



A Flexible System for Building Data Labeling Tools



Yu Zhang^{1,2}



Yun Wang²



Haidong Zhang²



Bin Zhu²




Siming Chen³



Dongmei Zhang²


¹University of Oxford ²Microsoft Research Asia ³Fudan University


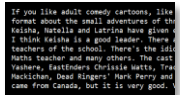

Data Labeling

- ML needs labeled data
- labeled data on *Paper with Code* 
 - > 30 modalities*
 - > 400 tasks*

* *Paper with Code*, <https://paperswithcode.com/datasets>, accessed 2022-03-18

Data Labeling

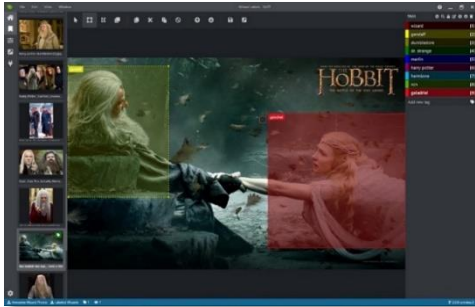
- ML needs labeled data
- labeled data on *Paper with Code* 
 - > 30 modalities*
 - > 400 tasks*

Modality	Task
Image 	Classify
Text 	Segment
Video 	NER
...	...

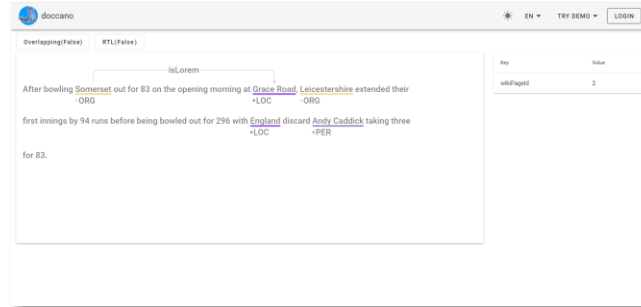
* Paper with Code, <https://paperswithcode.com/datasets>, accessed 2022-03-18

Data Labeling Tools

- Accommodate modalities and tasks



VoTT, for image



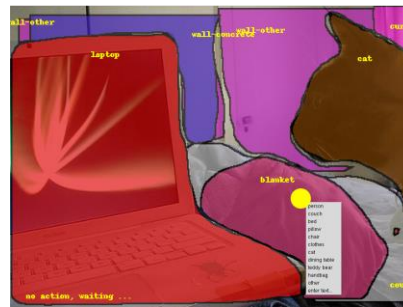
Doccano, for text

...

- Integrate machine-assistance



EasyAlbum, CHI 2007, clustering



Fluid Annotation, MM 2018, propose labels

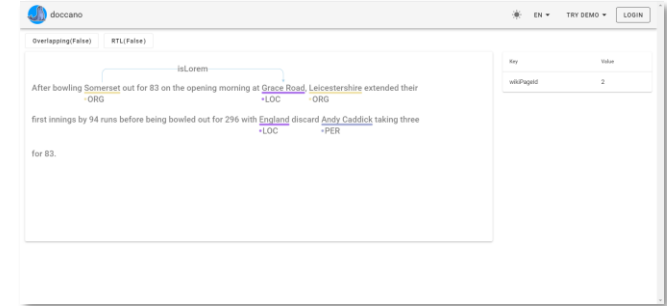
...

Required Skills

- Interface design
- Machine learning
- Software development
- ...



VoTT, for image



Doccano, for text



EasyAlbum, CHI 2007, clustering



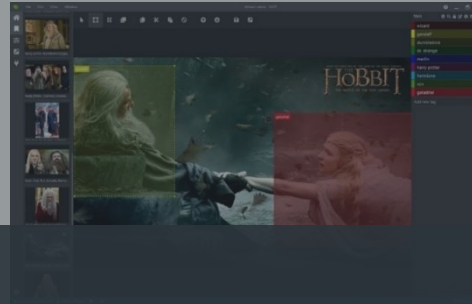
Fluid Annotation, MM 2018, propose labels

...

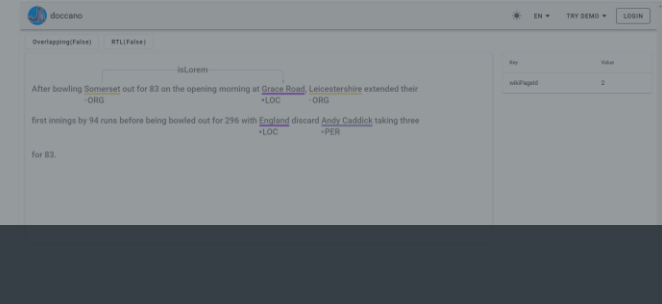
...

Required Skills

- Interface design
- Machine learning
- Software development
- ...

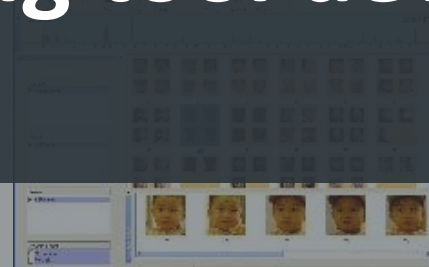


VoTT, for image

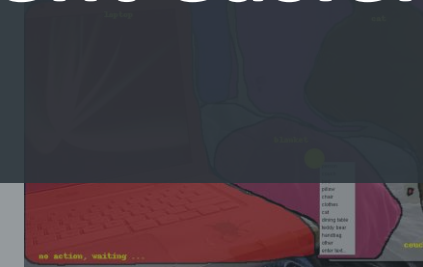


Doccano, for text

How to make labeling tool development easier?



ISSE, CHI 2007, clustering



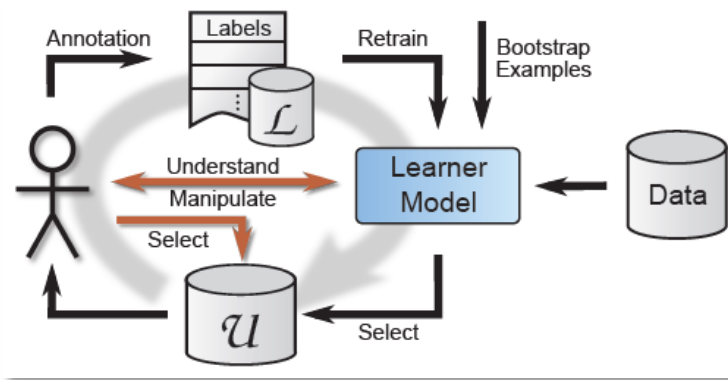
Fluid Annotation, MM 2018, propose labels

Design Requirements

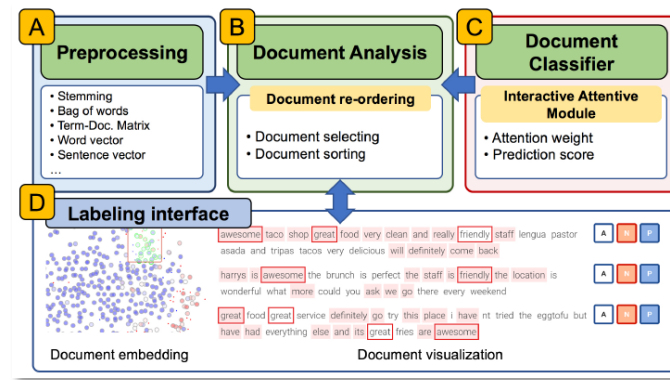
- R1: Unified module APIs
- R2: Module composition
- R3: Reuse modules/tools
- R4: Visual programming
- R5: Static checking

Identifying Reusable Building Blocks

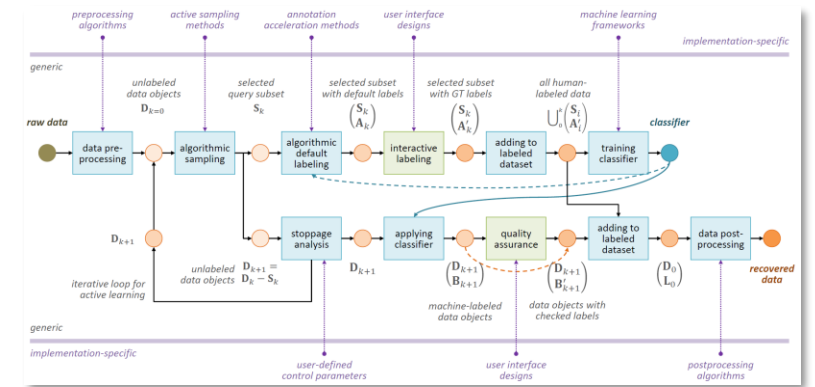
- Approach: Code existing modules in the literature
- Corpus: 36 flowchart figures



Höferlin et al., VAST 2012

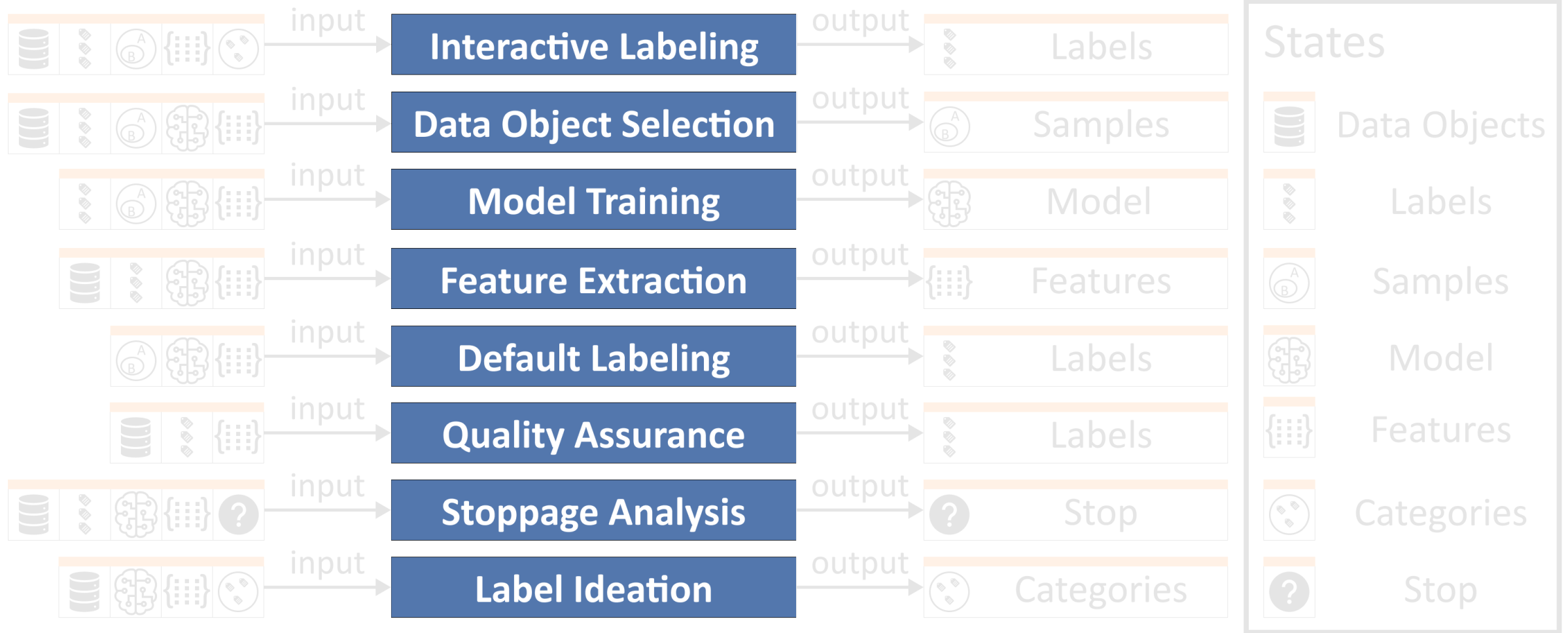


Cody, CHI 2021

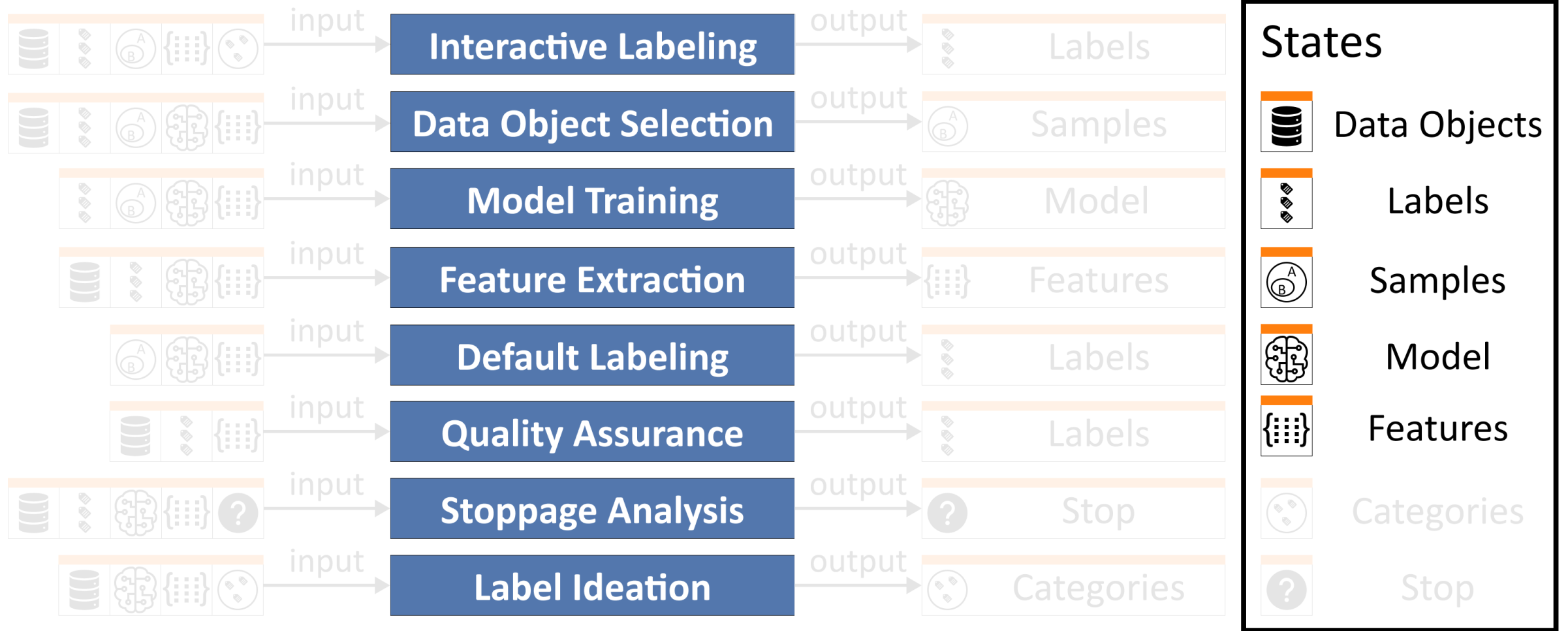


MI3, TIIS 2021

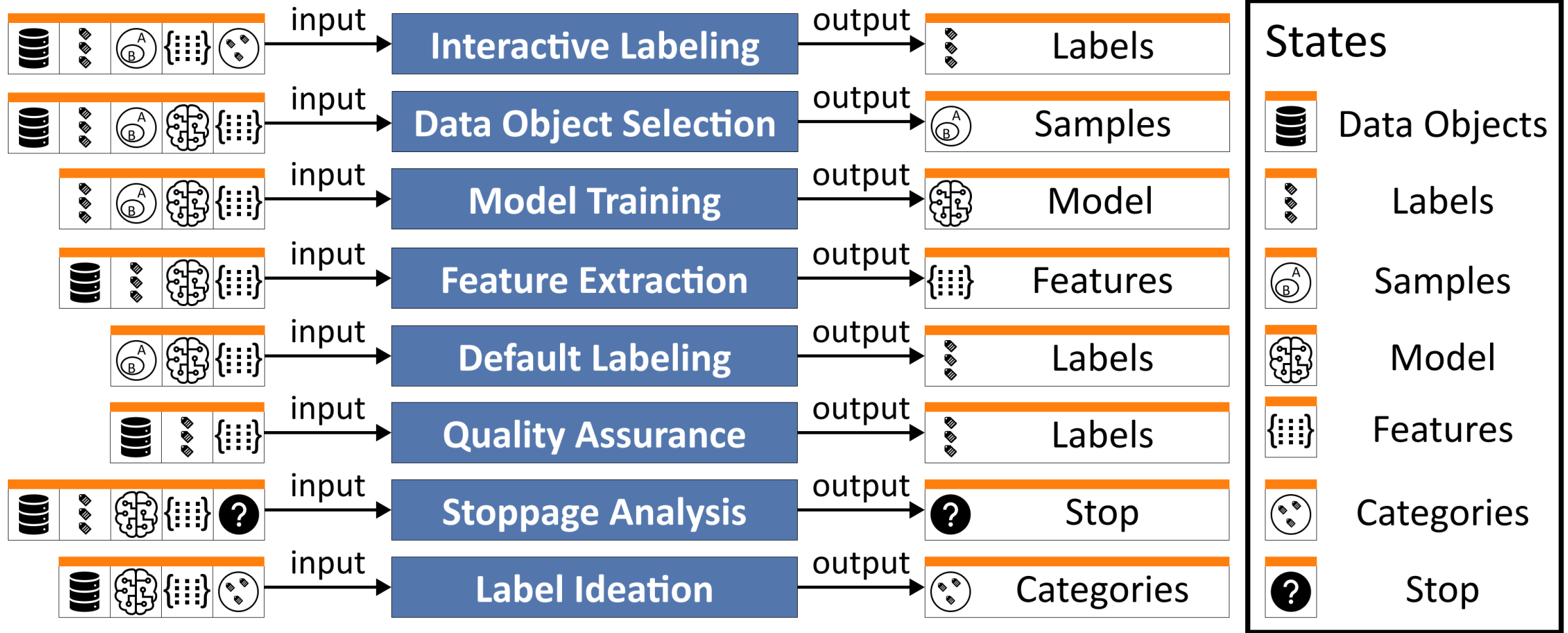
Coding Result



Coding Result



Coding Result



Coding Result



- **Data Object Selection:** decide data objects' labeling priority

Module Instances



- **Data Object Selection:** decide data objects' labeling priority
 - Implement as an **active learning strategy**

```
import numpy as np
from scipy.stats import entropy

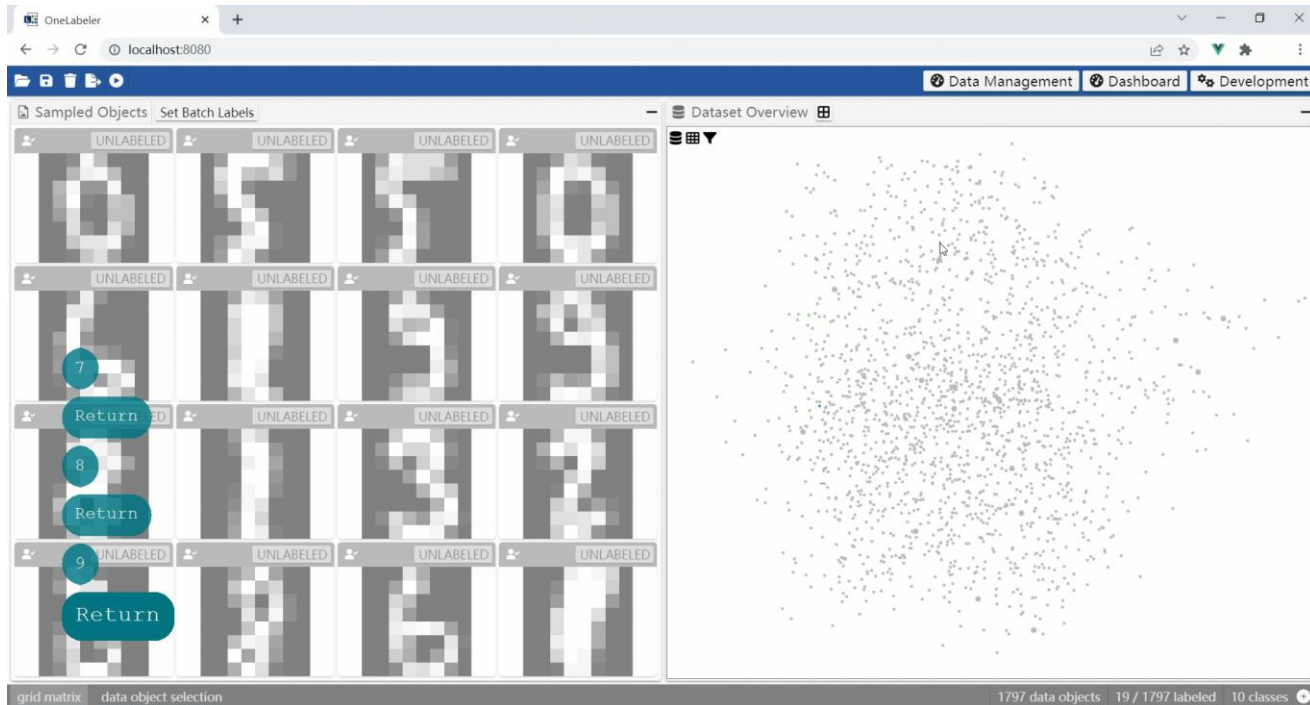
uncertainties = model.predict_proba(X)
entropies = np.transpose(entropy(np.transpose(uncertainties)))
idx = np.argmax(entropies)
```

Machine Computation

Module Instances



- **Data Object Selection:** decide data objects' labeling priority
 - Implement as an **interactive projection**



Human Computation

OneLabeler: A System for Building Labeling Tools

The image illustrates the OneLabeler system's workflow and components. At the top, a process flow shows a file named "my-labeling-tool Setup 0.1.0.exe" being compiled and then launched to create a "Generated Labeling Tool". This tool is shown as a window with a scatter plot of data points and a grid of sampled handwritten digits (0-9).

The main interface is the "OneLabeler" application, which features a "Workflow Canvas Panel" (A) for designing workflows. It includes an "Initial Workflow" (a1) and an "Edited Workflow" (a2). The initial workflow starts with "initialization", followed by "random sampling", "single object display", and "check all labeled". The edited workflow adds "SVD features", "clustering", "projection", "model training", "decision tree prelabel", and "interactive labeling" nodes, with decision diamonds for "stop?".

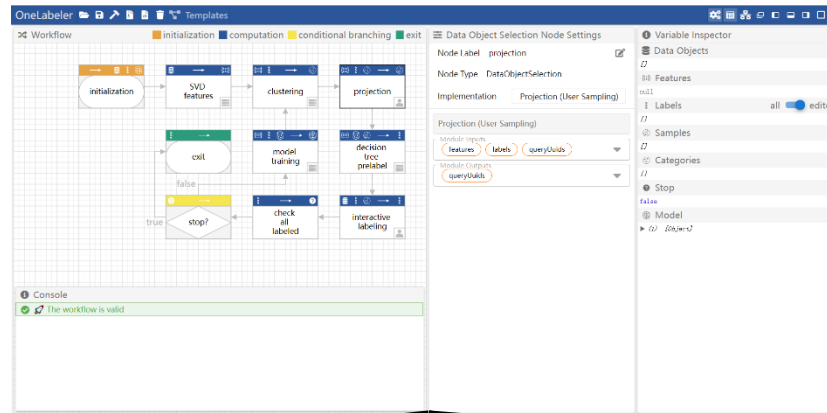
Other panels include the "Node Configuration Panel" (B) for "projection" (Node Type: DataObjectSelection, Implementation: Projection (User Sampling)), and the "Variable Inspector Panel" (D) showing state values for "Data Objects", "Features", "Labels", "Samples", "Categories", "Stop", and "Model".

A "Console Panel" (C) displays an "Example Error Message" (C1):

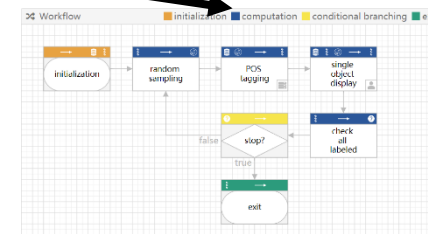
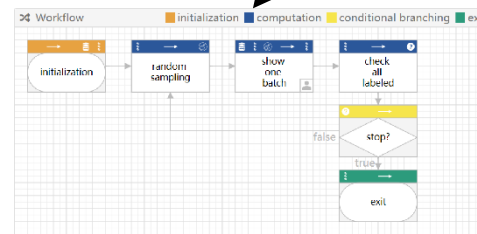
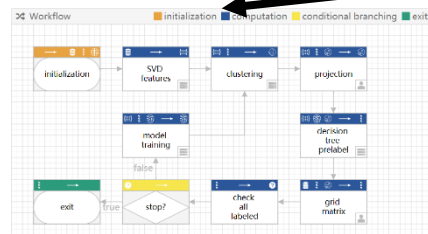
```
Console
The workflow is valid

computation node with label "projection" has indegree 0
Violated Rule: flowchart-node-indegree-range
Recommended fix: create an inward edge to the node with label "projection" to make its indegree nonzero
computation node with label "projection" has outdegree 0
```

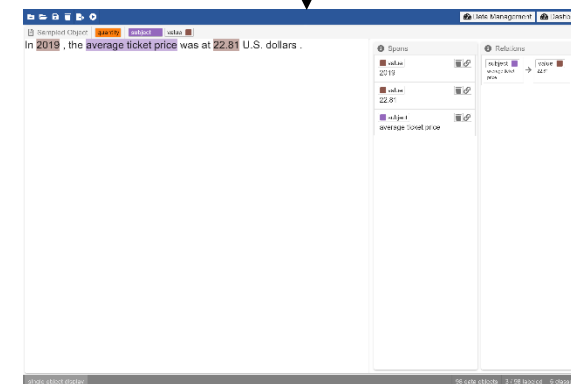
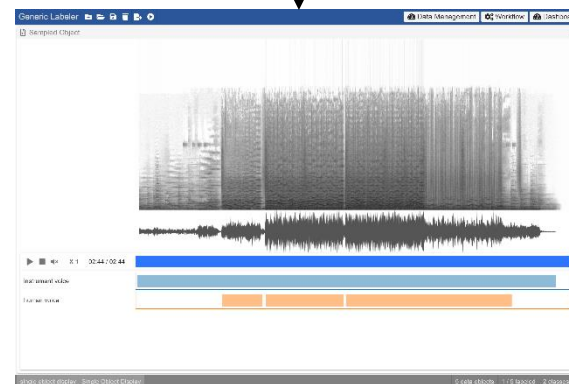
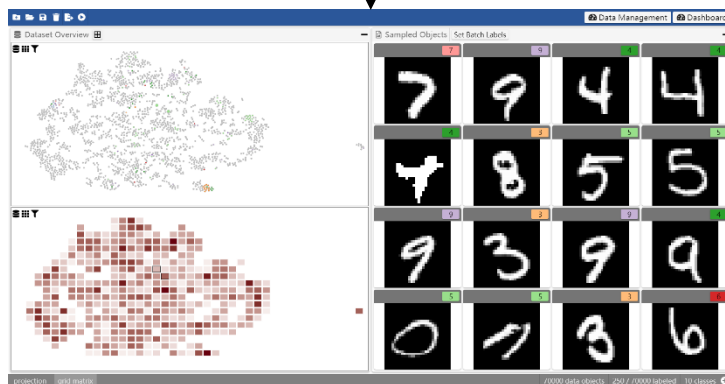
Visual Programming



Visual Programming Interface



Workflow Config



Labeling Tool

Built Data Labeling Tools

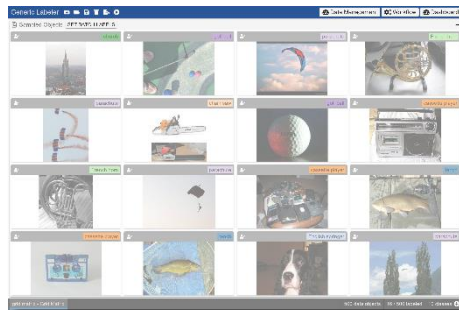


Image classification

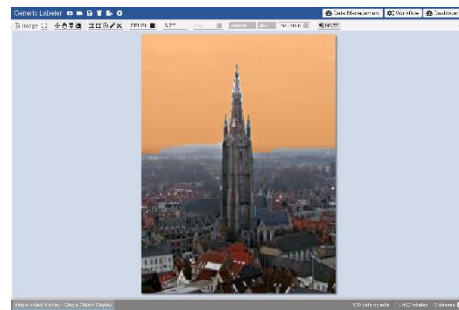
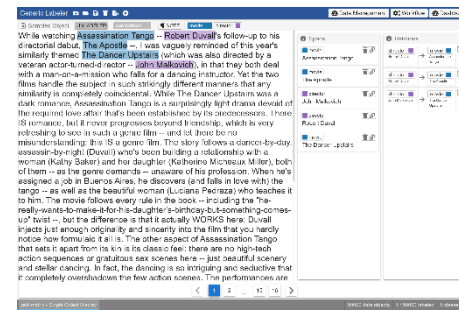
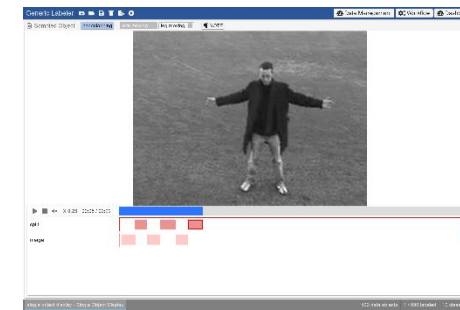


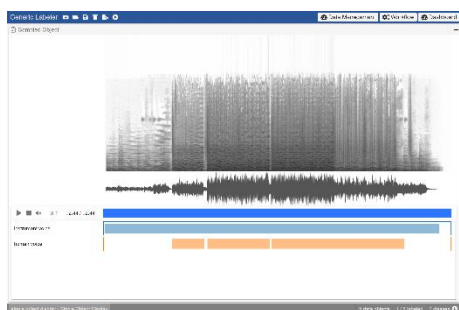
Image segmentation



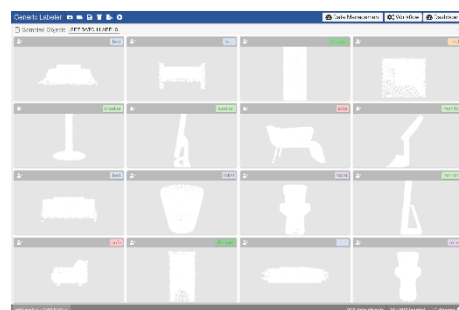
Text named-entity recognition



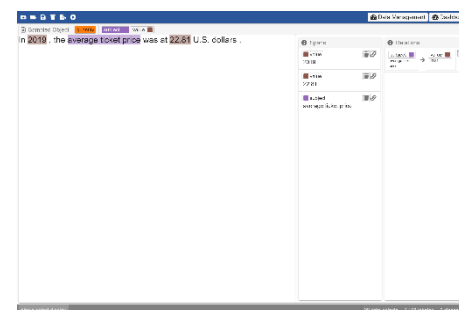
Video temporal segmentation



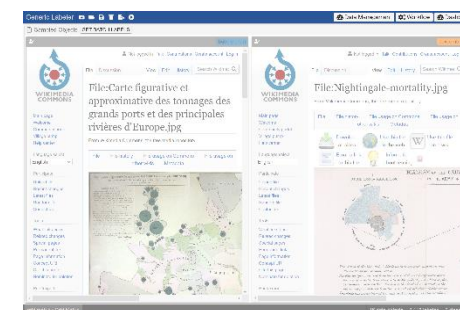
Audio temporal segmentation



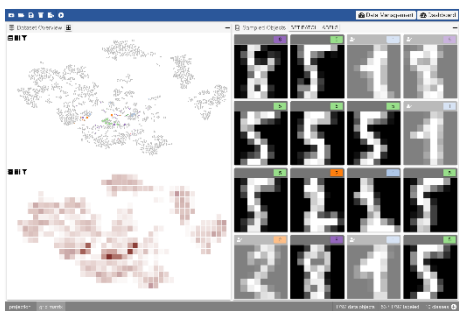
Point cloud classification



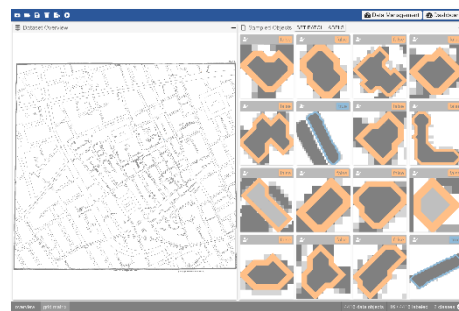
Machine-aided named-entity



Webpage classification



Mixed-initiative classification



Reverse engineering visualization

Modality – Image

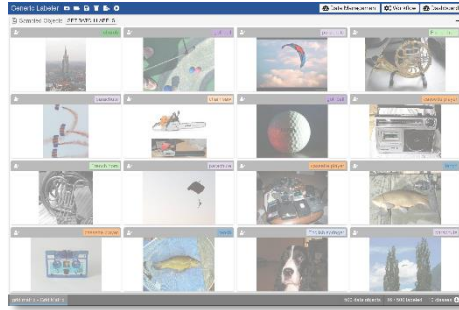


Image classification

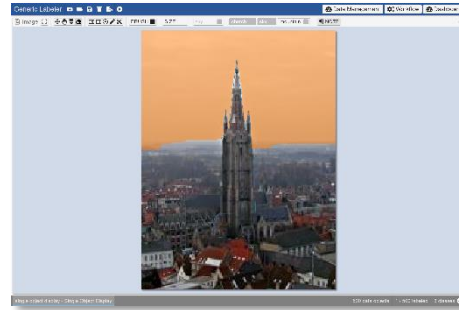
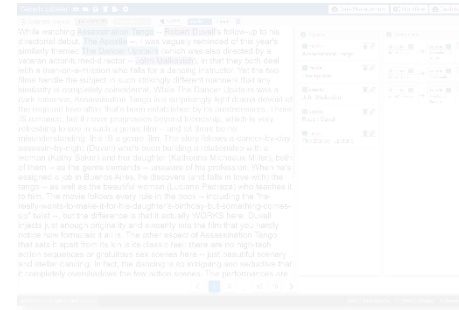
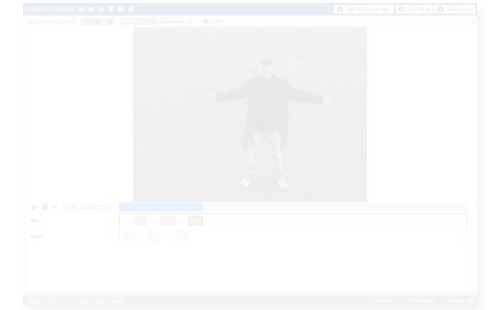


Image segmentation



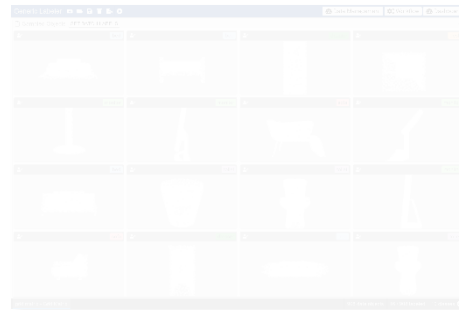
Text named-entity recognition



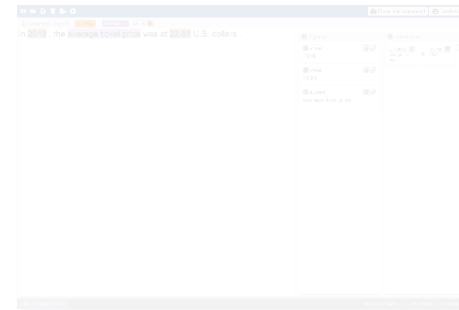
Video temporal segmentation



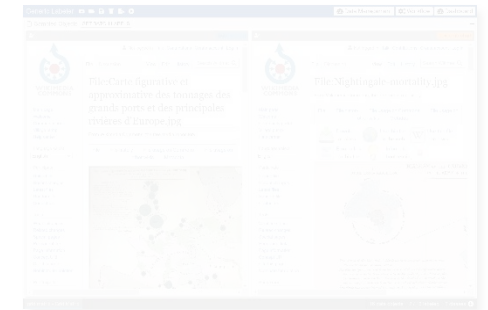
Audio temporal segmentation



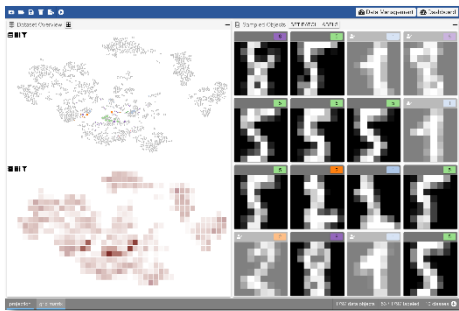
Point cloud classification



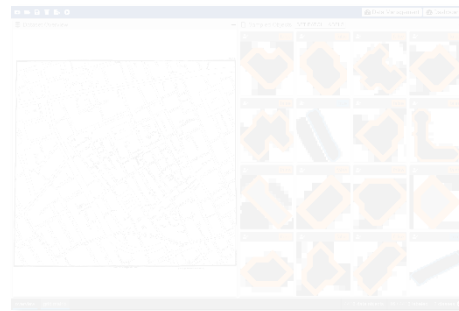
Machine-aided named-entity



Webpage classification



Mixed-initiative classification



Reverse engineering visualization

Modality – Text

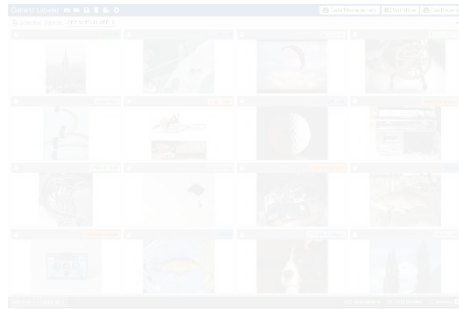


Image classification

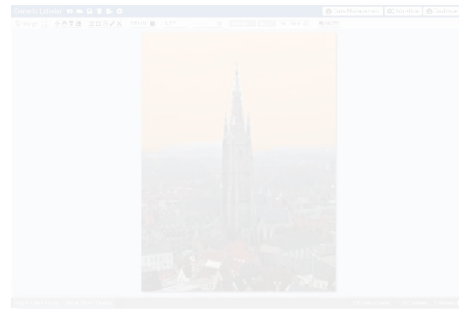
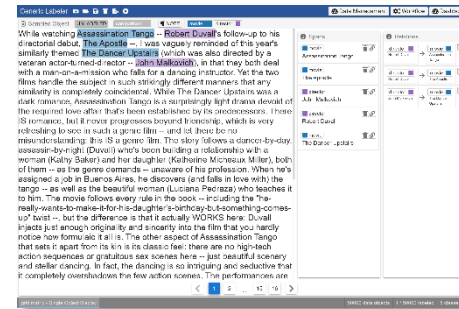
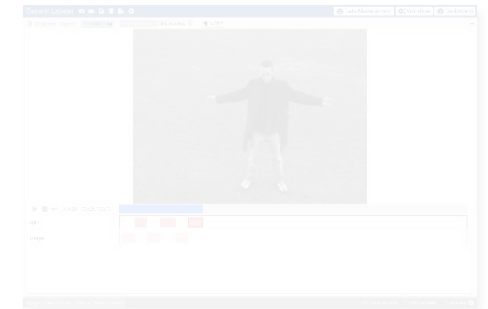


Image segmentation



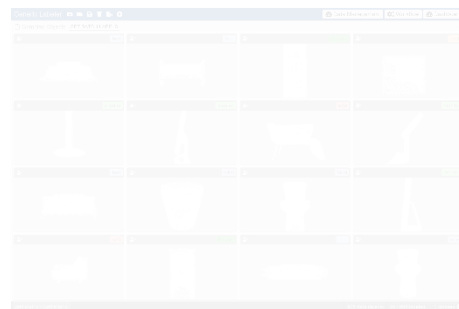
Text named-entity recognition



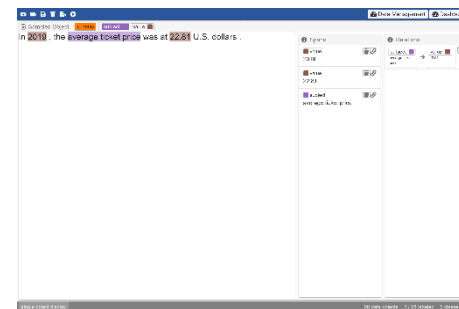
Video temporal segmentation



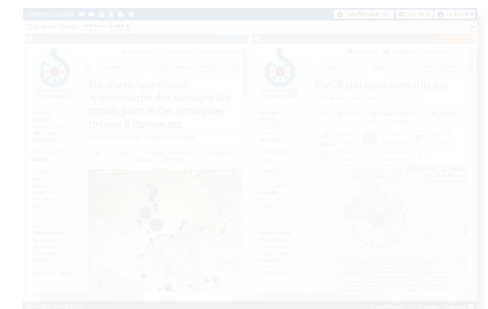
Audio temporal segmentation



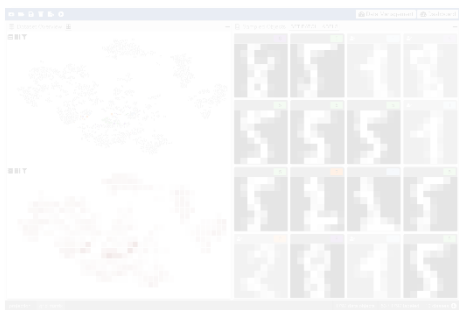
Point cloud classification



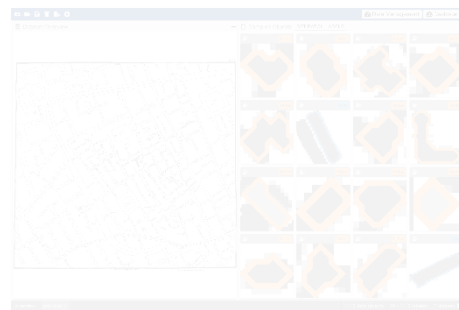
Machine-aided named-entity



Webpage classification



Mixed-initiative classification



Reverse engineering visualization

Modality – Video

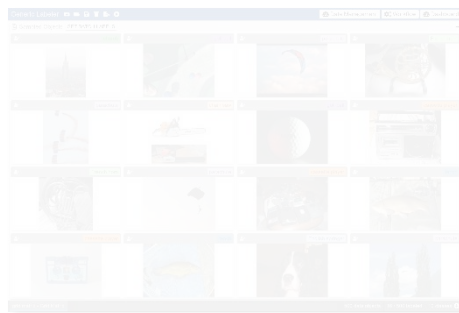


Image classification

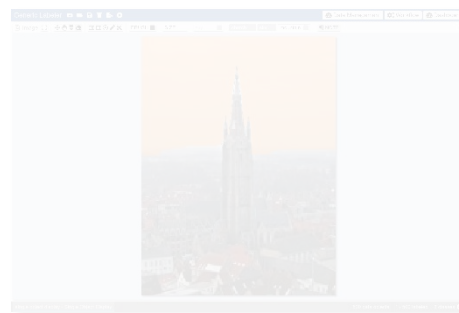
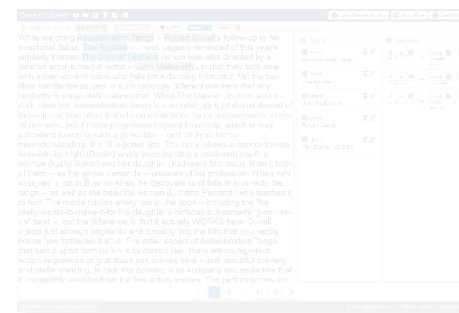
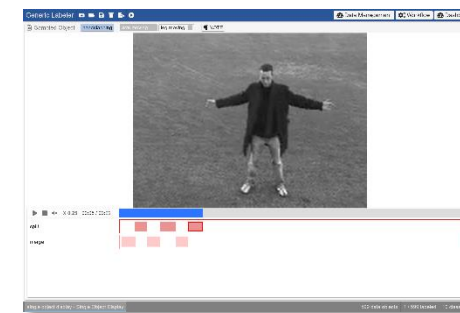


Image segmentation



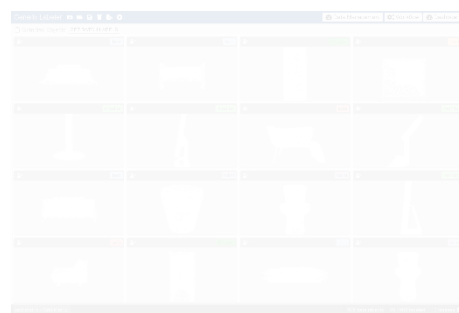
Text named-entity recognition



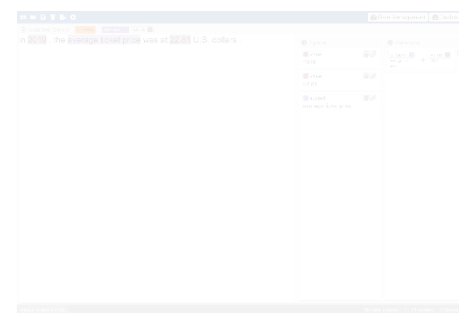
Video temporal segmentation



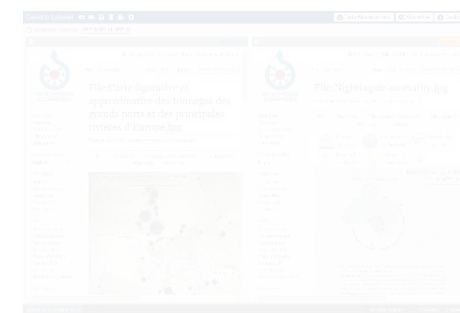
Audio temporal segmentation



Point cloud classification



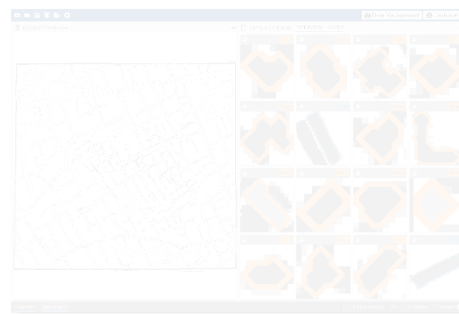
Machine-aided named-entity



Webpage classification



Mixed-initiative classification



Reverse engineering visualization

Modality – Audio

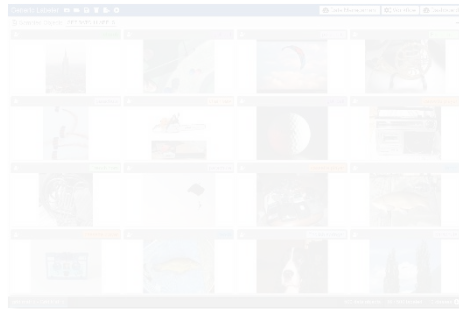


Image classification

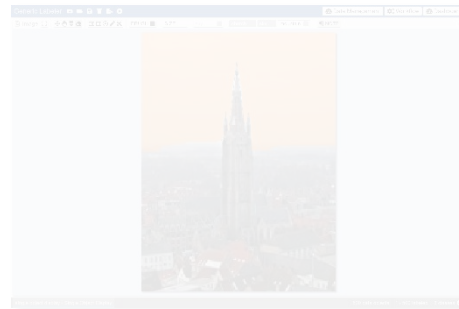
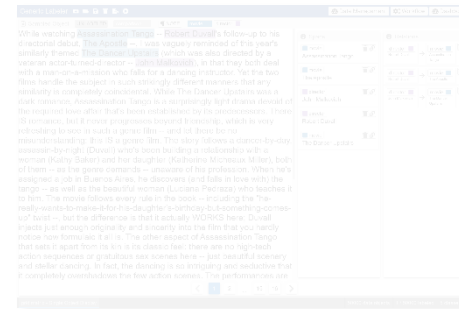
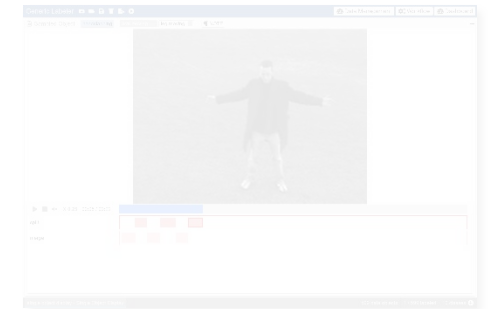


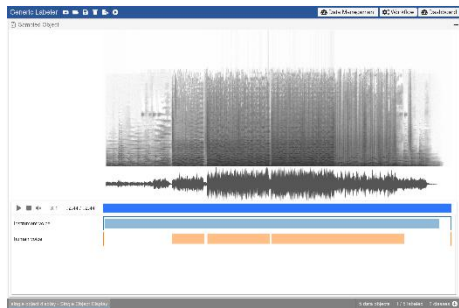
Image segmentation



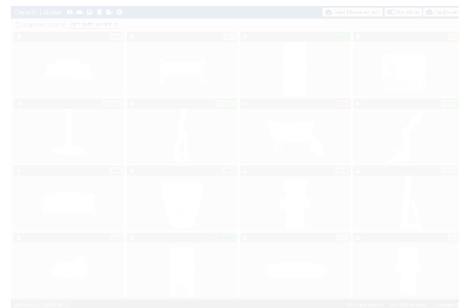
Text named-entity recognition



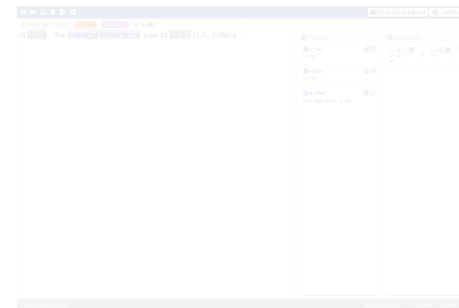
Video temporal segmentation



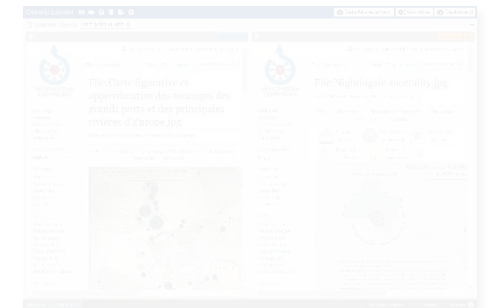
Audio temporal segmentation



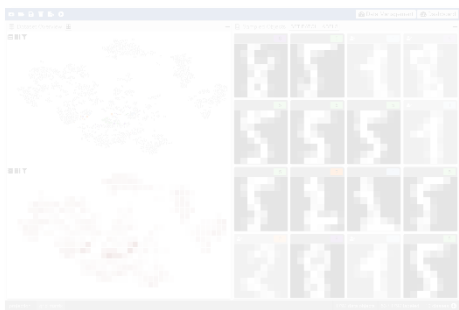
Point cloud classification



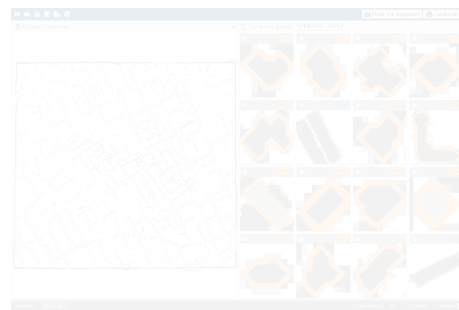
Machine-aided named-entity



Webpage classification



Mixed-initiative classification



Reverse engineering visualization

Modality – Point Cloud

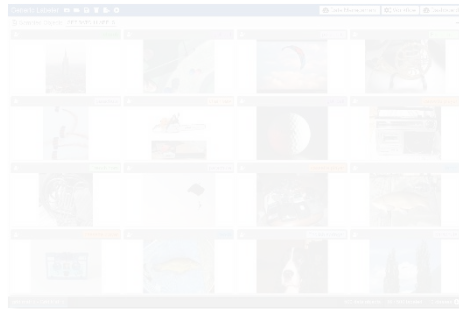


Image classification

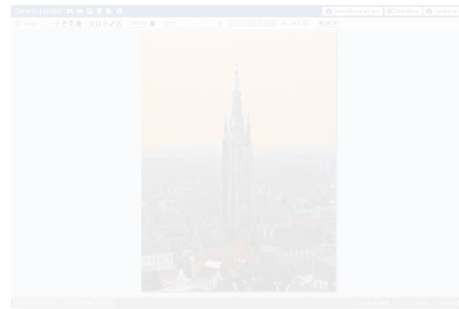
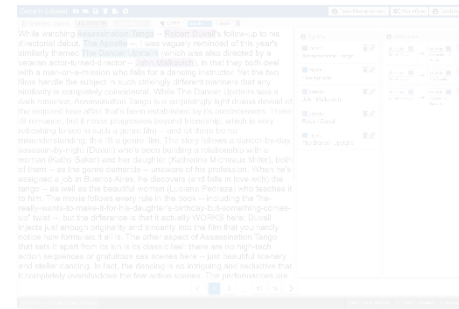
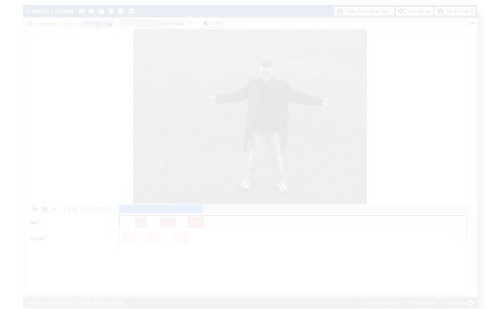


Image segmentation



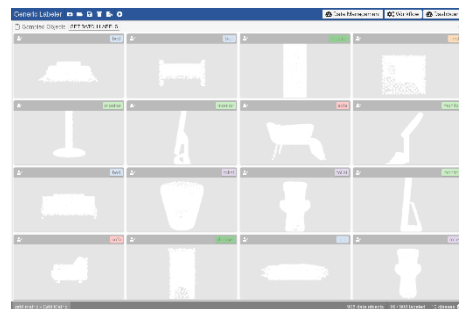
Text named-entity recognition



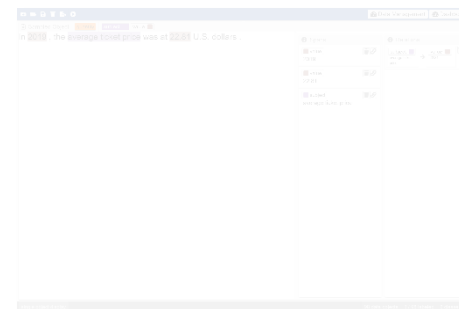
Video temporal segmentation



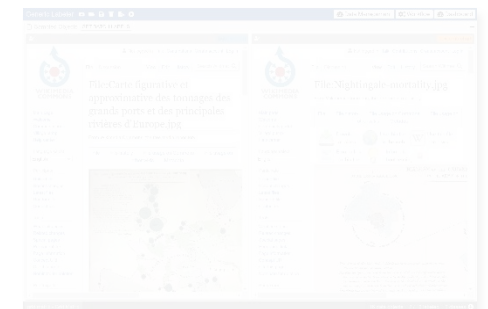
Audio temporal segmentation



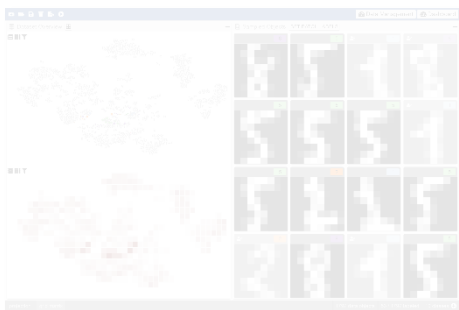
Point cloud classification



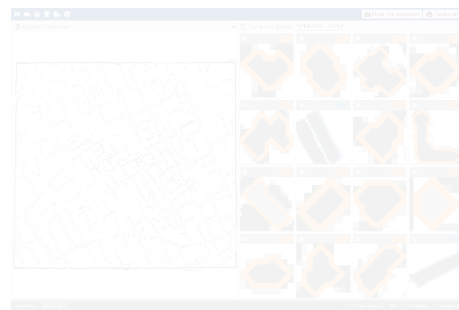
Machine-aided named-entity



Webpage classification



Mixed-initiative classification



Reverse engineering visualization

Modality – Webpage

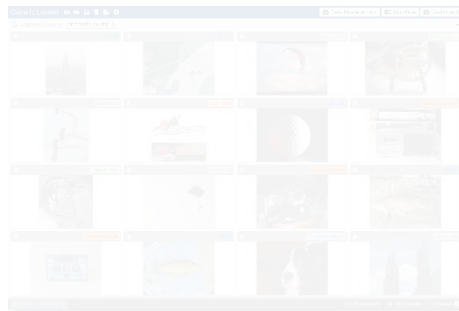


Image classification

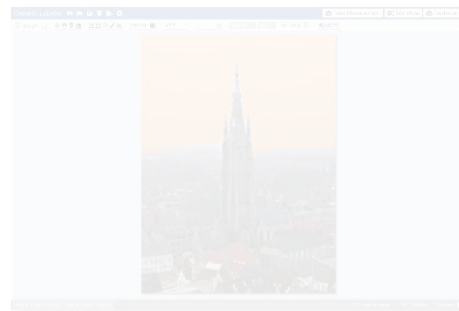
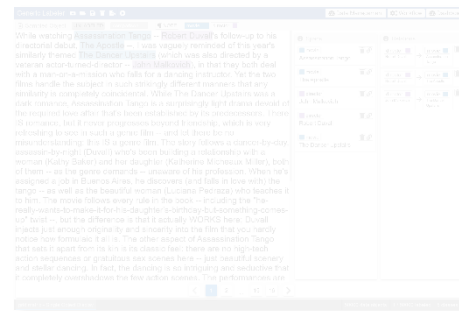
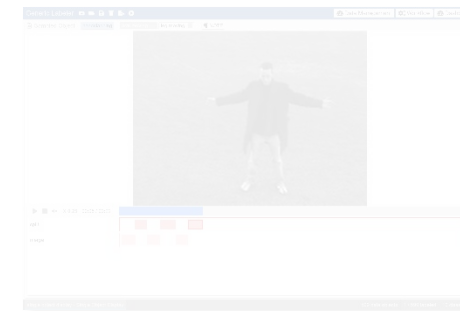


Image segmentation



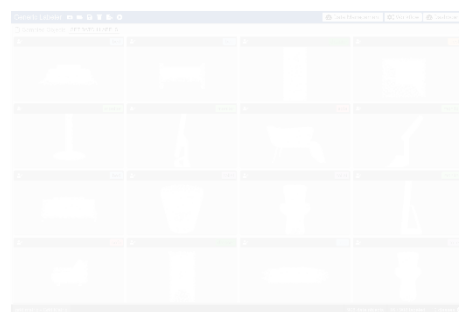
Text named-entity recognition



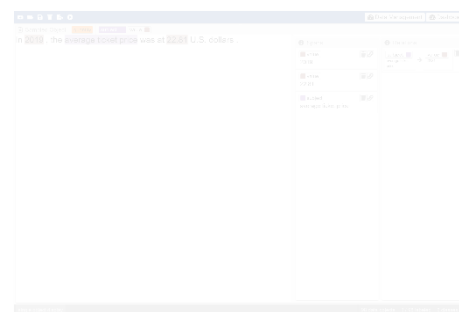
Video temporal segmentation



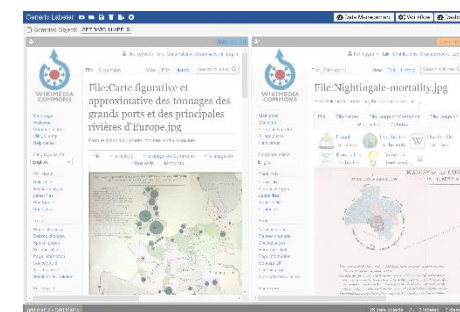
Audio temporal segmentation



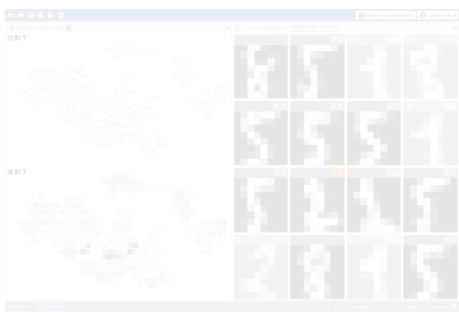
Point cloud classification



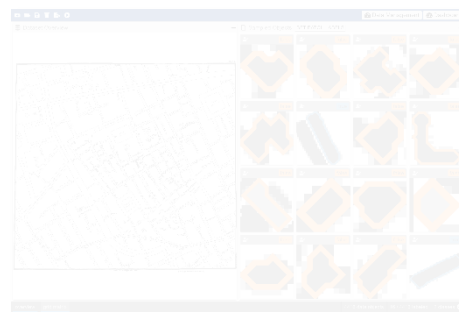
Machine-aided named-entity



Webpage classification



Mixed-initiative classification



Reverse engineering visualization

Task – Classification

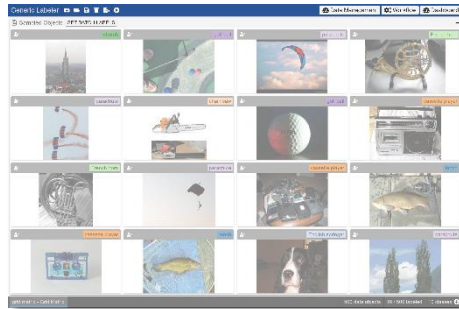


Image classification

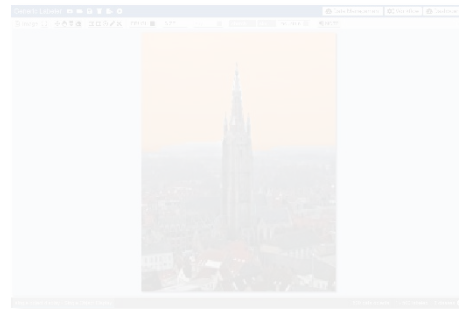
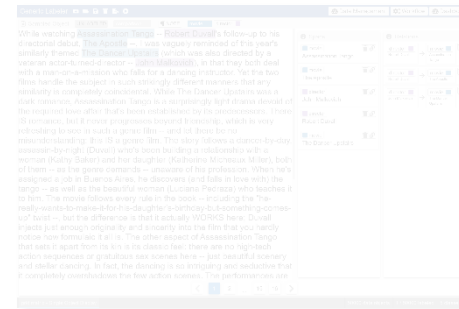
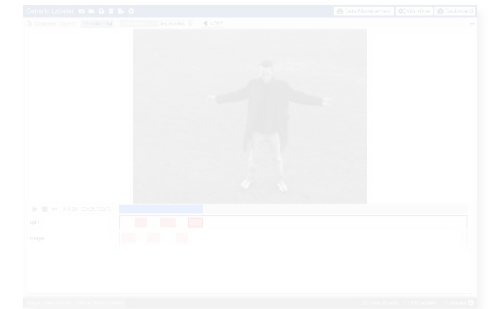


Image segmentation



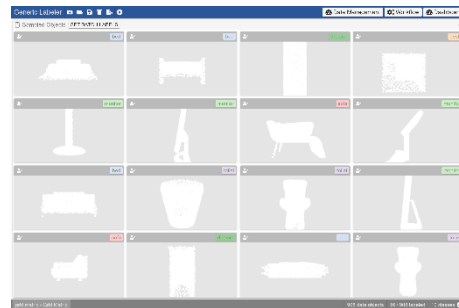
Text named-entity recognition



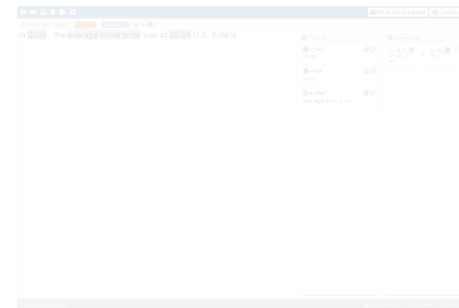
Video temporal segmentation



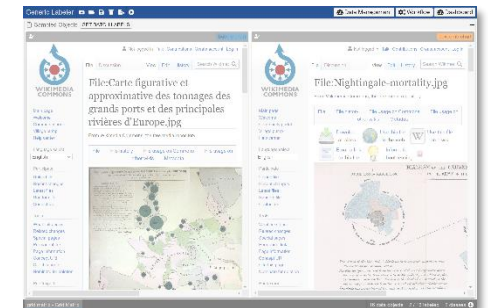
Audio temporal segmentation



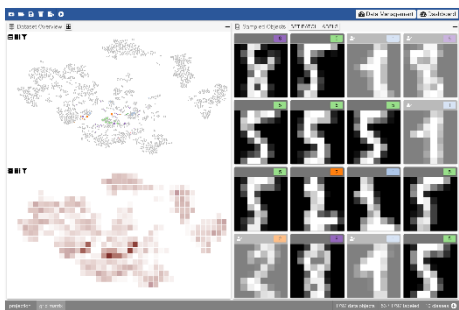
Point cloud classification



Machine-aided named-entity



Webpage classification



Mixed-initiative classification



Reverse engineering visualization

Task – Segmentation

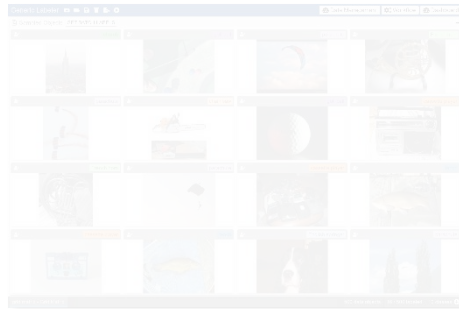


Image classification

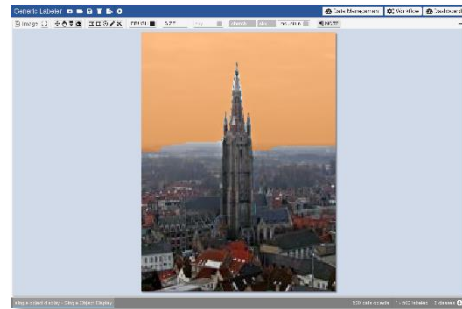
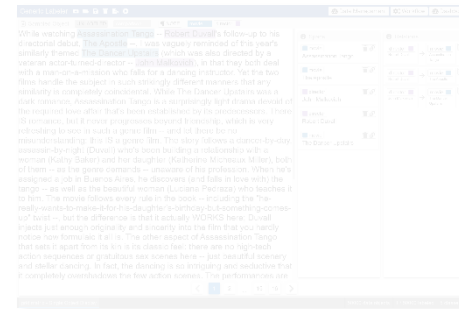
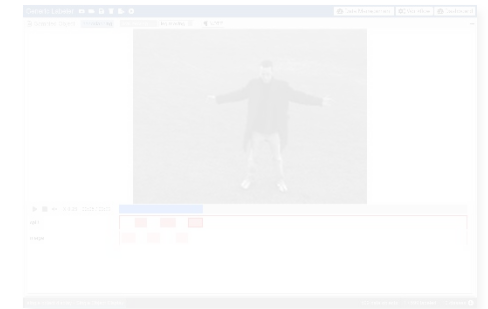


Image segmentation



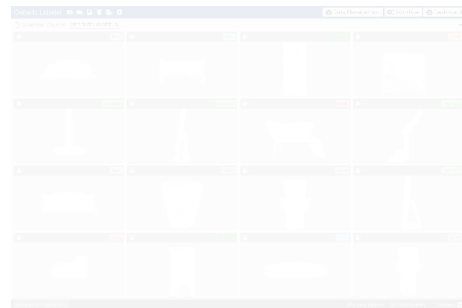
Text named-entity recognition



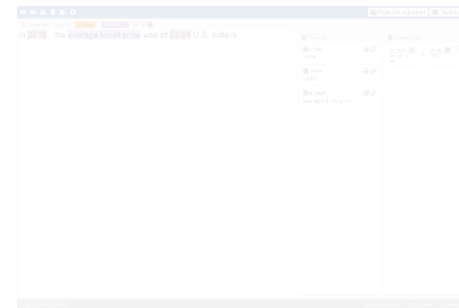
Video temporal segmentation



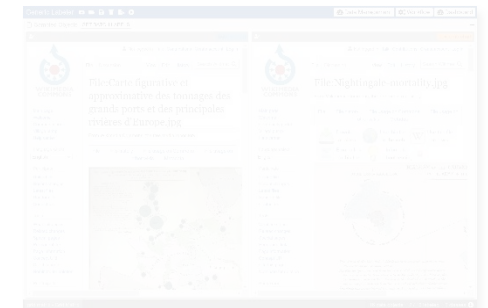
Audio temporal segmentation



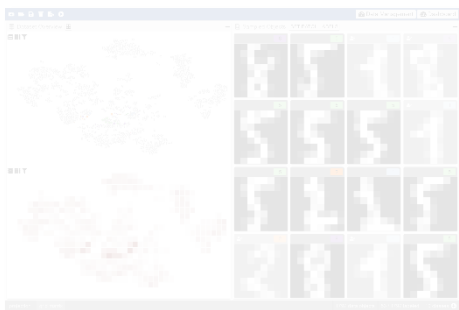
Point cloud classification



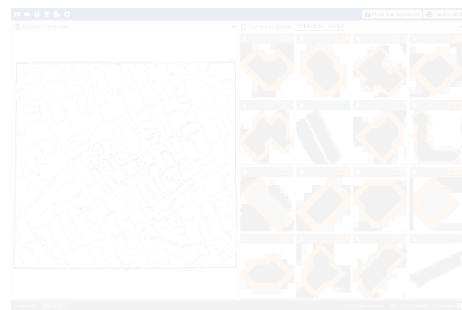
Machine-aided named-entity



Webpage classification



Mixed-initiative classification



Reverse engineering visualization

Task – Temporal/Textual Span Tagging

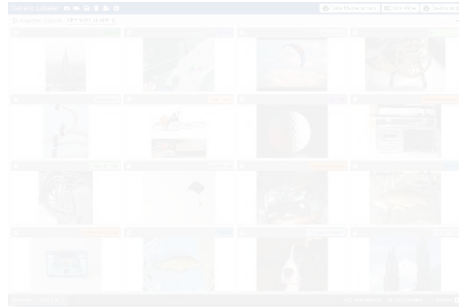


Image classification

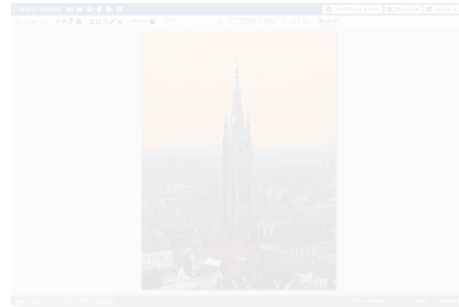
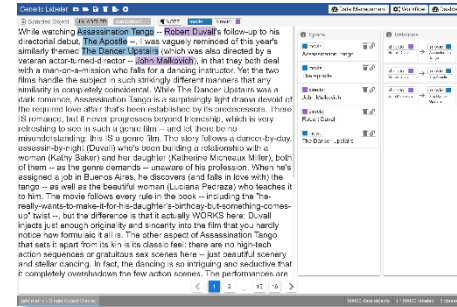


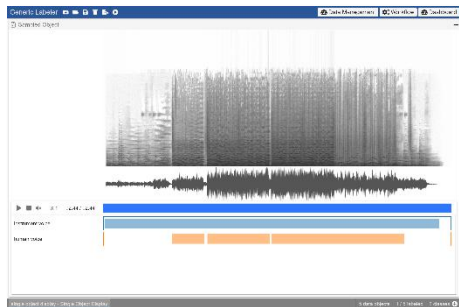
Image segmentation



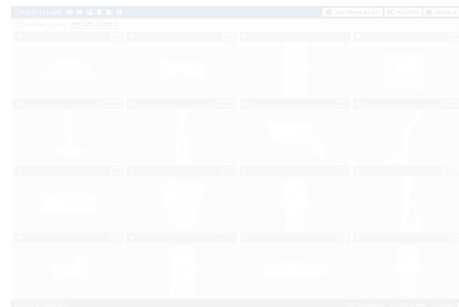
Text named-entity recognition



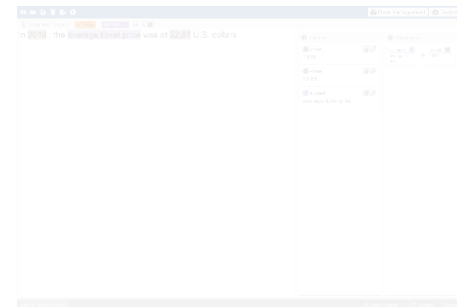
Video temporal segmentation



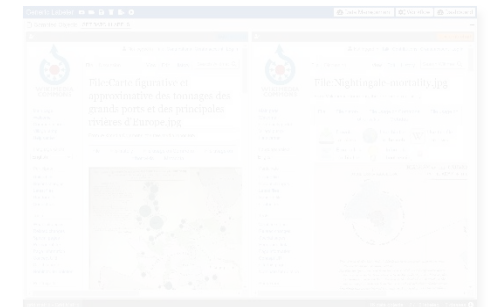
Audio temporal segmentation



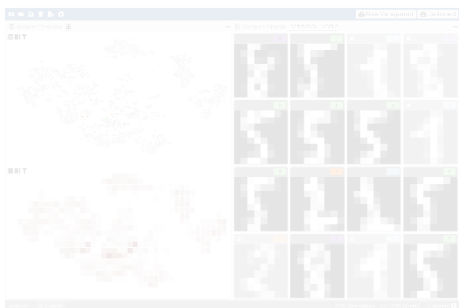
Point cloud classification



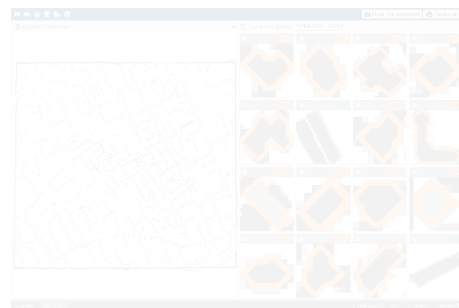
Machine-aided named-entity



Webpage classification



Mixed-initiative classification



Reverse engineering visualization

Adding Machine Assistance

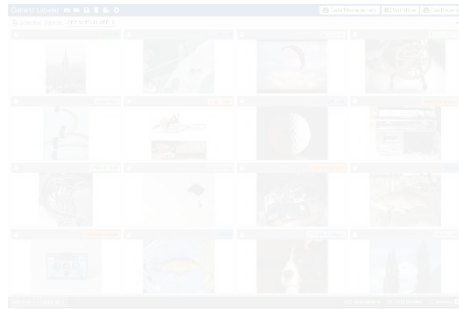


Image classification

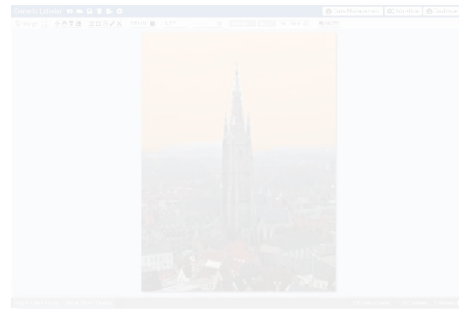
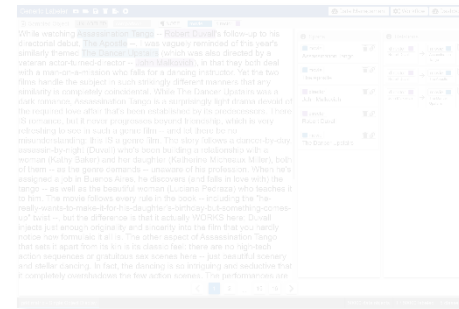
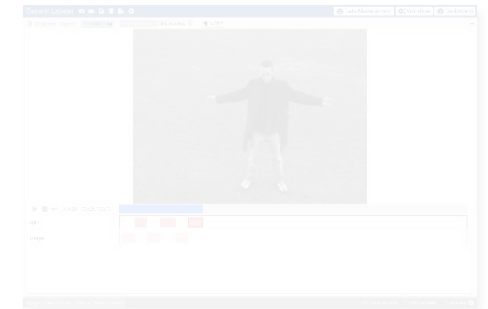


Image segmentation



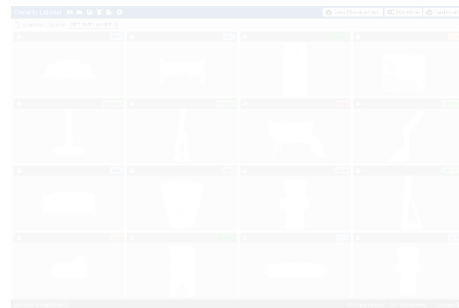
Text named-entity recognition



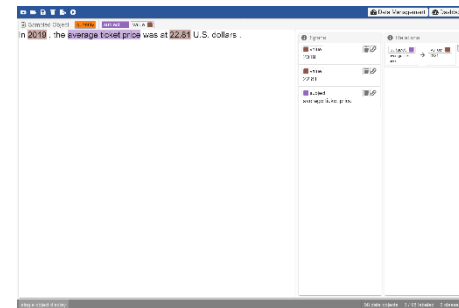
Video temporal segmentation



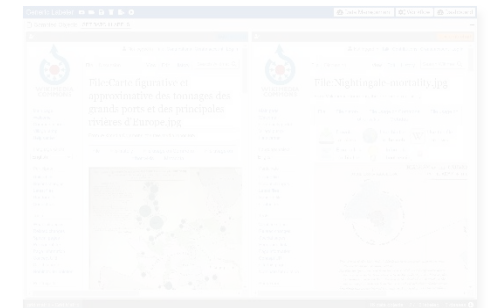
Audio temporal segmentation



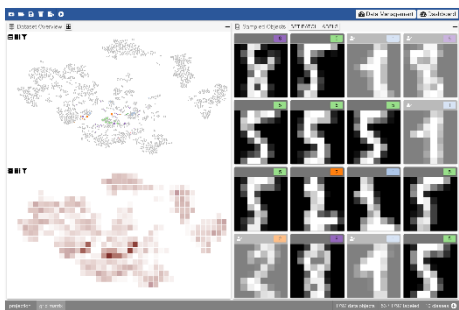
Point cloud classification



Machine-aided named-entity



Webpage classification



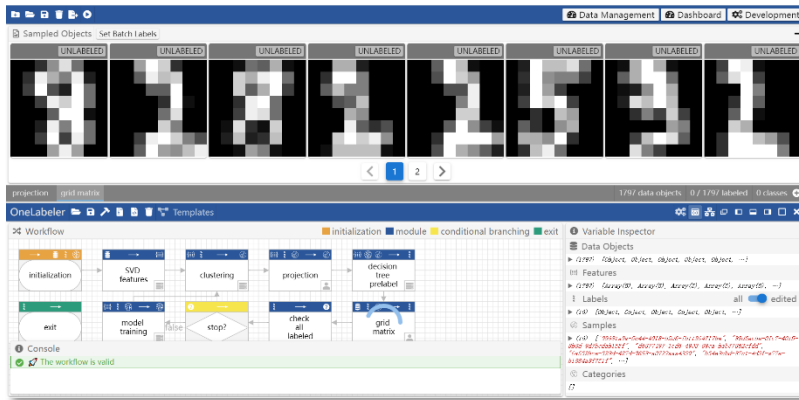
Mixed-initiative classification



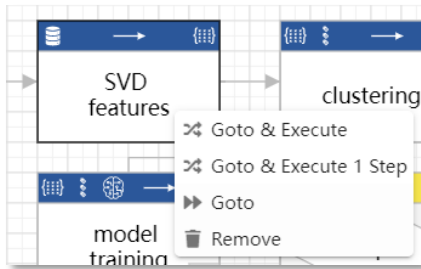
Reverse engineering visualization

System Features

Labeling Tool Preview

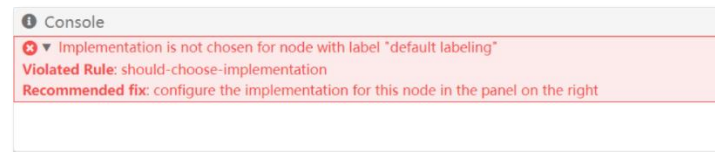


Variable Inspector

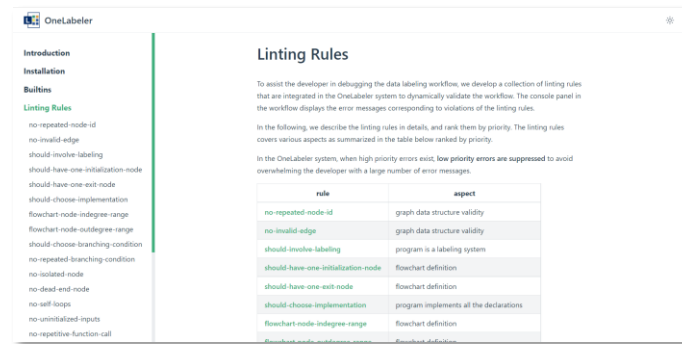


Single Stepping

Static Checking

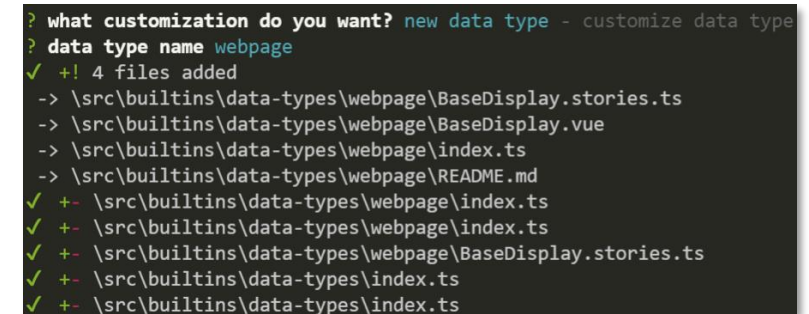
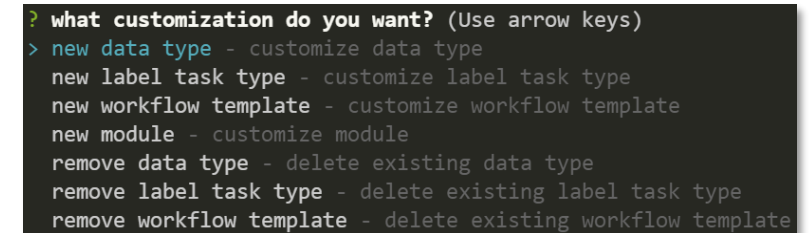


Console panel



Error Codes

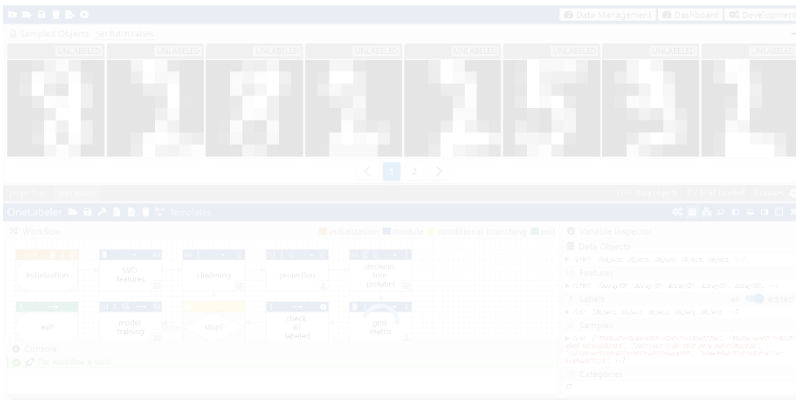
Customized Extension



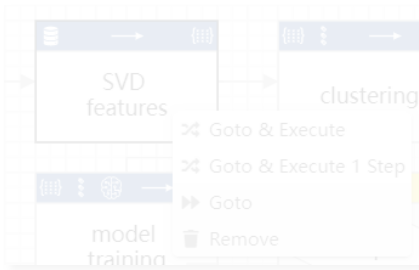
Command-line Interface

System Features

Labeling Tool Preview

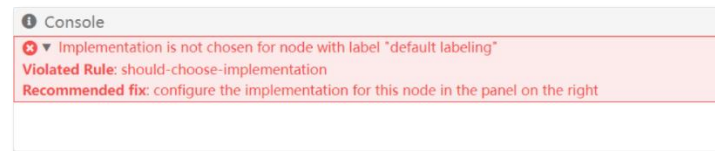


Variable Inspector

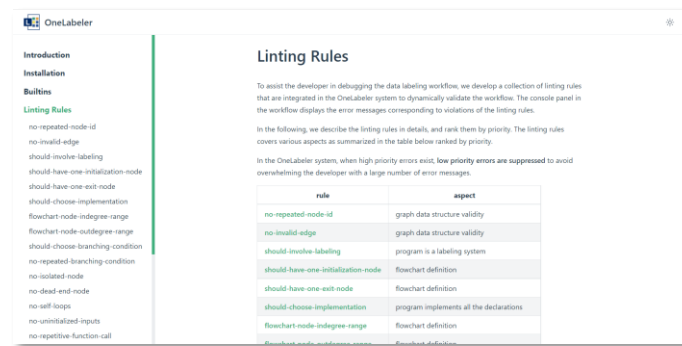


Single Stepping

Static Checking

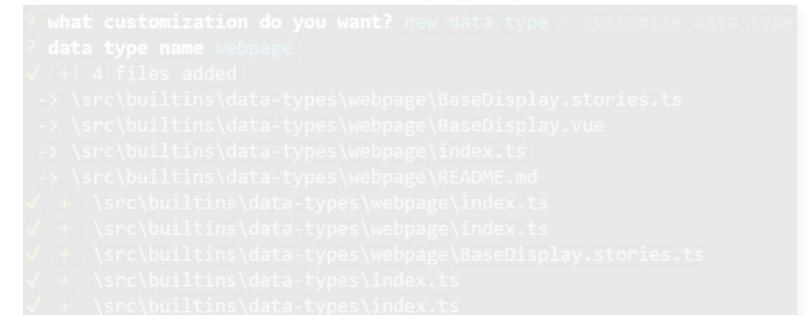
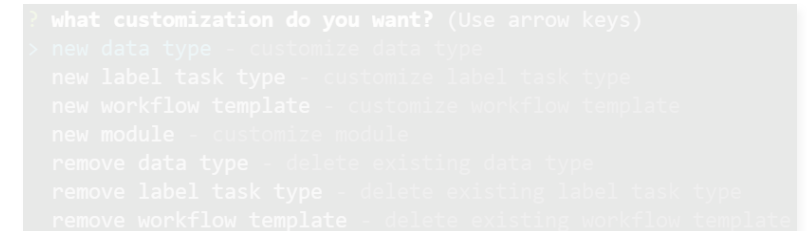


Console panel



Error Codes

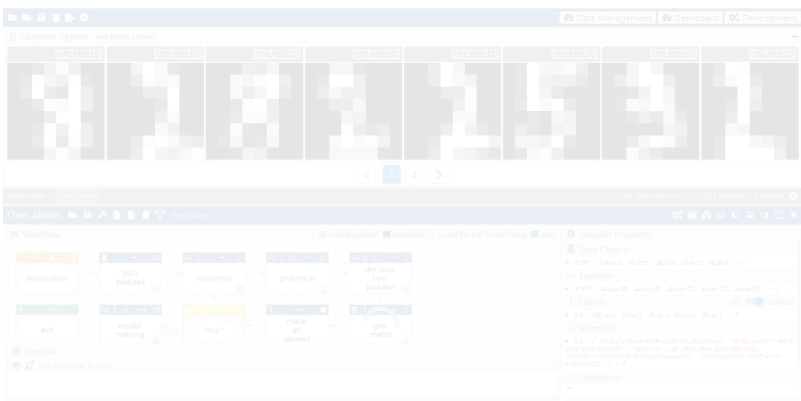
Customized Extension



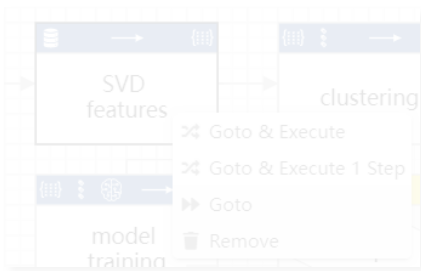
Command-line Interface

System Features

Labeling Tool Preview



Variable Inspector

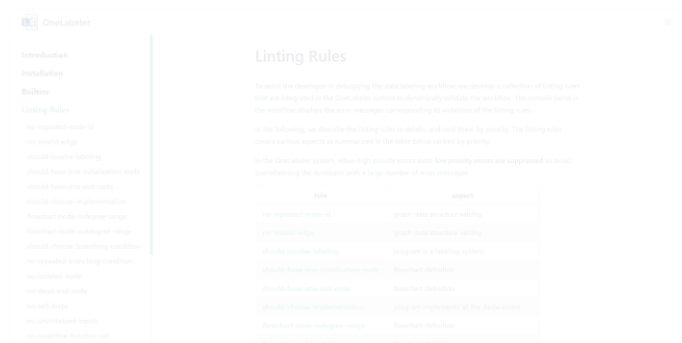


Single Stepping

Static Checking



Console panel



Error Codes

Customized Extension

```
? what customization do you want? (Use arrow keys)
> new data type - customize data type
new label task type - customize label task type
new workflow template - customize workflow template
new module - customize module
remove data type - delete existing data type
remove label task type - delete existing label task type
remove workflow template - delete existing workflow template
```



```
? what customization do you want? new data type - customize data type
? data type name webpage
✓ +! 4 files added
-> \src\builtins\data-types\webpage\BaseDisplay.stories.ts
-> \src\builtins\data-types\webpage\BaseDisplay.vue
-> \src\builtins\data-types\webpage\index.ts
-> \src\builtins\data-types\webpage\README.md
✓ +- \src\builtins\data-types\webpage\index.ts
✓ +- \src\builtins\data-types\webpage\index.ts
✓ +- \src\builtins\data-types\webpage\BaseDisplay.stories.ts
✓ +- \src\builtins\data-types\index.ts
✓ +- \src\builtins\data-types\index.ts
```

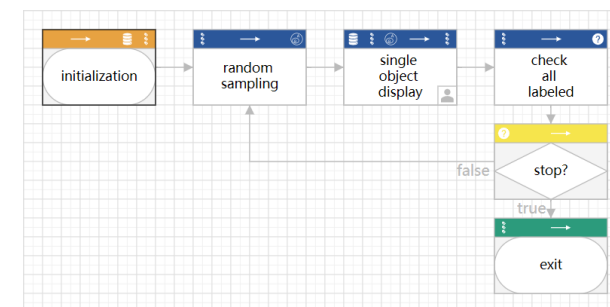
Command-line Interface

User Study

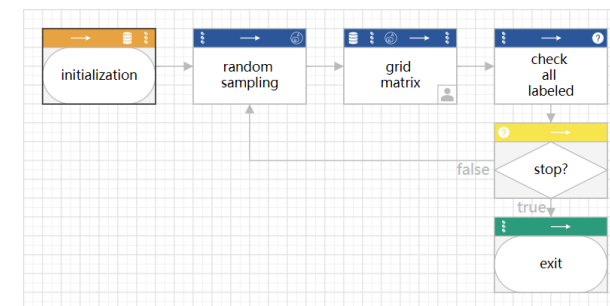
- Participants: 8 developers
- Tasks: 4 reproduction tasks
- Training: 30 minutes reading documentation

```
1 <template>
2   <iframe :src="dataObject.content" />
3 </template>
4
5 <script lang="ts">
6   import type { PropType } from 'vue';
7   import type { IDataObject } from '@/commons/types';
8
9   export default {
10    name: 'BaseDisplay',
11    props: {
12      /** The data object to be rendered. */
13      dataObject: {
14        type: Object as PropType<IDataObject>,
15        required: true,
16      },
17    },
18  };
19 </script>
```

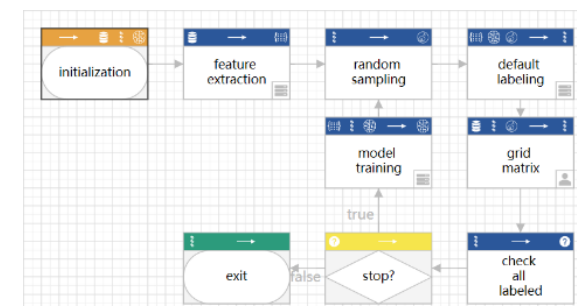
Task 4's reference customization code



Task 1's reference workflow



Task 2's reference workflow



Task 3's reference workflow

User Study

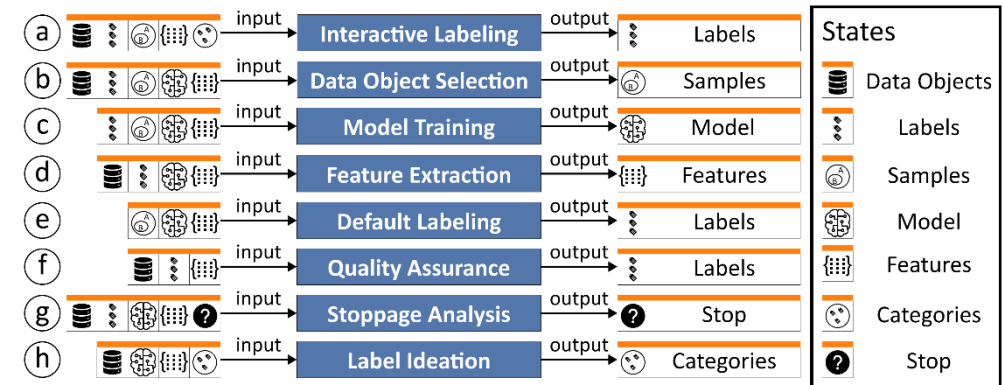
- Result: OneLabeler shows potential to accelerate development
 - Labeling tool development was reported to typically take **weeks**
 - Each participant **passed 3 – 4 tasks** with on average around **50 minutes**

User Study

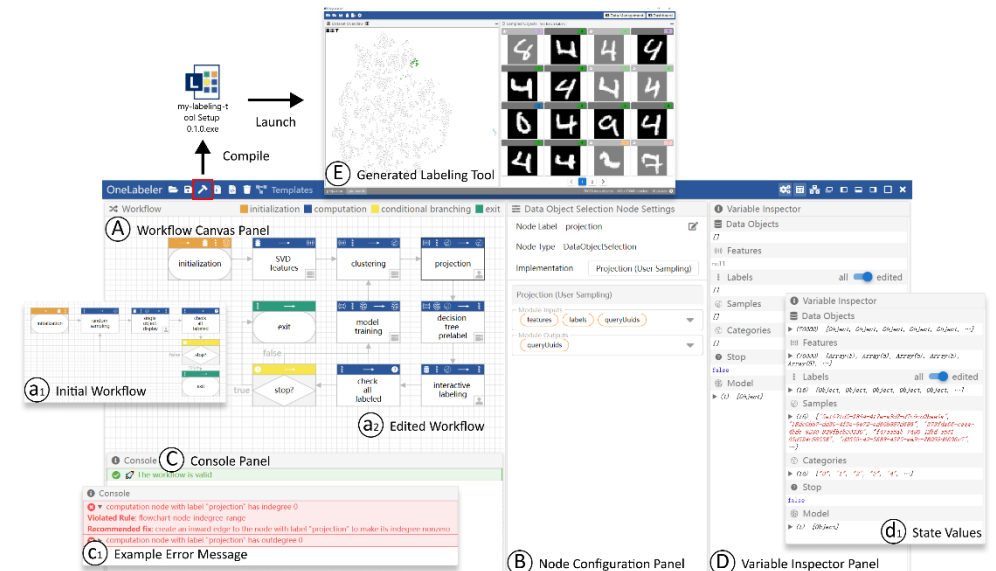
- Result: OneLabeler shows potential to accelerate development
 - Labeling tool development was reported to typically take **weeks**
 - Each participant **passed 3 – 4 tasks** with on average around **50 minutes**
- Failure Analysis: Customization demands expertise in web development

Summary

- Common modules in labeling tools
 - Identified by coding flowcharts



- Visual programming toolkit
 - Build in implementations
 - Development support features





OneLabeler



Code & Documentation

<https://github.com/microsoft/OneLabeler>



Yu Zhang^{1,2}



Yun Wang²



Haidong Zhang²



Bin Zhu²



Siming Chen³



Dongmei Zhang²

¹University of Oxford ²Microsoft Research Asia ³Fudan University

