة (ببلد) الألمانية في عددها الصادر. أمس أن البايرن يرغب في تدعير صف
	Three years DATE × after Argentina GPE × was hit by
	a deadly MONEY × economic crisis, there is fresh hope. The country's
	economy is set to grow about 8% PERCENT × this year DATE ×
	after seeing 9% PERCENT × growth last year DATE × , a sharp
O TECH	turnaround from 2002 DATE × when output fell 11% PERCENT × . The
	unemployment rate is improving, too: It is set to slip below
	13% PERCENT × by the end of the year DATE × , down from
	20% PERCENT × in May 2002 DATE × . True, problems remain, but
	the overall picture is one of vast improvement. Even

#### **Intelligent Automation**

Reduce manual annotation time throughout ML-assisted and weak supervision labeling. Bring your own ML models to speed up preannotation and increase your training datasets with data augmentation techniques.

Project Cor		0	Task (	Task ID	Task Name	Consensus	~
25,4%							
			+	1	Bath and England ce		100%
Task Conse	ensus Su	mmary (i)	-	2	UK firms are embrac		80%
100%	456	50.5%			🦞 Jenny Murtaugh	Culture	
75%	202	10.5%			Morganne Flaherty	Sports	
50%	147	9.0%			Pedro Pérez	Culture	
25%	153	30.0%			🚇 Marina Rull	Culture	
2.5%	155	50,0%			Joan Santiago	Sports	
Annotators	5	(i)	+	3	Three years after Arge		100%
Annotator		% Agreement	+	4	Shares in Manchester Un		100%
Jenny Mu	urtaugh	75,8%	+	5	Lord of the Rings dir		80%
Morgann		63.4%		6	Coronation Street h		80%

#### **Workforce & Project Management**

Monitor annotator's stats, individual skill sets, scoreboards and cost metrics. Collaborate with large teams for a more productive effort. Let your team annotate, review, comment and quantify consensus and progress.

General Info Project	ts Insights	Issues T	imeline						
Projects									
Active Closed									
T 2021 Events							rogress		
Norred-Entity Recogni Overall Progress	ition	Tasks Done			Avo. Time o	er Task Total Time			
58,1%		698 / 1.254		P 14	lm 27		n 52s		
+1,7		+11.4		<b>Review Flogged</b>	-25s	+24h im 21	is.		
D Political State							Progress		
Political State Part of Speech (POS 1 Overal Progress		20				<b>B</b>	hogenst		
Part Of Speech (POS 1	o38,		Tosks	Progress	Labels	Last Annotation		Total Time	Action
Part Of Speech (POS 1 Overall Progress 81,7%	Performant	rs	Tosks 1.041		Lobels 25.874			Total Time	Action
Part Of Speech (POS 1 Overall Progress 81,7%	Performant Annotato	rs		- 70%			Avg. Time 4h 36m	36 75h 36m	¢
Part Of Speech (POS 1 Overall Progress 81,7%	Performante Annotato All project T NER	rs	<b>1.041</b> 53	- 70%	25.874	Last Annotation	Avg. Time 4h 36m 2m 45s	<ul> <li>75h 36m</li> <li>8h 15m</li> </ul>	¢

#### **Data Quality**

Optimize your training datasets with our Quality features: consensus metrics, error detection, label distribution, bias and outliers identification. Organize and visualize your training data in one single place. Import your pre-annotated datasets and perform QA on your Model performance.



## **ELG Benefits**

- 1. Language resources to reduce entry barrier for developing NLP projects in multiple European countries.
- 2. Evangelize and educate startup ecosystem about the importance of developing local language based AI systems.
- 3. Availability of **multipurpose language datasets** to implement state-of-theart ML Models

#### What does the ELG have to provide to strengthen our business?

- 1. Access to a pool of organizations and companies developing Language Technologies and in need of generating Training Data.
- 2. Connect needs from the ELG community to solution & service providers.
- 3. Support dataset and training data creation for **underrepresented languages**.

### Contributions to the ELG and its marketplace

- 1. Provide the **framework** for public institutions, Universities and organizations to **consolidate all their Language Data Labeling Operations.**
- 2. Offer **discounts and customized plans** to organizations tied to the ELG program.
- 3. Provide access to the ELG database directly from our platform so that resources (datasets and Models) can be used to generate new Training Data.

#### **Benefits for others**

- 1. Reduce development costs of custom annotation tools or maintainance of open source
- 2. Access to a production grade Data Annotation Platform that reduces the time and cost of creating, managing and labeling data at scale.
- 3. Easily reuse language resources from ELG for your Training Data needs.

## Motivation and goals to engage in the ELG, benefits and positive effects

- 1. Acquire knowledge working with R&D institutions that would let us implement new state-of-the-art functionality needed by leading organizations.
- 2. Increase platform usage, visibility and exposure to the European market.

#### The ELG wishlist

- 1. Improve the 'community' based functionality. It should be easy to create accounts, upload resources, post comments, fork models and datasets.
- 2. Increase ELG platform embedded functionality so it's not only a comprehensive database of resources. Models playground and API accessibility , view/edit/modify datasets, provide hooks to training data, etc.
- **3.** Increase promotion and adoption across european language/startup ecosystem

## **Steps in the right direction!**

ELG has the potential to become the 'defacto' platform for language data-centric Al in Europe and should continue to drive and support the local community in developing resources that can be adopted by SMEs across Europe in the mid term.



## Get in touch!

enric@m47labs.ai

m47.ai