



European Language Grid – META-FORUM 2020

Dominik Riemer (FZI Research Center for Information Technology, Germany)

ELG Pilot Project:

Streaming Language Processing in Manufacturing

SLAPMAN: Streaming Language Processing in Manufacturing

Dominik Riemer, Steffen Thoma
01.12.2020



Motivation

Exploiting text data in the manufacturing domain



Idea #1

Machine log
data

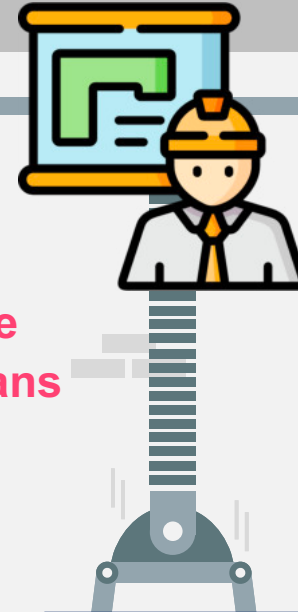
How to utilize
machine log data?

How can I conduct flexible
data analysis without deep
IT skills?

Idea #2

Service
technicians

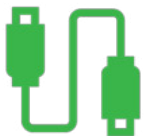
How to utilize
knowledge of service
technicians?



What is Apache StreamPipes?

„A self-service IIoT toolbox to enable non-technical users to connect, explore and analyze IoT data streams“

Developed @FZI since 2015, transition to Apache Incubator in 2019



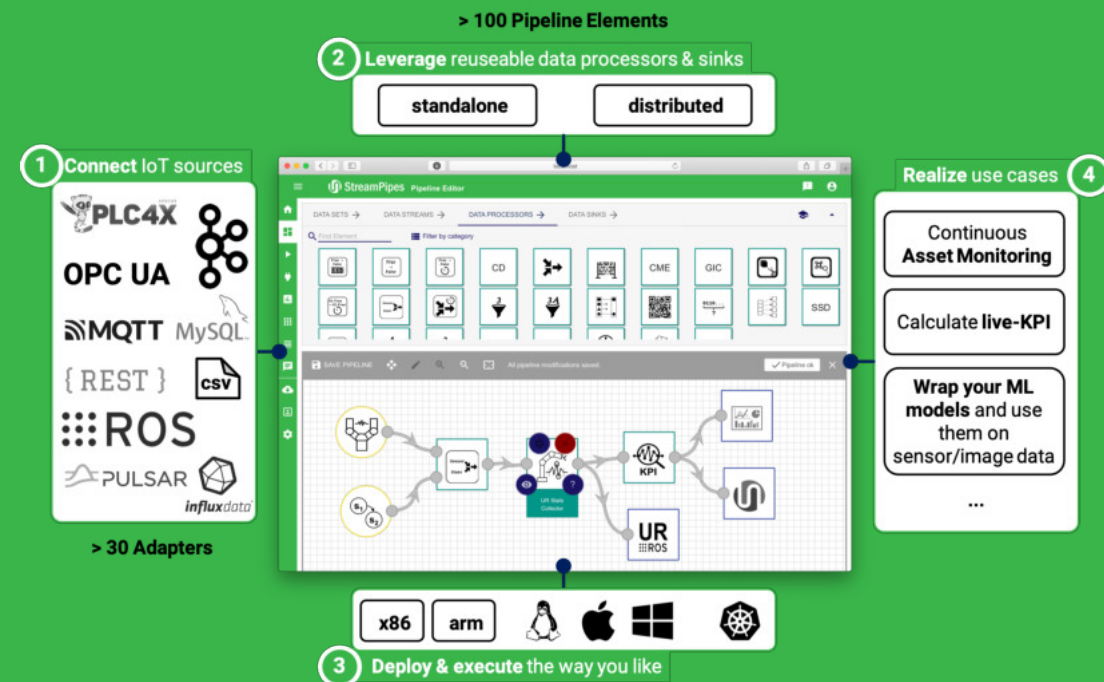
> 30 industrial data connectors to connect industrial assets within minutes



> 100 included algorithms for self-service data analytics



live dashboard & historical data explorer



Apache
streampipes



THE
APACHE[®]
SOFTWARE FOUNDATION

Project objectives

Investigating the role of industrial textual (e.g., log) messages in the manufacturing domain to improve product and process quality

Main Objectives



Develop/Integrate various text analytics features into Apache StreamPipes



Develop methods to easily train new text analytics models for machine assets



Provide a customized StreamPipes version to the ELG platform

1

ELG NLP Components for StreamPipes

- NER (BERT, Gate, Cogito), Sentiment, Translation, ...
- Integration with StreamPipes using the provided SDK and runtime wrappers

2

NLP Learning Component for StreamPipes

- Starting with an initial model, i.e. pre-trained general model or learned on a specified text corpus
- Continuous adaptation to the use case at hand via adaptation in regards to the incoming textual data stream

Let's connect!



streampipes.apache.org



streampipes.apache.org/docs



[apache/incubator-streampipes](https://github.com/apache/incubator-streampipes)



[@streampipes](https://twitter.com/streampipes)

... and leave a star on Github if you like it 😊

{riemer|thoma}@fzi.de

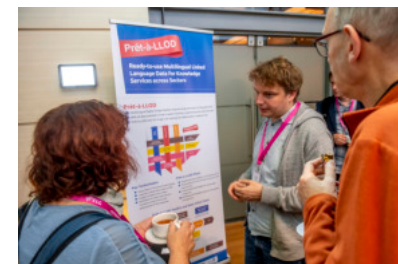


Day 1 Day 2 Day 3

1 December 2020 – META-FORUM 2020 Day 1

12:00 – 13:30 (CET)

	E3C (European Clinical Case Corpus)	Link to the Virtual meeting room
	LSDISCO (Lingsoft Solutions as Distributable Containers)	Link to the Virtual meeting room
	MKS-LLOD (MKS as Linguistic Linked Open Data)	Link to the Virtual meeting room
	Smart euSpeaker (Basque-speaking smart speaker based on Mycroft AI)	Link to the Virtual meeting room
	EVALITA4ELG (Italian EVALITA Benchmark Linguistic Resources for the ELG platform)	Link to the Virtual meeting room
	OPUS-MT (Open Translation Models, Tools and Services)	Link to the Virtual meeting room
	Text2TCS (Extracting Terminological Concept Systems from Natural Language Text)	Link to the Virtual meeting room
	PARA4DLM (Textual paraphrase dataset for deep language modeling)	Link to the Virtual meeting room
	YouTwinDi (Virtual Personal Assistant Prototype)	Link to the Virtual meeting room
	SLAPMAN (Streaming Language Processing in Manufacturing)	Link to the Virtual meeting room
	ELG (European Language Grid)	Link to the Virtual meeting room



Virtual Project Expo

<https://www.european-language-grid.eu/meta-forum-2020/project-expo/>



European Language Grid

Thank you!



The European Language Grid has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement № 825627 (ELG).

01/02/03-12-2020 META-FORUM 2020 – Piloting the European Language Grid (virtual conference)
<http://www.european-language-grid.eu>