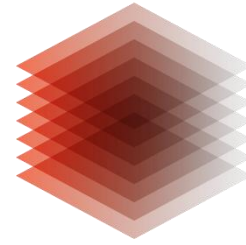

LEIBNIZ INFORMATION CENTRE
FOR SCIENCE AND TECHNOLOGY
UNIVERSITY LIBRARY

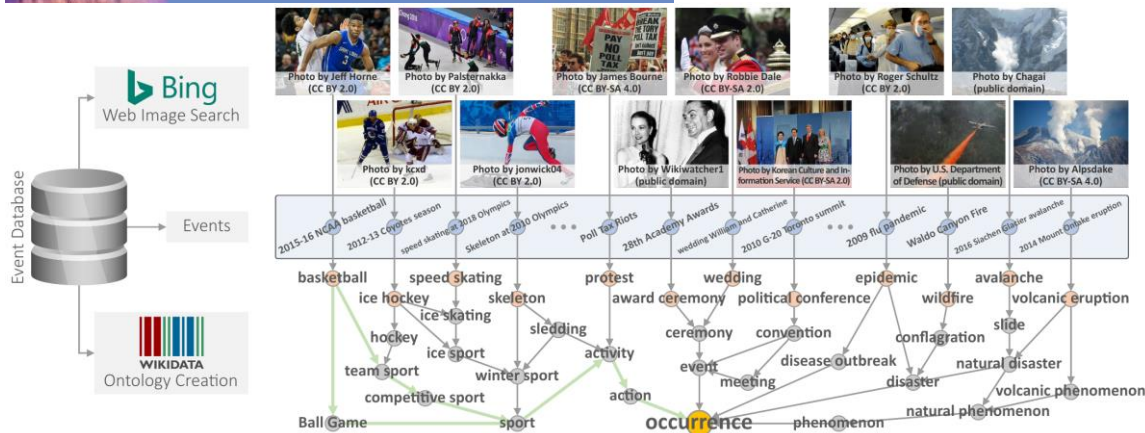
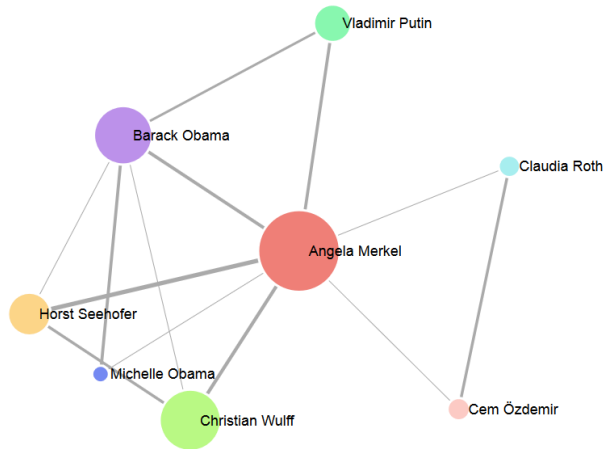


TIB

Multimodal Patent Retrieval with Vision-Language Models

Dr.-Ing. Eric Müller-Budack
TIB – Leibniz Information Centre for Science and Technology
Research Group Visual Analytics (Lead: Prof. Dr. Ralph Ewerth)

My Background

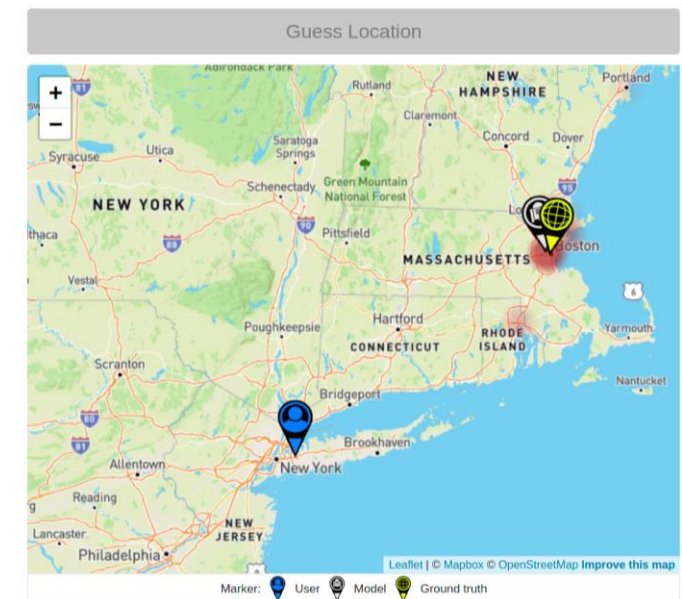
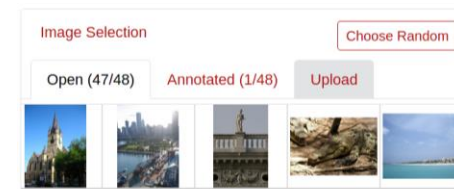


sigir21



[...] have been appalled yesterday to witness the behavior of Barack Obama PER 0.89, when, at a news conference with David Cameron PER 0.87 in Brussels LOC 0.63 he was asked whether Scotland LOC 0.59 should remain in the UK LOC 0.76 or leave. [...] (504 words)

[...] Steven Gerrard PER 0.85 believes he proved he 'is not finished yet' after his tenth Merseyside derby EVENT 0.97 goal in the 1-1 draw with [...] Brendan Rodgers PER 0.65 the Liverpool LOC 0.97 manager [...] (552 words)

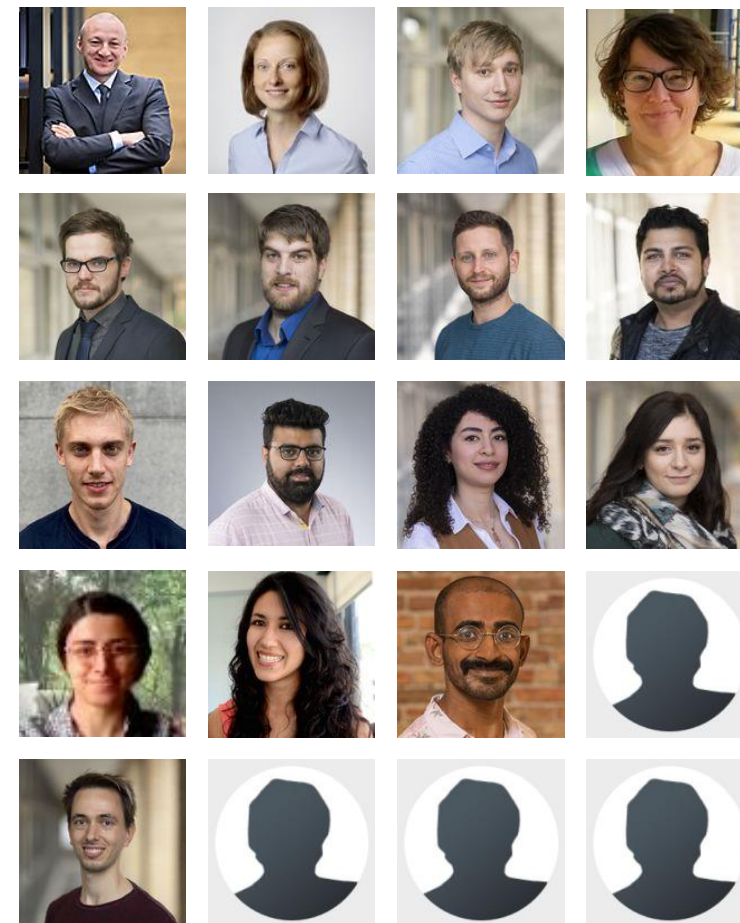


Research Group Visual Analytics (TIB)

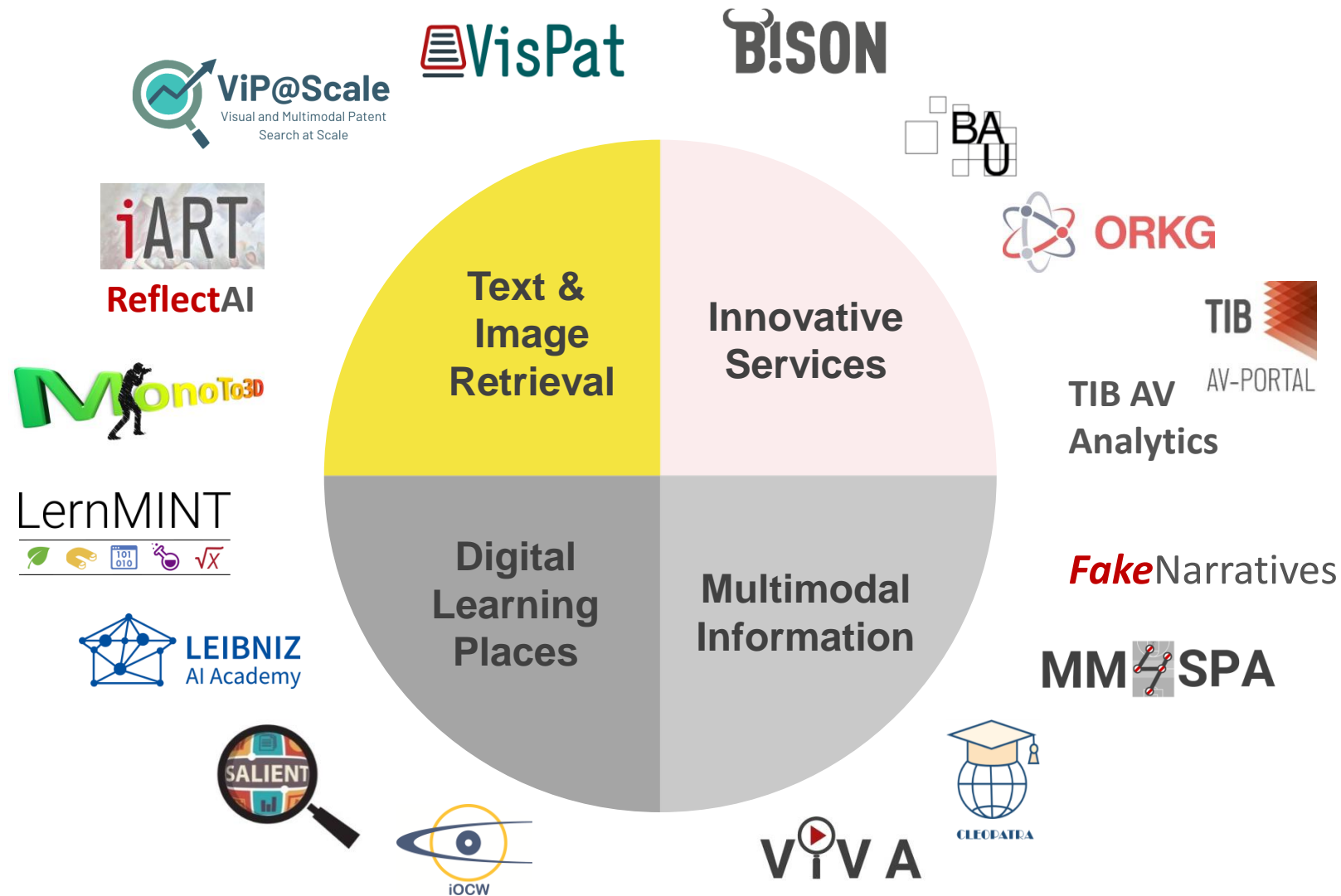
- Since November 2015 at TIB
- Lead: Prof. Dr. Ralph Ewerth
- 20 members (thereof 3 postdocs, 13 PhD students)

Research interests:

- Multimodal & cross-modal information retrieval
- Deep learning for the analysis and annotation of media data
- Interactive search, exploration, and visualization of media data
- Information extraction for the Open Research Knowledge Graph
- Digital library as a virtual place of learning



Research Projects (past 5 years)



Tagesschau CC BY-SA 4.0



00:12:15.4271.00

00:00:0000:01:54.3500:03:49.0500:05:43.6300:07:38.1800:09:32.7200:11:27.27




a

d

SHOTSANNOTATIONS




Shot 1

Begin: 00:00:00.000
End: 00:00:00.160
Duration: 00:00:00.160






Shot 2

Begin: 00:00:00.200
End: 00:00:29.480
Duration: 00:00:29.280



Shot 3

Begin: 00:00:29.520
End: 00:00:36.800
Duration: 00:00:07.280



b1

SHOTSANNOTATIONS

Full Shot

Neutral

/s/stadium/soccer

sports fields, parks, leisure spaces

outdoor, man-made

b2

OPTIONS

Newscaster

Face Identification Lahm

Celebrating [X-CLIP]

FIFA WC Trophy [CLIP]

Celebrating & Lahm & Trophy

— Face Emotions

Happy

Disoust

Newscaster

Newscaster

Newscaster

Ne

New

ad

Neutr

Neutral

Sad

Angr

Fe

Fea

Angr

Sa

Angr

Sa

Neutral

Hi

Happy

Neutr

Ne

Ne

Sa

Ha


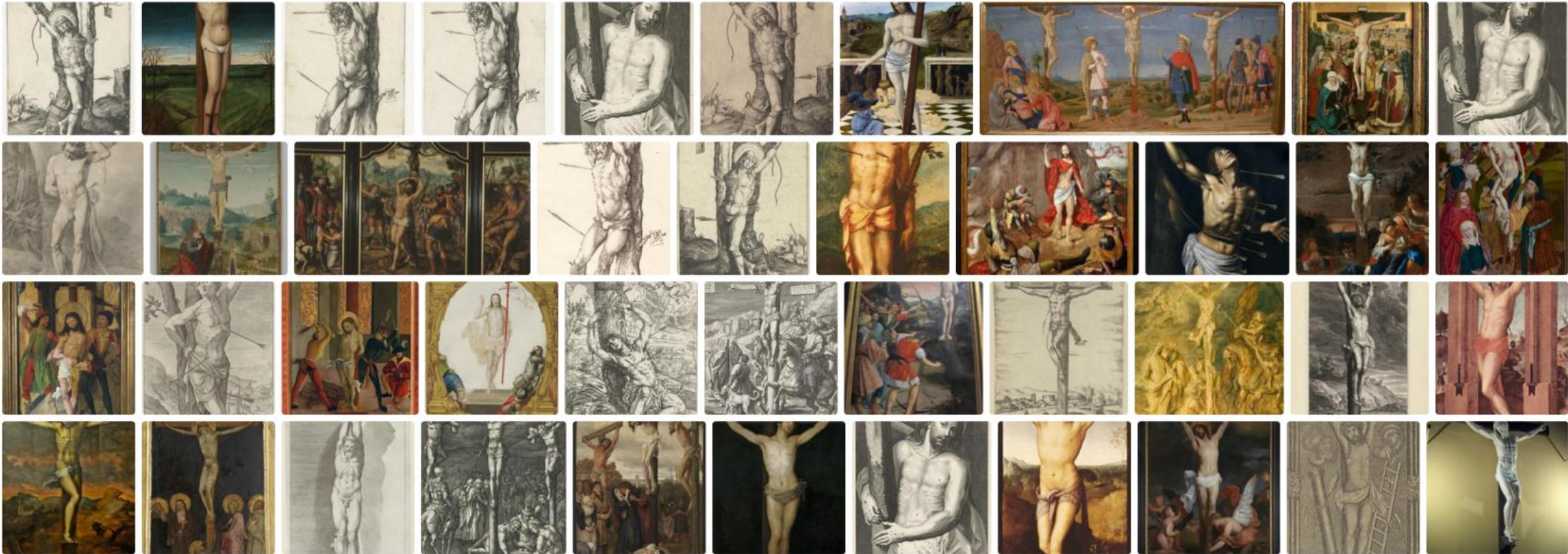
Net

Ne

Neutral

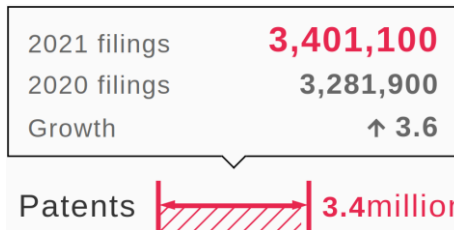
Neutral

c

 Global Weights Result View Cluster Display 

Motivation

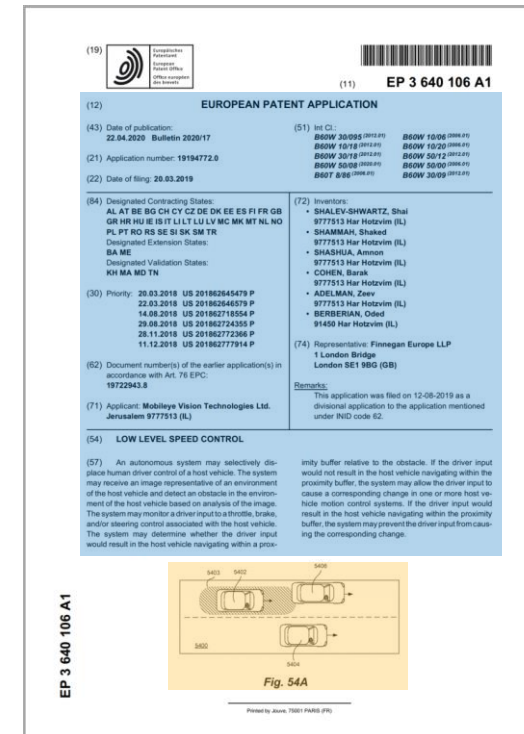
- Number of patent applications rapidly growing



Source (WIPO¹): <https://www.wipo.int/en/ipfactsandfigures/patents>

- **Patent search** is an important task:
 - Assessment of the novelty of an invention
 - Retrieval of related work
 - Identification of plagiarism

- Current **search systems** and approaches **mainly rely on text or metadata** [Krestel et al. 2021]
- But patents are multimodal and figures contain important information!**

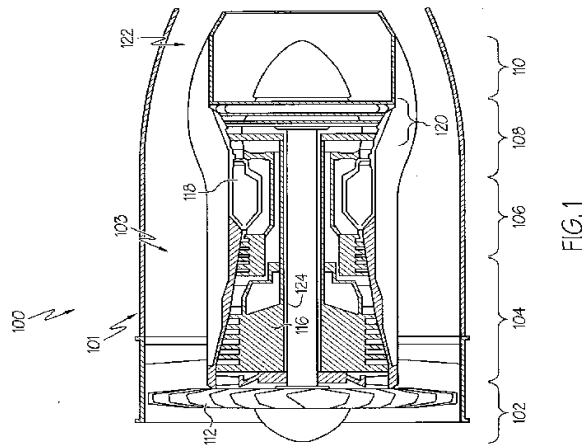


¹WIPO: World Intellectual Property Organization

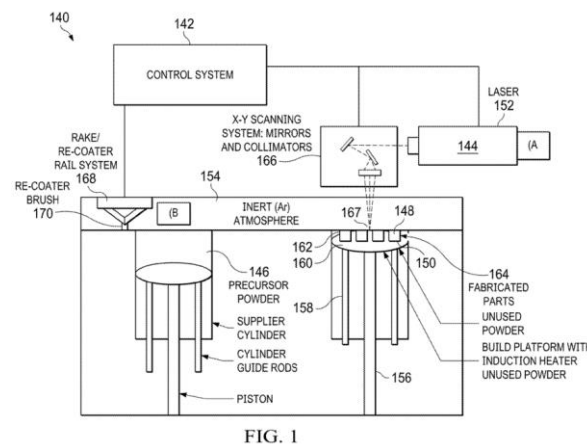
Motivation

Image-based and multimodal patent retrieval systems are required!

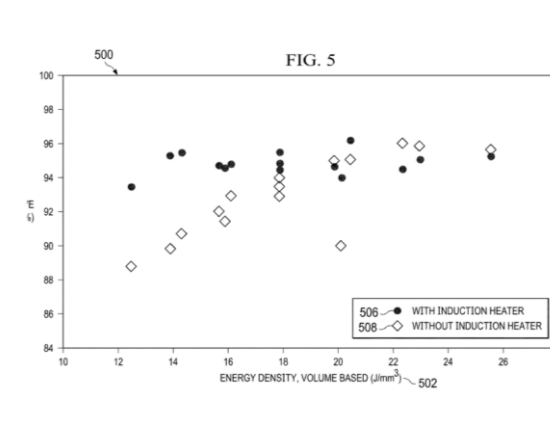
Technical specifications
(e.g., technical drawings)



Experimental settings
(e.g., structured diagrams)



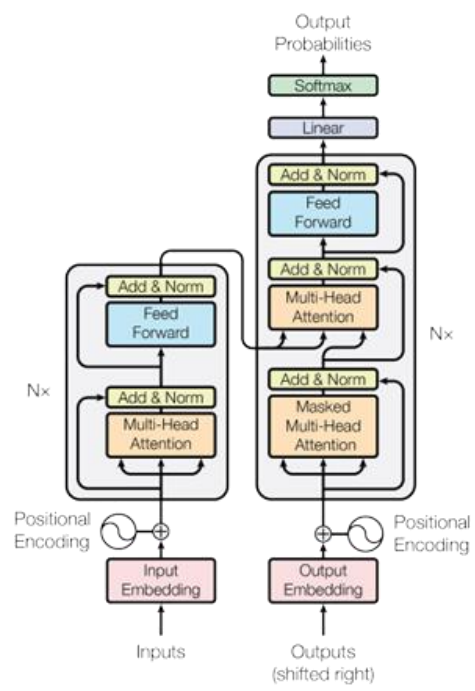
Experimental results
(e.g., charts)



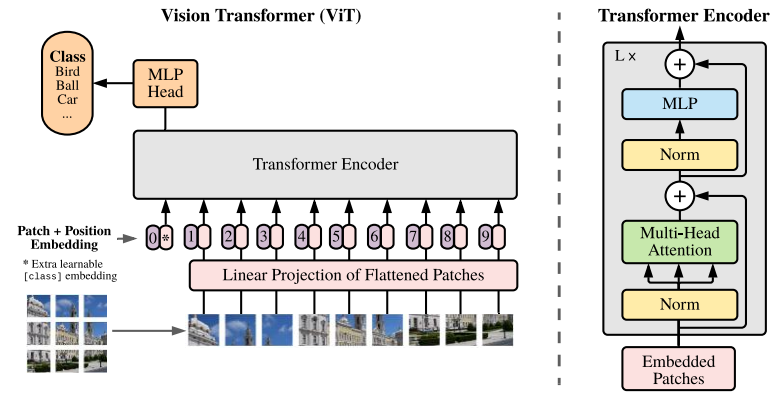
- Current **search systems** and approaches **mainly rely on text or metadata** [Krestel et al. 2021]
- **But patents are multimodal and figures contain important information!**

New Opportunities

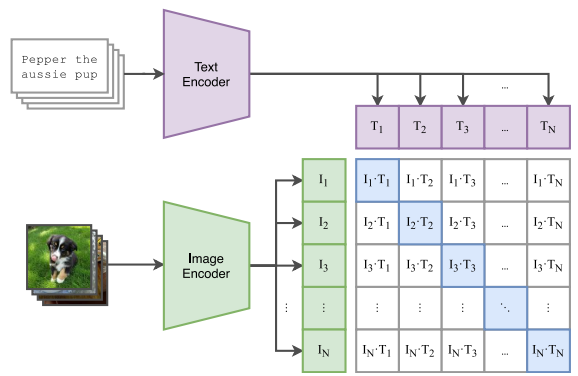
Vision Transformer [Dosovitskiy et al. 2021]



Transformer
(Attention Mechanism)
[Vaswani et al. 2017]



(1) Contrastive pre-training



Contrastive Language-Image Pretraining
[Radford et al. 2021]

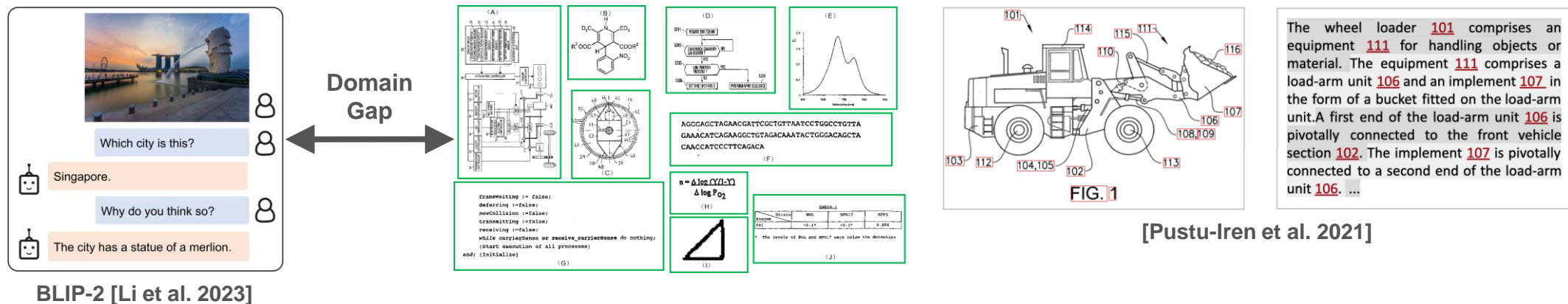
Large Language Models (LLMs)



Vision-Language Models (VLMs)

Challenges

- Multimodal patent retrieval is not a trivial task
- Machine learning approaches, particularly from Computer Vision, are **not optimized for patents**



- **Finetuning necessary** to adopt models to the patent domain
- **Multimodal training** (and test) **data is sparse** and descriptive image captions rarely provided
 - Time-consuming **manual annotation** required
 - Automatic approaches to **map semantically related text to images**

ExpResViP

Exploitation of research results through visual patent retrieval

VisPat Home

Search Results (5539 Patents)

Query Image

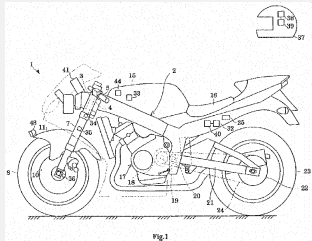
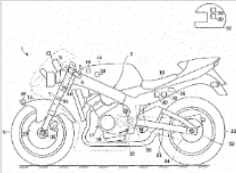
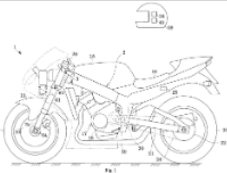
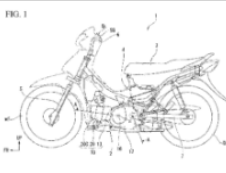
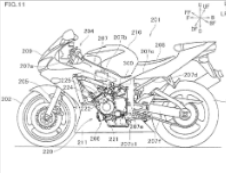
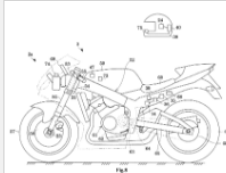


Fig. 1

<input type="checkbox"/> EP17181433A1 1 100.00%	<input type="checkbox"/> EP14850731A1 2 91.86%	<input type="checkbox"/> EP13877372A1 3 88.33%	<input type="checkbox"/> EP17747470A1 4 88.19%	<input type="checkbox"/> EP13755559A1 5 87.32%
 Fig. 1 RIDER CHARACTERISTICS DETERMINING METHOD	 Fig. 1 DRIVING SKILL EVALUATION METHOD, DRIVING SKILL EVALUATION PROGRAM, DRIVING SKILL EVALUATION DEVICE, AND VEHICLE EQUIPPED THEREWITH	 FIG. 1 GENERATOR MOTOR UNIT, POWER OUTPUT ENGINE, AND VEHICLE	 FIG. 11 LEANING VEHICLE	 Fig. 1 HOST COMPUTER, OPERATION- SKILL DETERMINATION SYSTEM, OPERATION-SKILL DETERMINATION METHOD, AND OPERATION-SKILL DETERMINATION PROGRAM

VisPat



Prototype

Project Overview

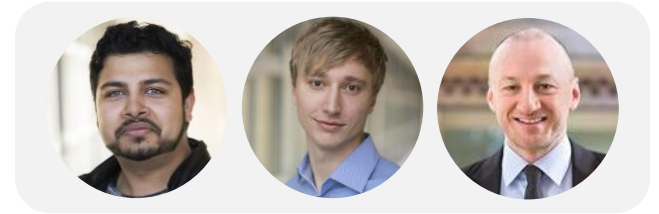


ExpResViP: Exploitation of research results through visual patent retrieval



Bundesministerium
für Bildung
und Forschung

- **Funding:** Federal Ministry of Education and Research (BMBF)
- **Duration:** July 2020 – August 2023
- **Homepage:** <https://projects.tib.eu/expresvip/>
- **Team:** Junaid Ghauri, Dr.-Ing. Eric Müller-Budack, Prof. Dr. Ralph Ewerth
- **Project Partners:**
 - Fraunhofer IAIS¹
 - Leibniz Headquarters
 - University Hildesheim (IWIST²)



Former team members:

- Kader Pustu-Iren
- Gerrit Bruns

¹Institute for Intelligent Analysis and Information Systems

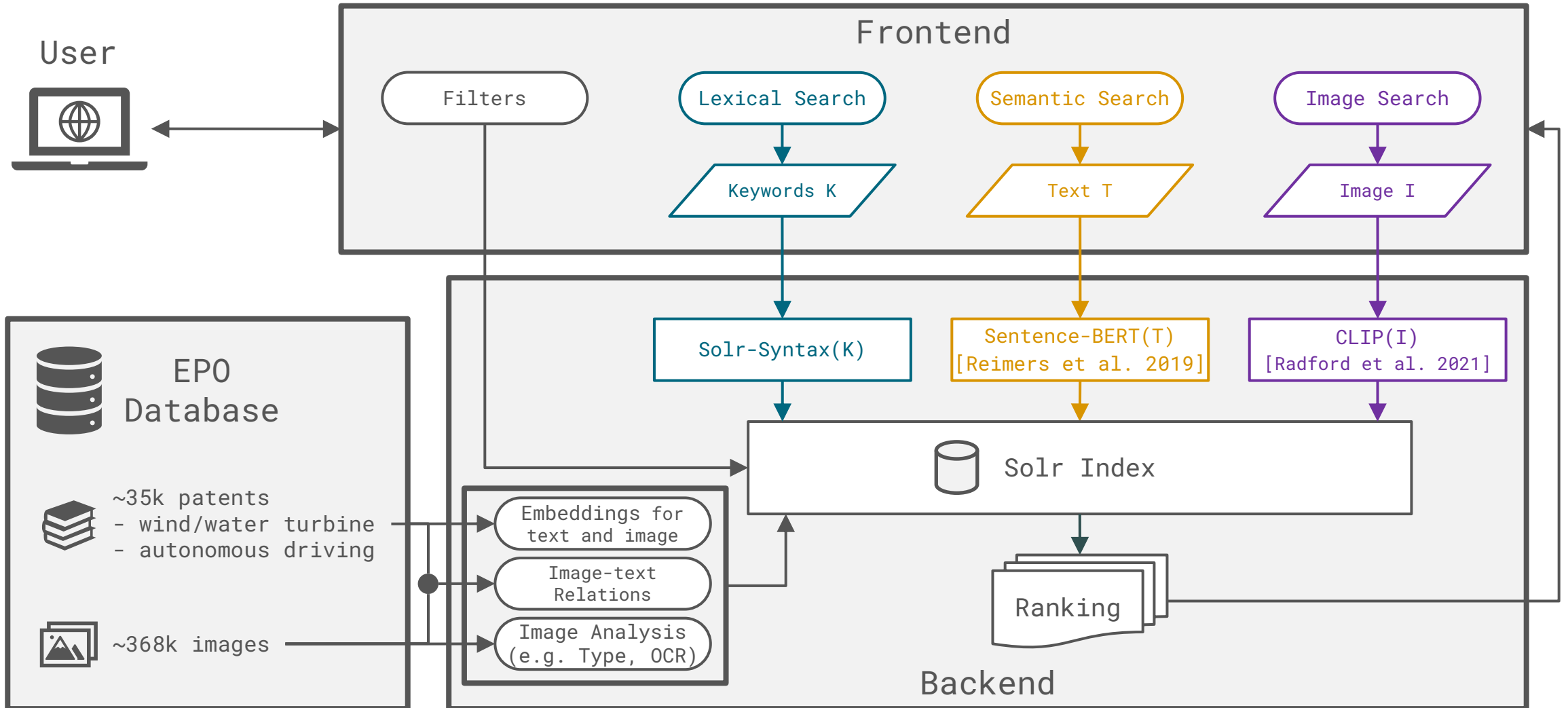
²Department for Information Science and Natural Language Processing

Project Goals

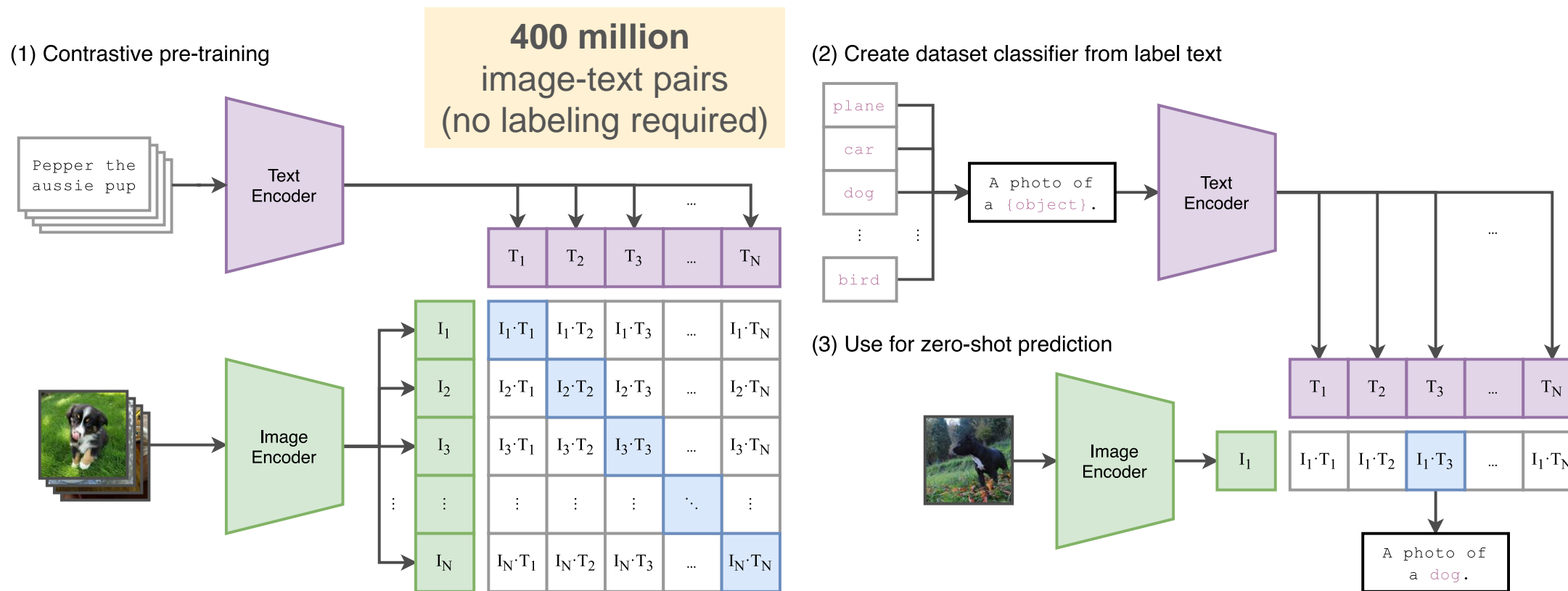


- **Visual and textual search** for patents
 - Lexical (keyword-based) text search
 - Semantic text search
 - Image-based semantic search
- Automatic **analysis of visual elements** in patents
 - Basic image analysis (e.g., illustration type, optical character recognition)
 - Advanced image analysis (e.g., visual concept recognition)
- Establishing **semantic image-text relationships** in patents
- Development of a **patent retrieval prototype**

Architecture

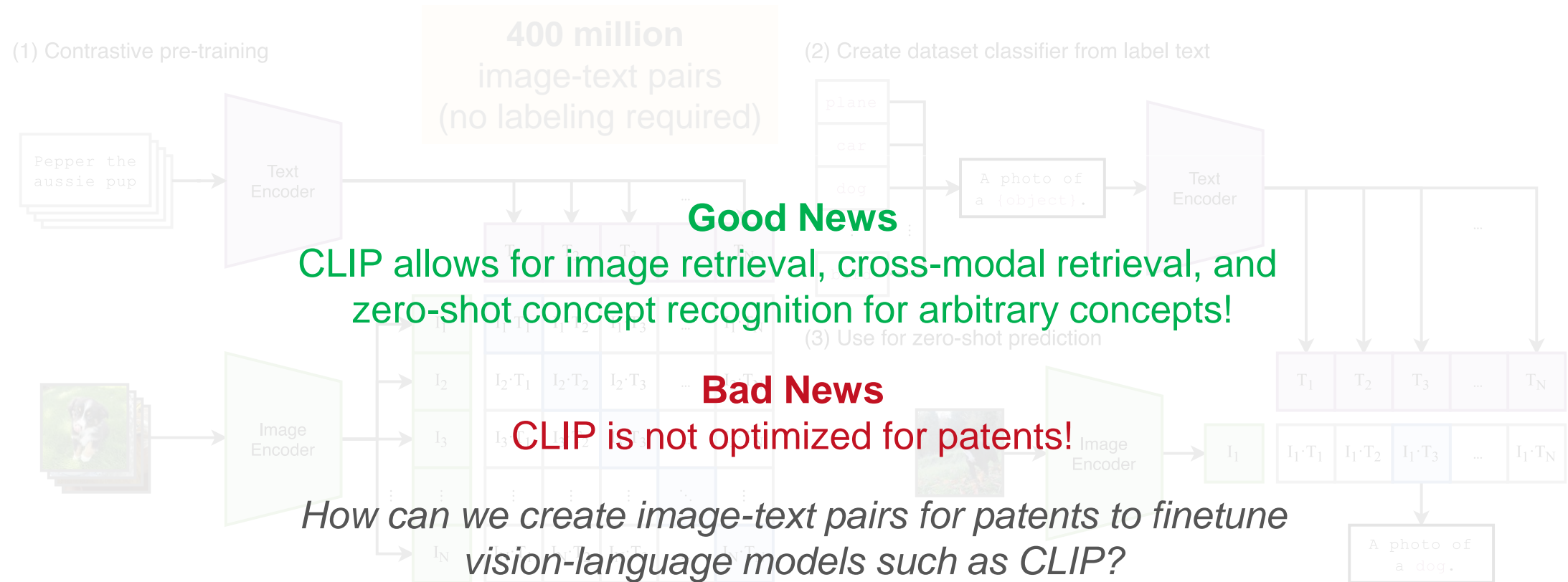


What's CLIP?



Contrastive Language-Image Pretraining
[Radford et al. 2021]

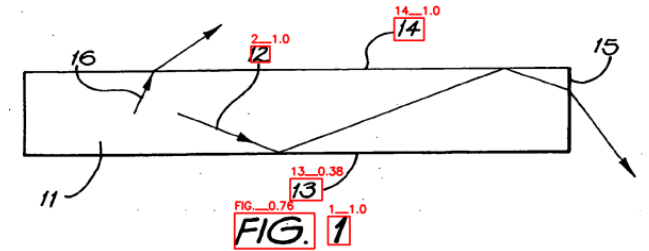
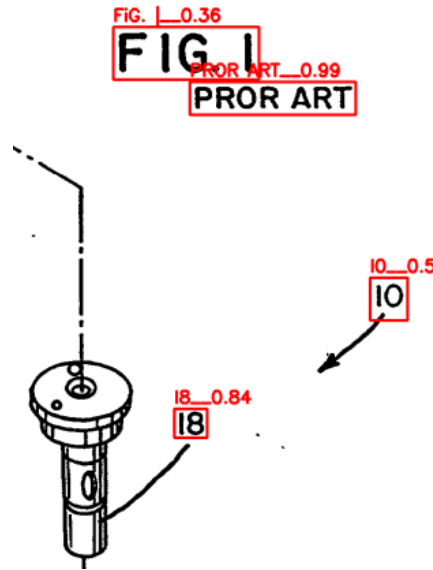
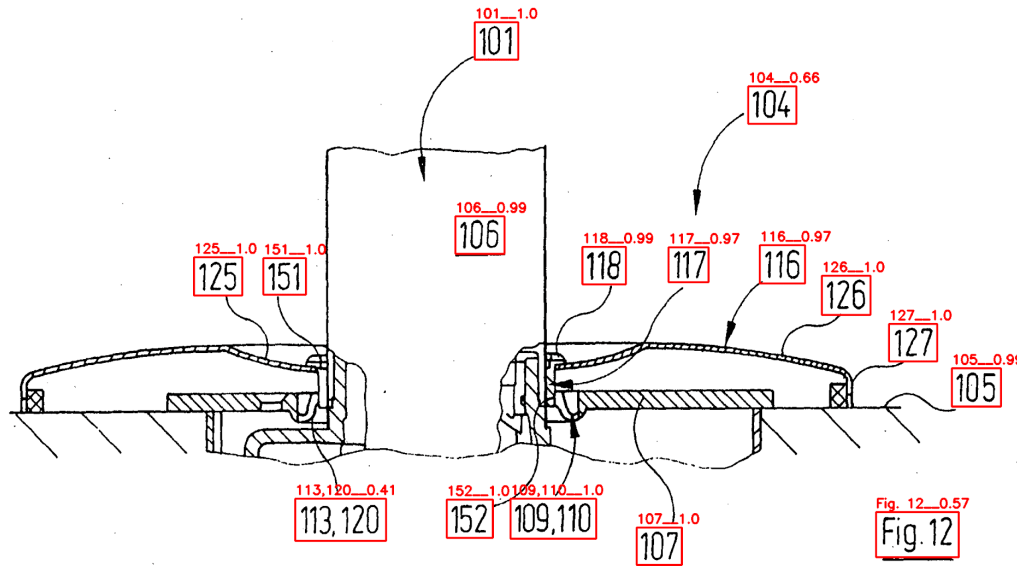
What's CLIP?



Contrastive Language-Image Pretraining
[Radford et al. 2021]

Automatic Identification of Image-Text Relations

- Optical Character Recognition using EasyOCR¹



Existing Problems:

- Missing detections
- Fragmented detections
- Wrong Recognition

Notation

recognition__confidence

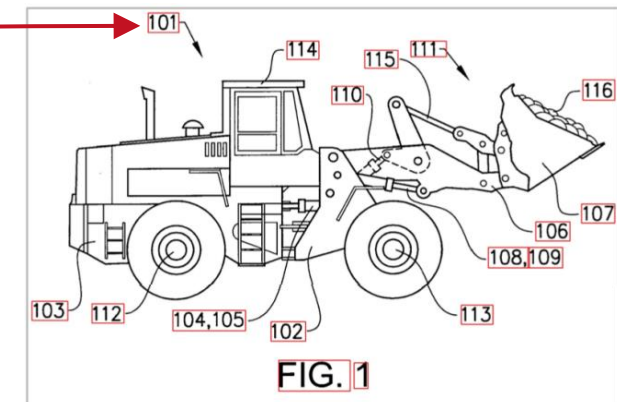
Overlaid Text

¹EasyOCR: <https://github.com/JaidedAI/EasyOCR>

Automatic Identification of Image-Text Relations

- Optical Character Recognition using EasyOCR¹
- Extraction of reference signs in the text
 - Text parsed from xml file of the patent
 - Removal of stop words
 - Regular expression to find numbers that are the candidates for reference signs
 - Mapping of reference signs to OCR outputs
 - Rule-based approach to assign concepts to the reference signs

“The wheel loader 101 comprises an equipment 111 for handling objects or material. The equipment 111 comprises a load-arm unit 106 and an implement 107 in the form of a bucket fitted on the load-arm unit. A first end of the load-arm unit 106 is pivotally connected to the front vehicle section 102. The implement 107 is pivotally connected to a second end of the load-arm unit 106.”...



¹EasyOCR: <https://github.com/JaidedAI/EasyOCR>

DEMO TIME

VisPat Home

Search Results (5539 Patents)

Query Image

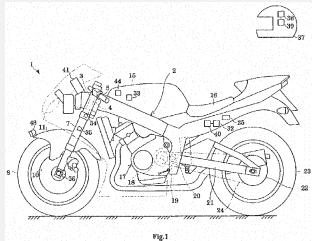
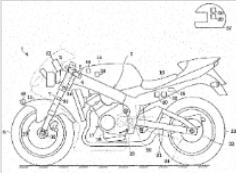
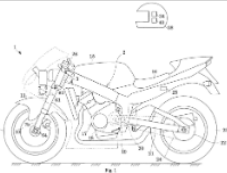
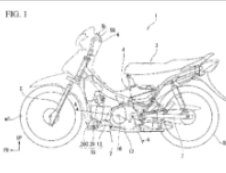
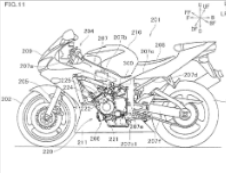
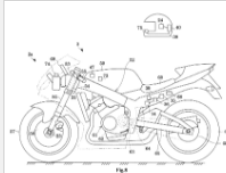
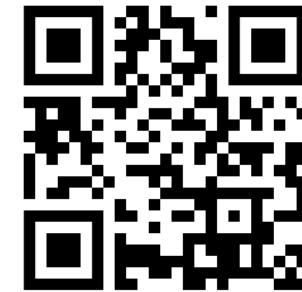


Fig. 1

<input type="checkbox"/> EP17181433A1 1 100.00%	<input type="checkbox"/> EP14850731A1 2 91.86%	<input type="checkbox"/> EP13877372A1 3 88.33%	<input type="checkbox"/> EP17747470A1 4 88.19%	<input type="checkbox"/> EP13755559A1 5 87.32%
 Fig. 1 RIDER CHARACTERISTICS DETERMINING METHOD	 Fig. 1 DRIVING SKILL EVALUATION METHOD, DRIVING SKILL EVALUATION PROGRAM, DRIVING SKILL EVALUATION DEVICE, AND VEHICLE EQUIPPED THEREWITH	 FIG. 1 GENERATOR MOTOR UNIT, POWER OUTPUT ENGINE, AND VEHICLE	 FIG. 11 LEANING VEHICLE	 Fig. 1 HOST COMPUTER, OPERATION- SKILL DETERMINATION SYSTEM, OPERATION-SKILL DETERMINATION METHOD, AND OPERATION-SKILL DETERMINATION PROGRAM

VisPat



Prototype

Example 1: Lexical Search

Search Results (332 Patents)



Current Query:

(text:windmill AND section:description)

EP09700678B1

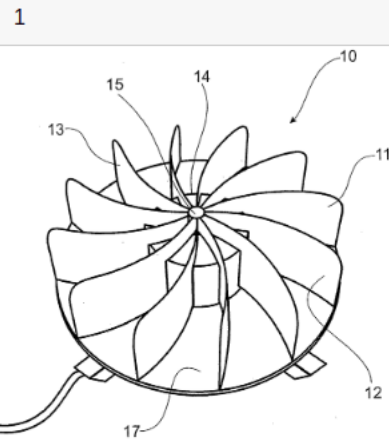
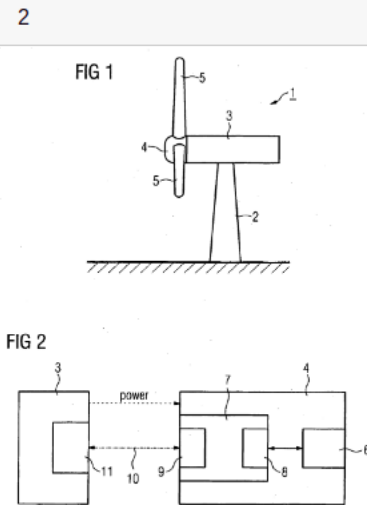


FIG. 1

TURBINE ASSEMBLY

EP07012714B1



Method for controlling of at least one element of a first component of a wind turbine, control device and use of the control device

EP07717965B1

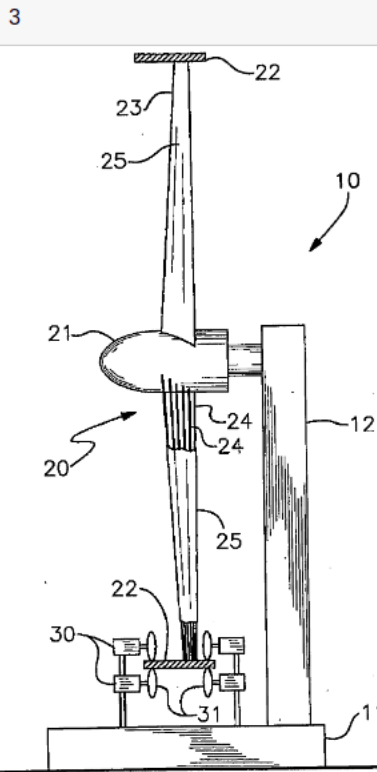
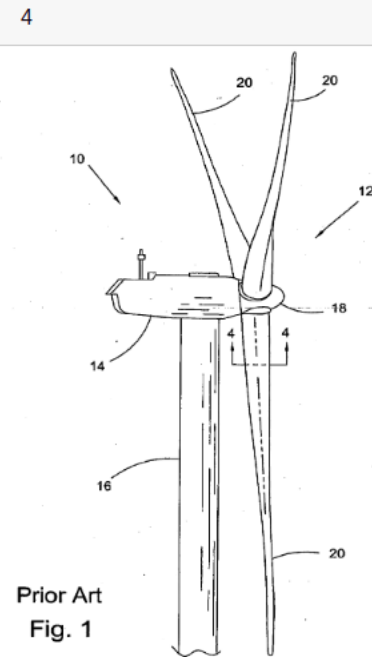


Fig. 1

WIND TURBINE

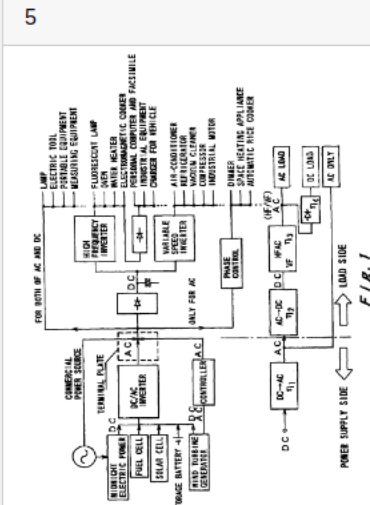
EP07796806B1



Prior Art
Fig. 1

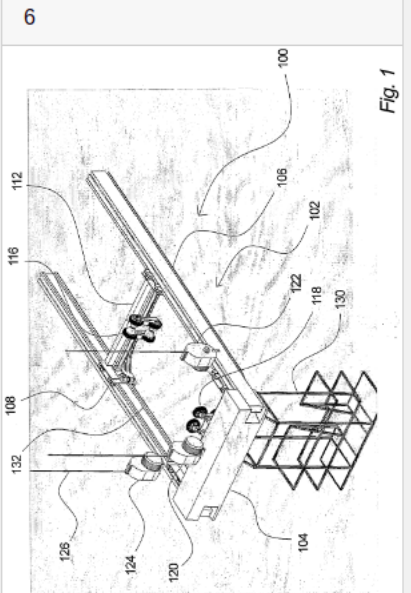
METHOD AND APPARATUS FOR GRID LOSS RIDE-THROUGH FOR WIND TURBINE PITCH CONTROL SYSTEM

EP03254832B1



Co-generated power supply system

EP07702460B1



DEVICE FOR ENABLING ACCESS TO A STRUCTURE ABOVE GROUND LEVEL

Example 2: Lexical Search (with bool connection)

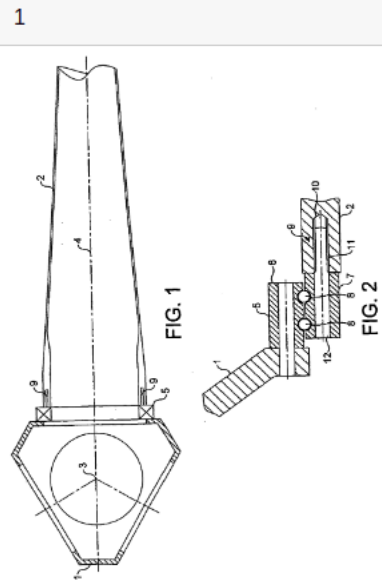
Search Results (168 Patents)



Current Query:

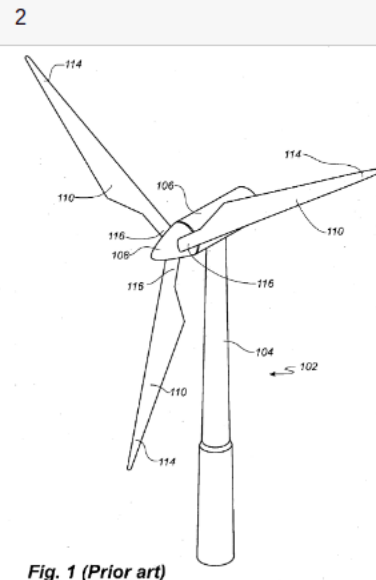
(text:windmill AND section:description) AND (text:turbine AND section:description)

EP09785200B1



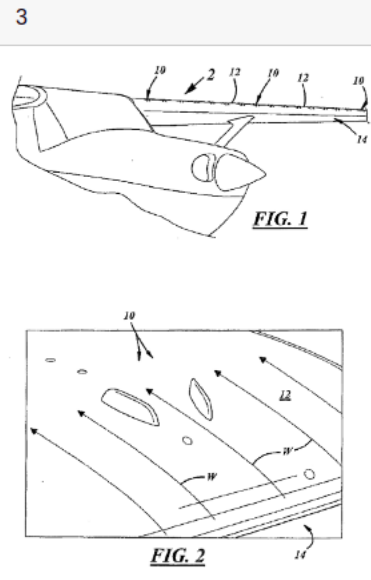
A WIND TURBINE ROTOR

EP09734051B1



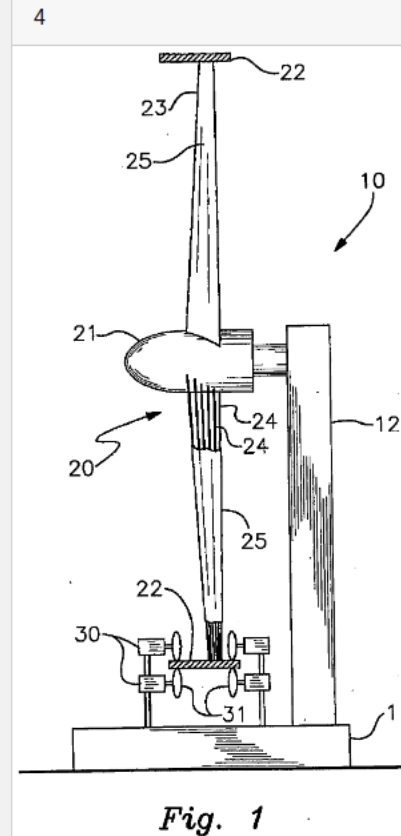
UPWIND WIND TURBINE WITH
BLADES SUPPORTED ON THE
LEEWARD SIDE

EP11190780B1



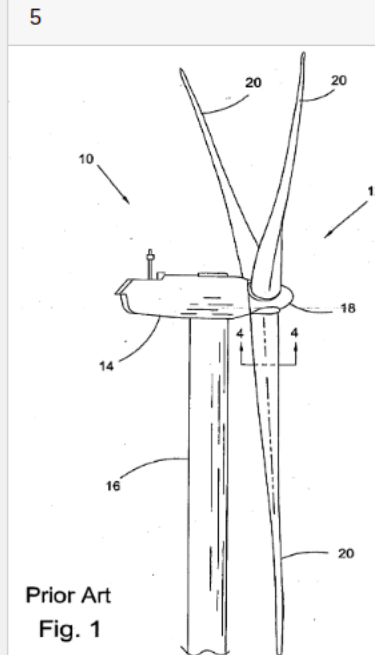
Plasma actuated vortex generators

EP07717965B1



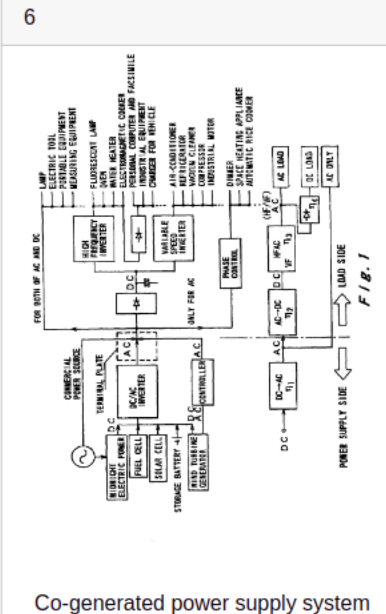
WIND TURBINE

EP07796806B1



METHOD AND APPARATUS FOR
GRID LOSS RIDE-THROUGH FOR
WIND TURBINE PITCH CONTROL
SYSTEM

EP03254832B1



Co-generated power supply system

Example 3: Image-Text Relations

	US-A- 4 183 715 US-A- 4 403 916	US-A- 4 316 699 US-A- 5 354 175
Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid.		

DESCRIPTION

RefName ▾ wind turbine

The search found 4 matches in 1 panels

Reference signs Wind Turbine 102 ✕

106 nacelle

110 blades

108 hub

104 tower

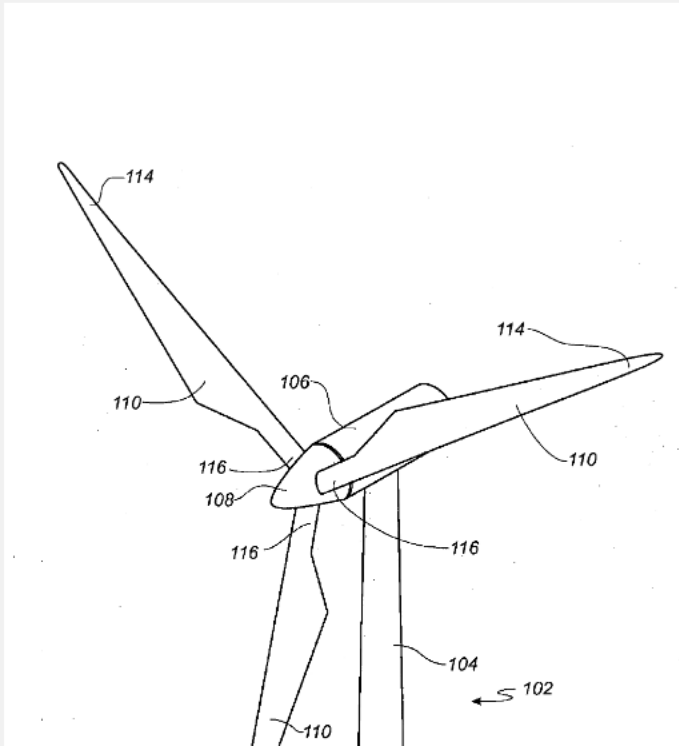
102 wind turbine

Add to indocument search

Best mode for carrying out the invention

[0034] The **wind turbine 102** illustrated in **Fig. 1** is a modern conventional turbine that comprises a **tower 104** supporting a **nacelle 106** (also called the windmill housing). A substantially horizontal main shaft projects from the **nacelle 106**, a rotor being mounted on said shaft, said rotor comprising a **hub 108** and two or more **blades 110**. The rotor can be made to rotate by the wind. Preferably, the **wind turbine 102** is a so-called upwind turbine, where the wind impinges on the rotor before it impinges on the **tower 104**, and where the **nacelle 106** is able to yaw, i.e. rotate around a vertical axis with respect to the **tower 104**, the rotor thereby adjusting itself to the wind direction at any given moment. Moreover, the wind turbine is preferably provided with three **blades 110** extending substantially radially outwards from the **hub 108** and being formed as so-called shell bodies made of a polymer material reinforced with glass fibres, carbon fibres, or other reinforcement fibres. Each blade 110 comprises a root section 116 near the **hub 108** and a blade tip 114.

Wind Turbine 102



Example 4: Semantic Text Search

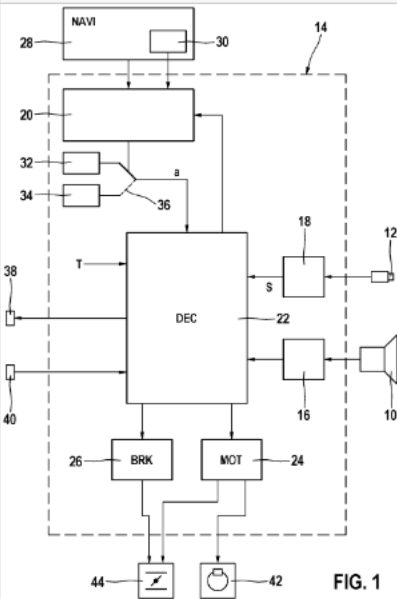
Search Results (2962 Patents)



Current Query:

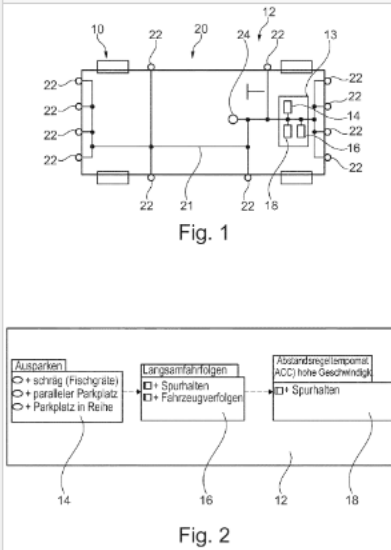
lane assistant for autonomous driving

☐ **EP10787353B1**
1 49.70%



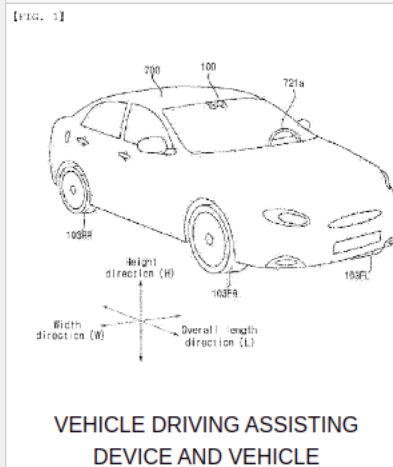
ANFAHRASSISTENT FÜR
KRAFTFAHRZEUGE

☐ **EP18168865B1**
2 43.33%

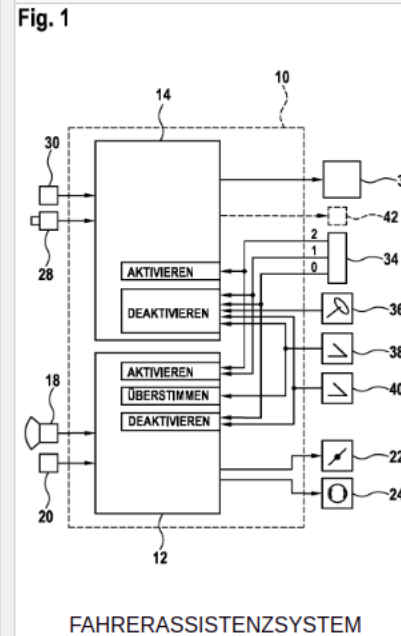


FAHRASSISTENZSYSTEM FÜR EIN
FAHRZEUG MIT EINEM
AUSPARKASSISTENTEN UND
EINEM LANGSAMFAHR-
FOLGEASSISTENTEN

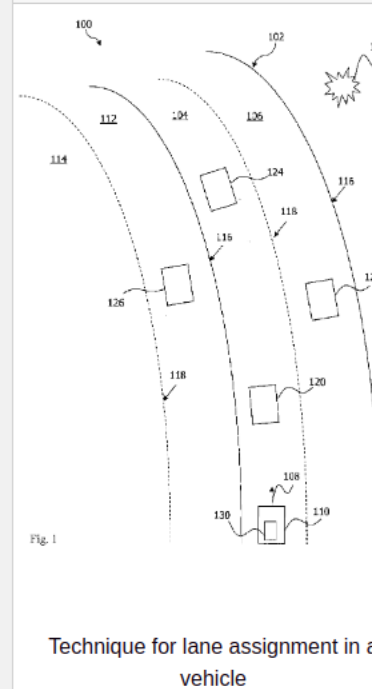
☐ **EP16796646A1**
3 42.32%



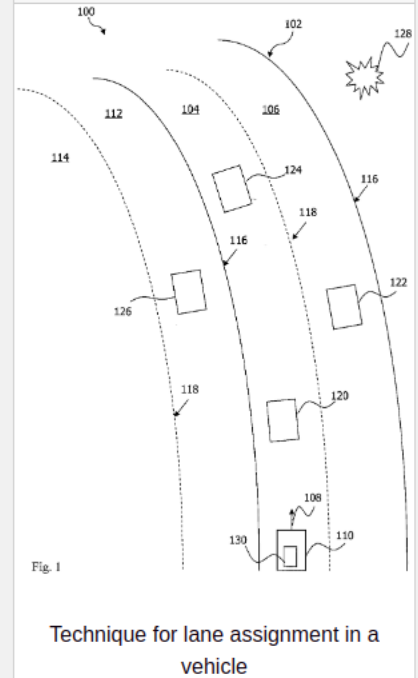
☐ **EP13753282B1**
5 41.21%



☐ **EP13176640A1**
6 41.19%



☐ **EP13176640B1**
7 41.19%



☐ **EP09717688B1**
8 41.18%

FIG. 1

☐ **EP16743313B1**
9 40.97%

FIG. 1

☐ **EP16743313A1**
10 40.97%

FIG. 1

Example 5: Image Search

Search (patents)

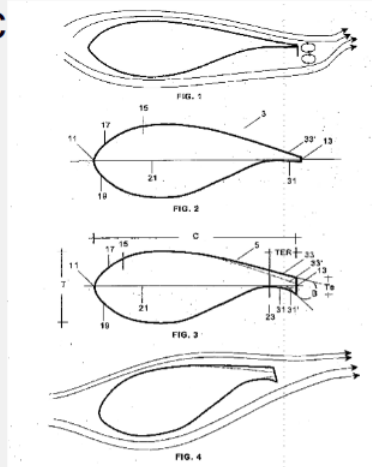
Filter Image Type

- ☐ Block Circuit
- ☐ Drawing
- ☐ Geneseq.
- ☐ Math
- ☐ Symbol
- ☐ Chemical
- ☐ Flowchart
- ☐ Graph
- ☐ Program
- ☐ Table

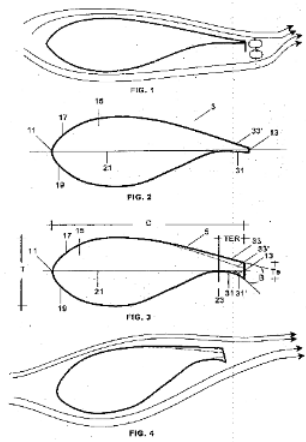
Filter Perspective

- ☐ Front/Rear View
- ☐ Perspective View
- ☐ Side View
- ☐ Plan View

Current Query:

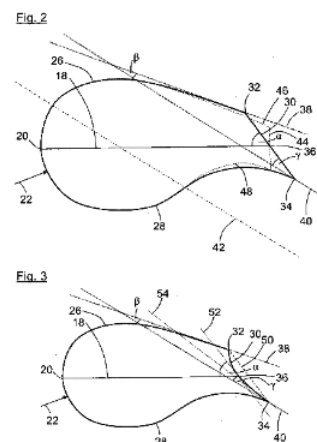


☐ **EP07788560B1**
1 100.00%



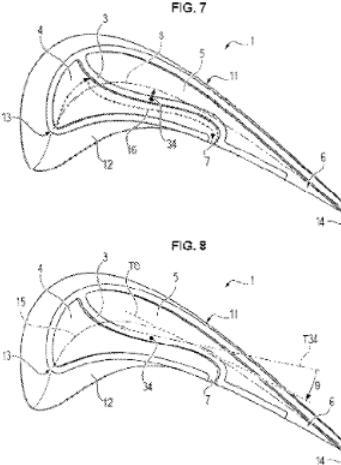
WIND GENERATOR BLADE WITH
DIVERGENT TRAILING EDGE

☐ **EP11007383B1**
2 96.25%



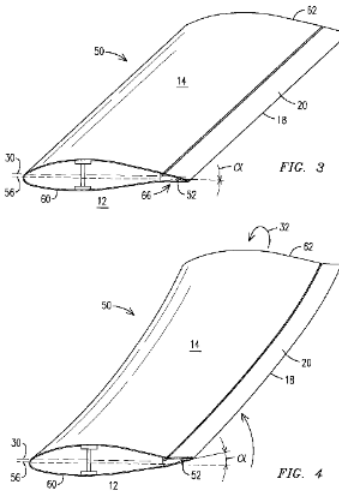
Windenergieanlagenrotorblatt mit
einer dicken Profilhinterkante

☐ **EP15763959B1**
3 95.58%



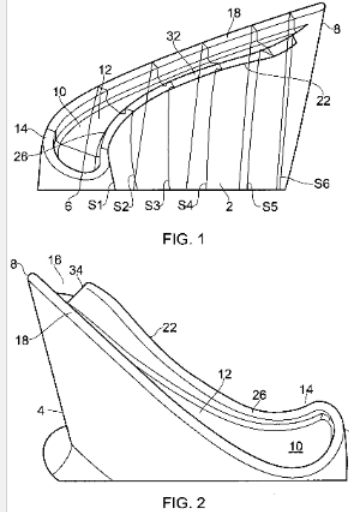
BAIGNOIRE DE SOMMET D'AUBES
D'UNE TURBINE DE
TURBOMACHINE

☐ **EP14167925B1**
4 95.31%



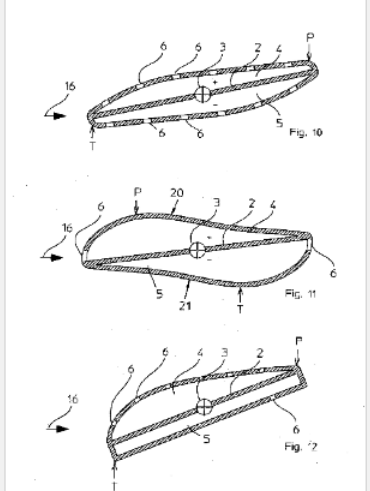
Wind turbine blade having a tensile-
only stiffener for passive control of
flap movement

☐ **EP09251339B1**
5 95.30%



A blade for a rotor having a squealer
tip with a partly inclined surface

☐ **EP12182807B1**
6 95.28%



Klappe zur Anordnung in einem
Strömungskanal einer
klimatechnischen Anlage

Example 6: Image Search

Search Results (20 items)

Filter Image Type

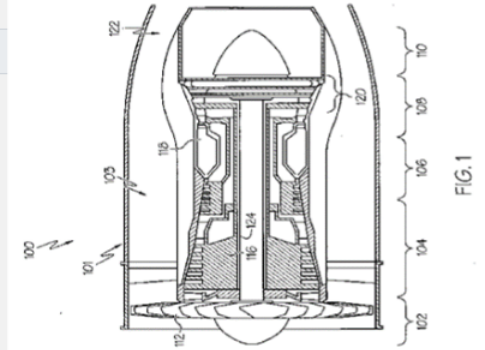
- ☐ Block Circuit
- ☐ Drawing
- ☐ Geneseq.
- ☐ Math
- ☐ Symbol
- ☐ Chemical
- ☐ Flowchart
- ☐ Graph
- ☐ Program
- ☐ Table

Filter Perspective

- ☐ Front/Rear View
- ☐ Perspective View
- ☐ Side View
- ☐ Plan View

Current Query:

Patent Application Publication Apr. 22, 2010 Sheet 1 of 5 US 2010/0095680 A1



☐ **EP10741899B1**

1 83.04%

VERFAHREN UND VORRICHTUNG ZUM TRANSPORT UND ZUR

☐ **EP16275119B1**

2 81.91%

LIGHTING ARRANGEMENT WITH BATTERY BACK-UP

☐ **EP07764453B1**

3 81.46%

A HANDLING SYSTEM FOR A WIND TURBINE NACELLE, METHODS FOR TRANSPORT AND VERTICAL

☐ **EP09826372B1**

4 81.42%

Fig. 13

☐ **EP18211829A1**

5 81.05%

ANORDNUNGEN ZUM TRANSPORT VON ROTORBLÄTTERN EINER WINDTURBINE

☐ **EP19174869A1**

6 80.54%

LUBRICANT DISPENSER FOR A

Summary



- **Platform for image-based, lexical, and semantic search** in patents
- Based on **state-of-the-art approaches** from natural language processing and computer vision
- **Automatic extraction of image-text relations** (linking of image components to keywords)
- **Faceted search** based on image information (e.g., illustration type)

Future Work

- **Multimodal and cross-modal search**
 - Proof-of-concept using a **multimodal re-ranking approach** [Pustu-Iren et al. 2021]
- **Adaption of vision-language models to patents** (e.g., via fine-tuning with image-text pairs)
- **Image retrieval based on image or object regions**
- **Explainability** of retrieval results

ViP@Scale

Visual and Multimodal Patent Search at Scale

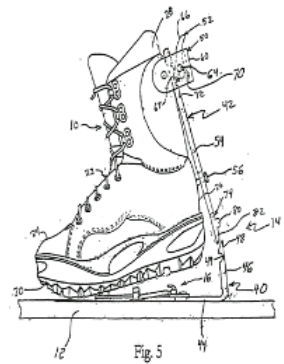


Enter your query here...

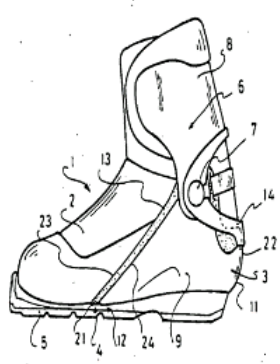
a shoe for skiing

OR

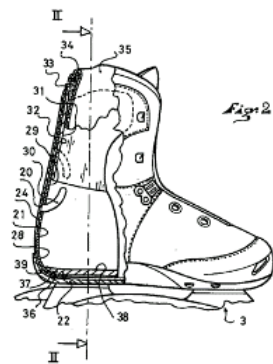
☐ Upload an image



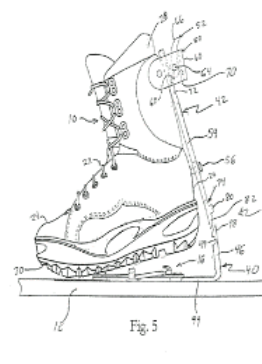
Patent: EP1033085



Patent: EP1166668



Patent: EP0956788



Patent: US0200100093



Patent: EP1120055



Demo Code

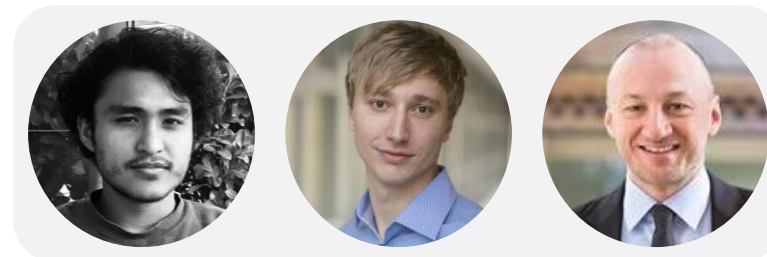
ViP@Scale – Project Overview



Visual and Multimodal Patent Search at Scale



- **Funding:** European Patent Office (EPO)
- **Duration:** April 2023 – March 2025
- **Management:** Prof. Dr. Ralph Ewerth
- **Team:** Sushil Awale, Dr.-Ing. Eric Müller-Budack
- **Project Goals:**
 - **Object localization and instance segmentation** in patent figures
 - **Large-scale (multimodal) representation models for patent images**
 - Application of the multimodal models to downstream tasks, e.g., **(cross-modal) image retrieval**
 - **Evaluation on large-scale test** sets using patent collections from the EPO



Current Status

Completed

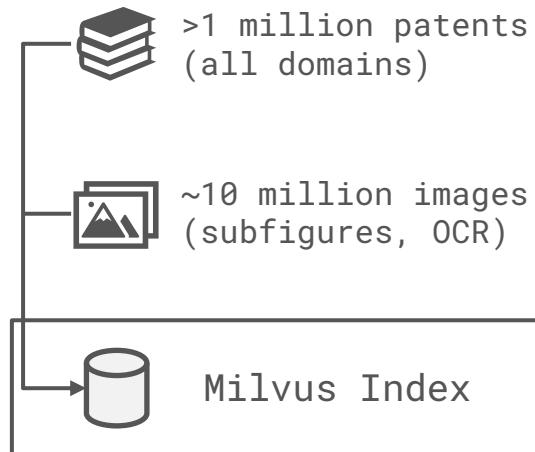
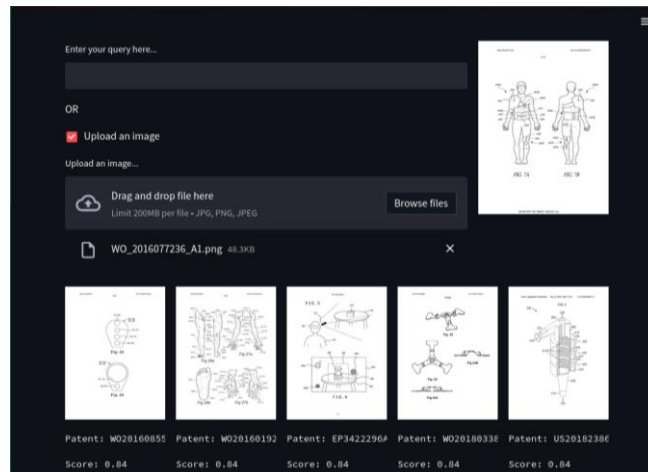
1. Data acquisition & preprocessing
2. Data analysis
3. Demo application for:
 - Image retrieval
 - Cross-modal image retrieval


In Progress

1. Index patent figures
2. Design evaluation pipeline
3. Establish baseline score

Next Steps

1. Generate image-text pairs
2. Adopt VLMs for patents
3. Other approaches (e.g.,
contrastive learning for images)



 **InstructBLIP**
[Dai et al. 2023]

 **Flamingo**
[Alayrac et al. 2022]

Vision-Language Models
(VLMs)

Preliminary Results – Cross-modal Retrieval

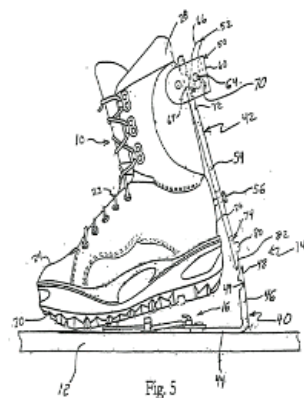
Patent Search

Enter your query here...

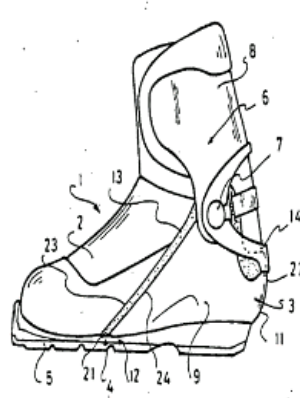
a shoe for skiing

OR

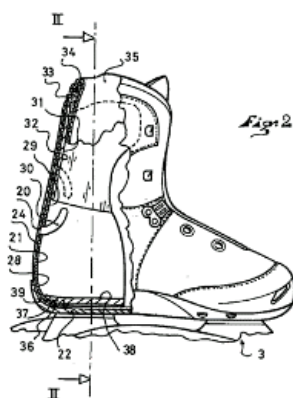
☐ Upload an image



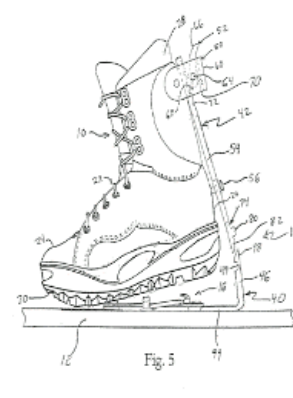
Patent: EP1033085



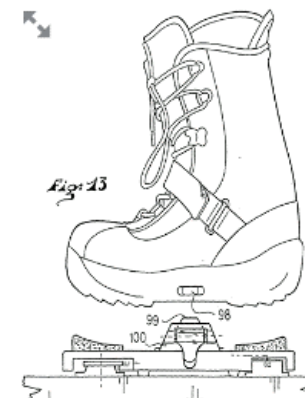
Patent: EP1166668



Patent: EP0956788



Patent: US0200100093



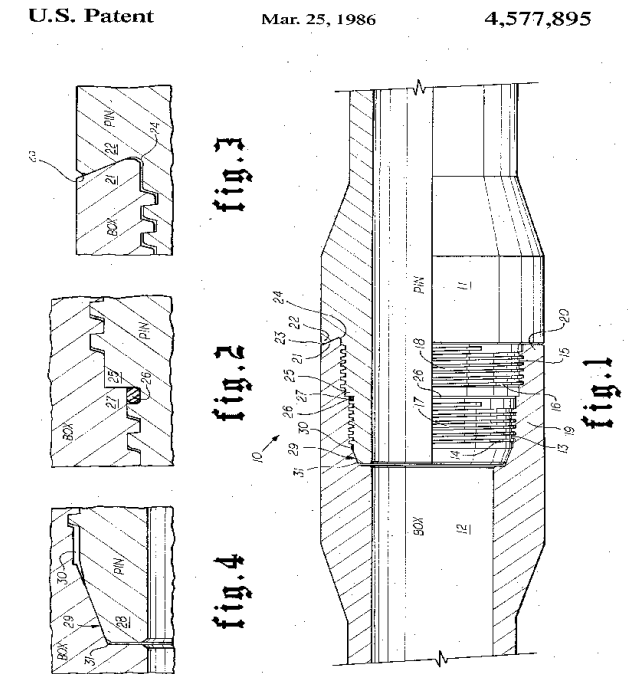
Patent: EP1120055

Fine-tuning of VLMs for Patents

- Adaption of CLIP (or other VLMs) to the patent domain
- Pairs of image and text required
- OCR outputs and **mapping to text** provided by EPO
- Text reference enriched with **context from full sentence**

FIG. 1 is a longitudinal sectional view taken through the joint showing it in engaged position

FIG. 1 shows a pipe connection 10 which includes a pin member 11 and a box member 12 having interengaged threads 13 and 14 in one step, and interengaged threads 15 and 16 in a second step.




- **Problem: Most text references are not descriptive** and not suitable for finetuning VLMs
- **Component labels** from OCR outputs can be used to generate captions for the figures

Enrichment of Image-Text Pairs using LLMs




E

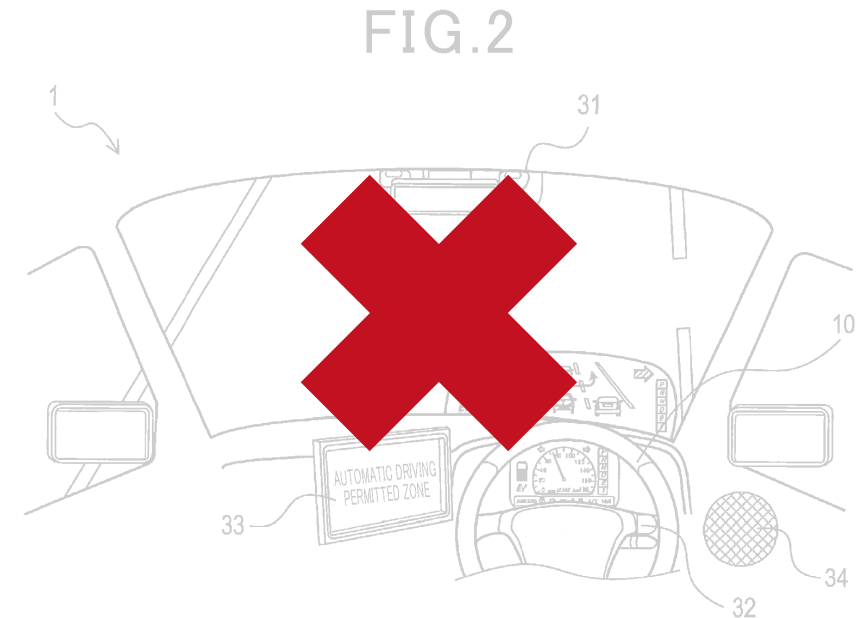
Please create a image caption for a patent with the following keywords:

1 vehicle
10 steering wheel
31 imaging device
32 changeover switch
33 display device
34 speaker



"An innovative vehicle steering wheel system with integrated imaging device (31), changeover switch (32), display device (33), and speaker (34) for enhanced control and communication capabilities."



Enrichment of Image-Text Pairs using LLMs

E

Given a patent on autonomous driving, please create a image caption for a patent with the following keywords:

1 vehicle


10 steering wheel

31 imaging device

32 changeover switch

33 display device

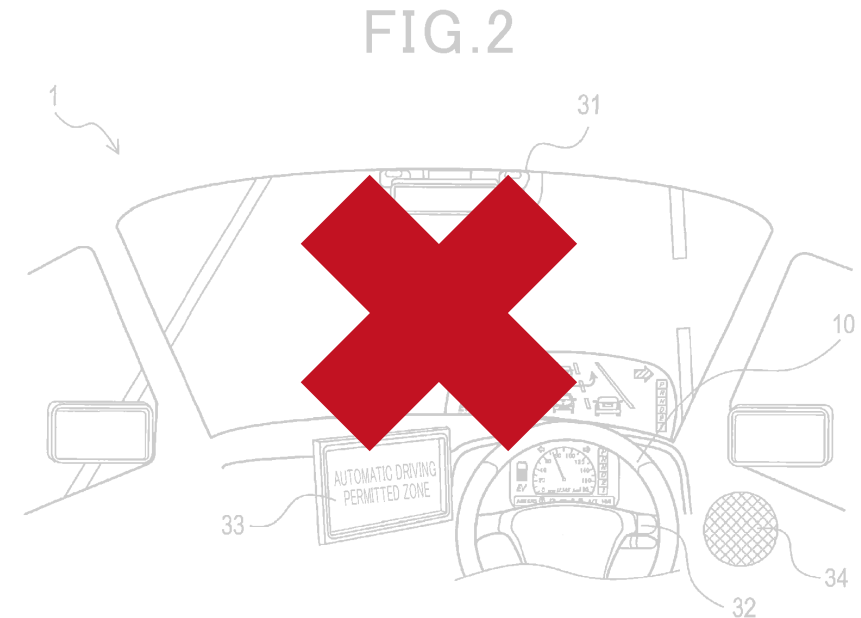
34 speaker



Title: Autonomous Driving System Patent

Image Caption:

Figure 1: An innovative autonomous driving system for vehicles. The patent showcases advanced technology including a steering wheel (10), an imaging device (31) for real-time environmental perception, a changeover switch (32) for manual and autonomous driving modes, a display device (33) providing critical information to the driver, and a speaker (34) for auditory feedback. This cutting-edge system redefines the future of automotive safety and convenience.



Enrichment of Image-Text Pairs using VLMs

E

Can you provide a detailed patent description of this image given that the image components are as follows?
1 vehicle; 10 steering wheel; 31 imaging device;
32 changeover switch; 33 display device; 34 speaker

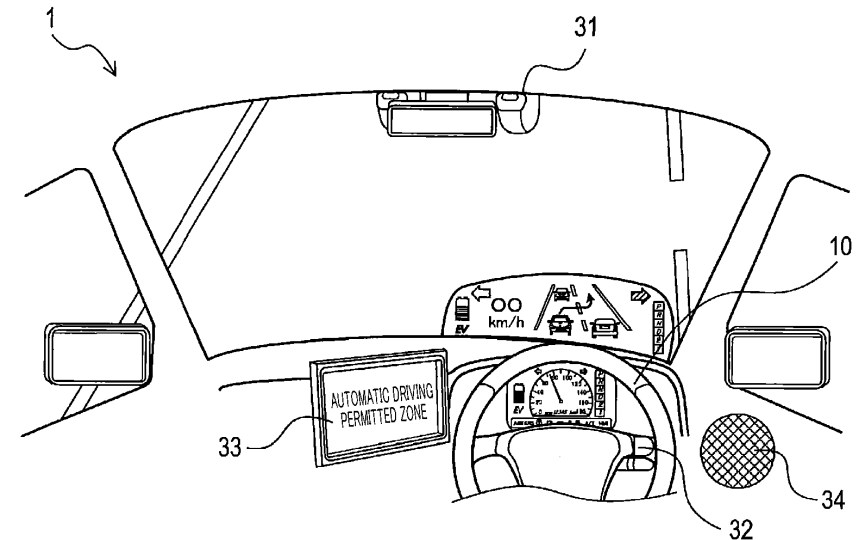


The steering wheel is located on the dashboard, and there is an imaging device located on the steering wheel. There is also a changeover switch located on the steering wheel, which can be used to switch between different modes of operation. A display device is also present on the dashboard, and there is a speaker located on the side of the steering wheel."

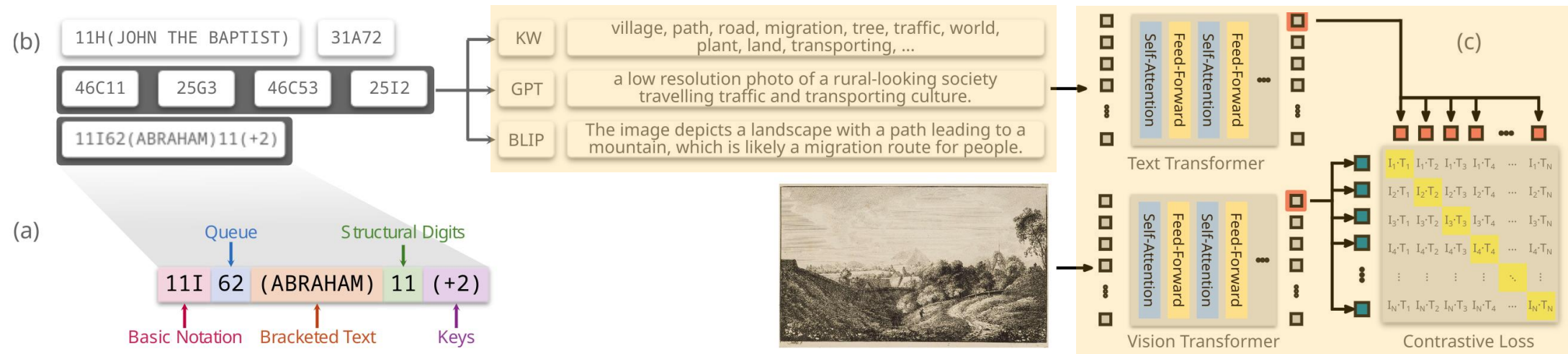
 **InstructBLIP**
[Dai et al. 2023]

Note: Much smaller LLM compared to GPT-3.5 applied (flan-t5-xl)

FIG.2



Large-scale Classification of Art-historical Images



Strategy	# of Training Images per <i>Iconclass</i> Concept			
	> 0	> 10	> 100	> 1000
KW	0.1862	0.2025	0.2545	0.3953
BLIP	0.1922	0.2106	0.2596	0.3961
GPT	0.1902	0.2063	0.2583	0.3916
Baseline (not finetuned) → LAION-400M	0.1845	0.2017	0.2540	0.3936

Outlook

Vision-Language Models (VLMs) for multimodal patent retrieval can provide:

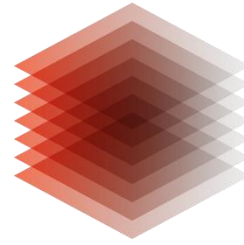
- **better retrieval results** by
 - adopting VLMs to the patent domain
 - re-ranking of the top-k results
- **improve search through user interactions:**

Please only include results that show technical drawings from autonomous driving!

- **better explainability** of retrieval results:

Can you tell me what these patents have in common and what differentiates them?

LEIBNIZ INFORMATION CENTRE
FOR SCIENCE AND TECHNOLOGY
UNIVERSITY LIBRARY



TIB

Thank you for your attention!

References

- [Alayrac et al. 2022] Flamingo: a Visual Language Model for Few-Shot Learning. NeurIPS 2022.
- [Dai et al. 2023] InstructBLIP: Towards General-purpose Vision-Language Models with Instruction Tuning. arXiv preprint 2023.
- [Ghuri et al. 2023] Classification of Visualization Types and Perspectives in Patents. International Conference on Theory and Practice of Digital Libraries,TPDL 2023.
- [Krestel et al. 2021] A survey on deep learning for patent analysis. World Patent Information, 65. 2021.
- [Pustu-Iren et al. 2021] A multimodal approach for semantic patent image retrieval. Patent Text Mining and Semantic Technologies, PatentSemTech 2021.
- [Radford et al. 2021] Learning Transferable Visual Models From Natural Language Supervision. International Conference on Machine Learning, ICML 2021
- [Springstein et al. 2021] iART: A Search Engine for Art-Historical Images to Support Research in the Humanities. ACM Multimedia Conference, ACMMM 2021
- [Springstein et al. 2023] TIB AV-Analytics: A Web-based Platform for Scholarly Video Analysis and Film Studies. International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2023.
- [Vaswani et al. 2017] Attention is All you Need. Annual Conference on Neural Information Processing Systems, NIPS 2017.