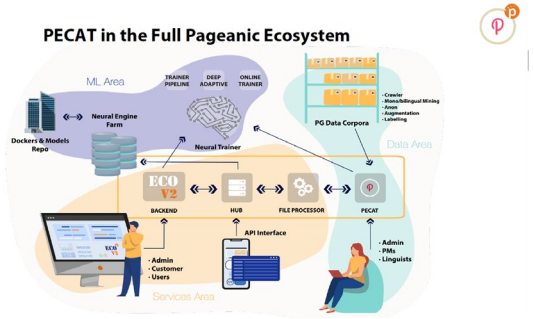
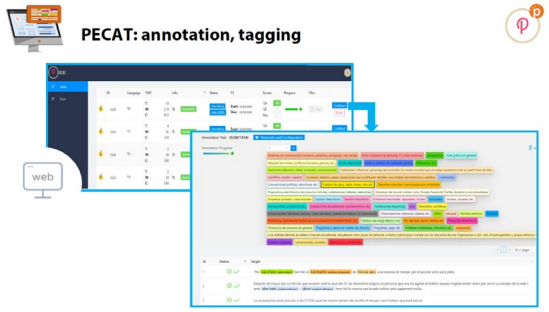
Workflow: From Data to Task-Specific Small Language Model

We are AI Language Technology
pangeanic

CLEANSE DOMAIN DATA

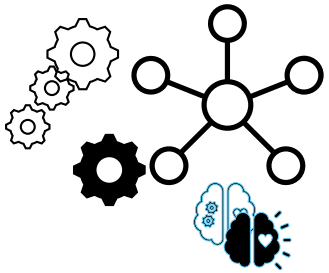


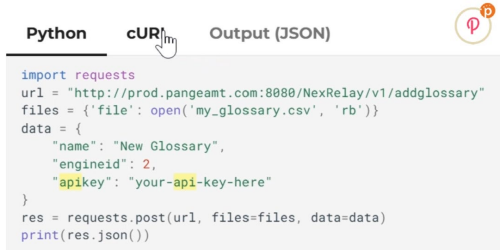
CURATE & LABEL FOR THE TASK



TRAIN / ADAPT SLM

RAG





DEPLOY & GOVERN

To cleanse domain data

- **Prepare** domain data for processing
- **Apply** governance controls
- (Optional) **Anonymize** sensitive data

To curate and label

- **Annotate** data for specific tasks
- **Identify** key features
- **Tasks**: Sentiment / Sensitivity, Probability, Classification, Translation, Summarization, Routing

To train and adapt an SLM

- **Optimize** the model with task-specific data
- **Utilize fine-tuning**, adapters (LoRA, etc.) RAG integration
- **Evaluation** harness for metrics

To deploy and govern

- **Launch** the SLM into controlled production
- **Establish** monitoring and controls
- **Human-in-the-loop**, Observability, Rollback, Versioning (v1, v2...)