

◆ Your membership will expire on February 9, 2024 [Reactivate membership](#)

Using Ontologies & Scene Graphs for Advanced Image Searching



Tharindu Vitharana · [Follow](#)

Published in [Sysco LABS Sri Lanka](#) · 9 min read · Nov 21, 2023

49



...



Search

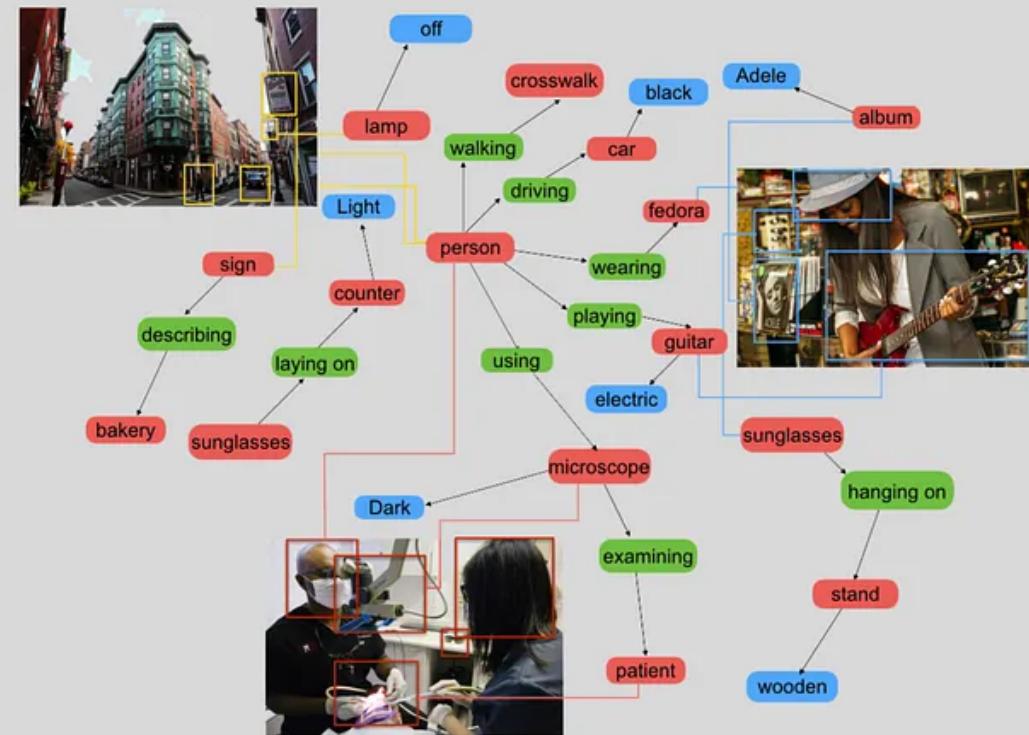
Write



Using Ontologies & Scene Graphs for Advanced Image Searching



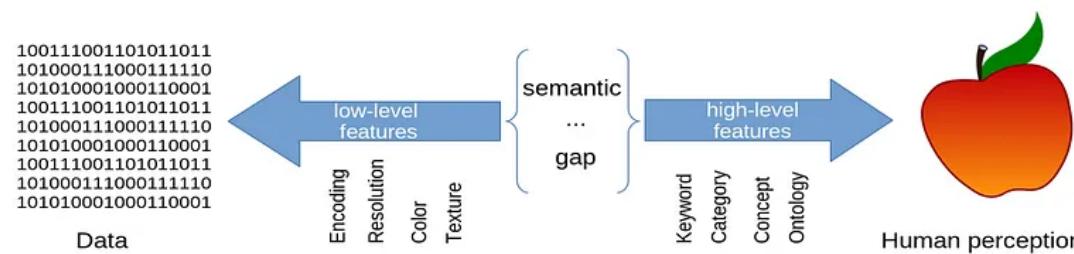
Tharindu Vitharana



In today's digital age, image searching has become an integral part of our online experiences. Whether it's finding the perfect image for a presentation or identifying objects within a photo, users are increasingly demanding more refined and efficient search capabilities. This is where ontologies come into play, offering a powerful tool for advancing image searching to new levels of precision and relevance. In this article, I will guide

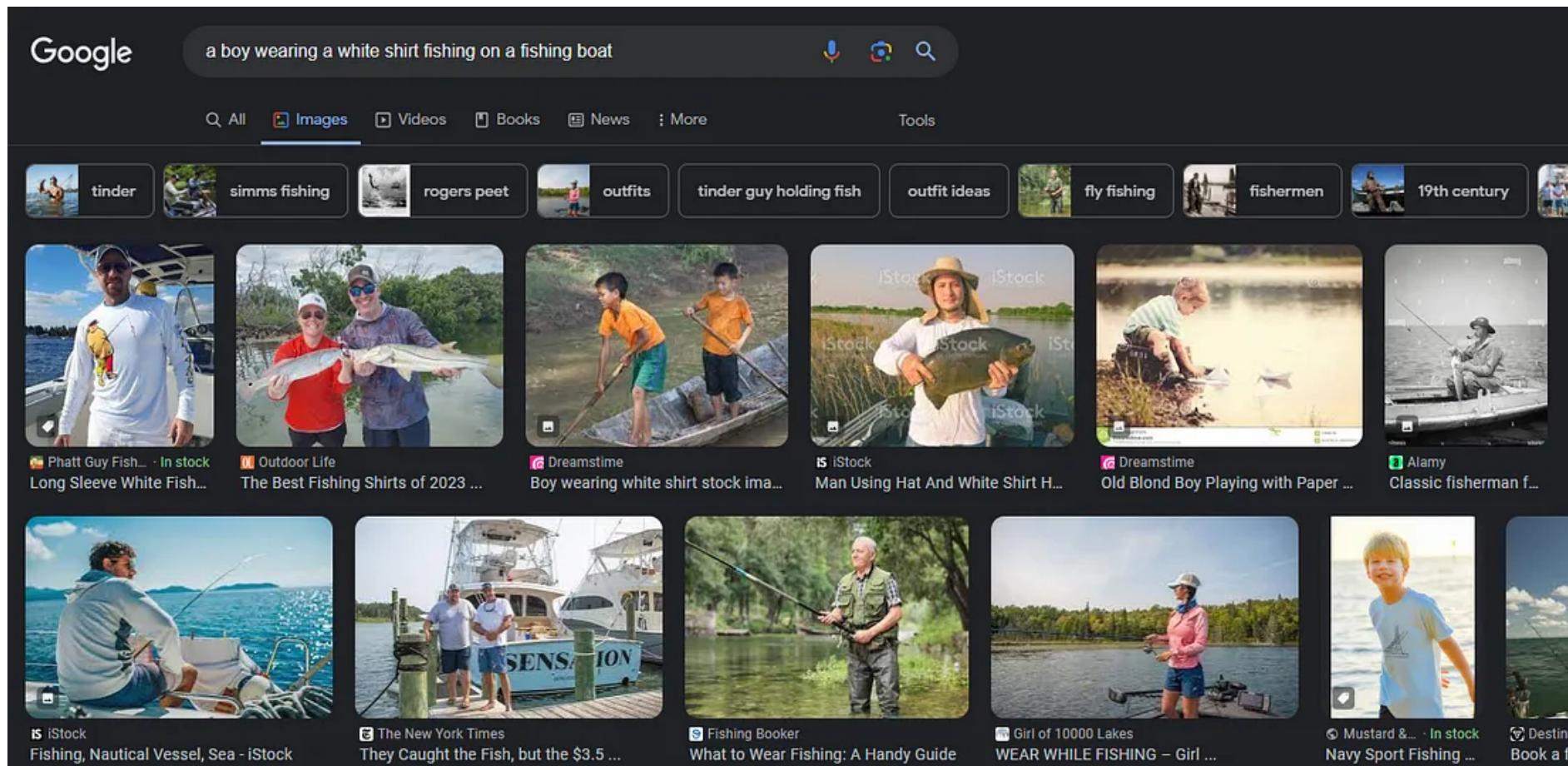
you through the basics of using ontologies and scene graphs to solve problems with current image searching.

Semantic gap is a major problem in image retrieval.



The semantic gap in image searching represents the disparity between the low-level visual features extracted from images, such as colors, textures, and shapes and the high-level semantic concepts that humans associate with those images. Bridging this gap is a critical challenge in the field of computer vision and image retrieval. While machines excel at analyzing pixel data, they struggle to understand the contextual meaning and nuances that humans effortlessly grasp when interpreting images. Addressing the semantic gap involves developing sophisticated algorithms and techniques, often incorporating artificial intelligence and machine learning, to enable computers to comprehend images on a more abstract and conceptual level, making image searching more accurate and context-aware.

The below image shows how Google provides image results for the search query “a boy wearing a white shirt fishing on a fishing boat”.

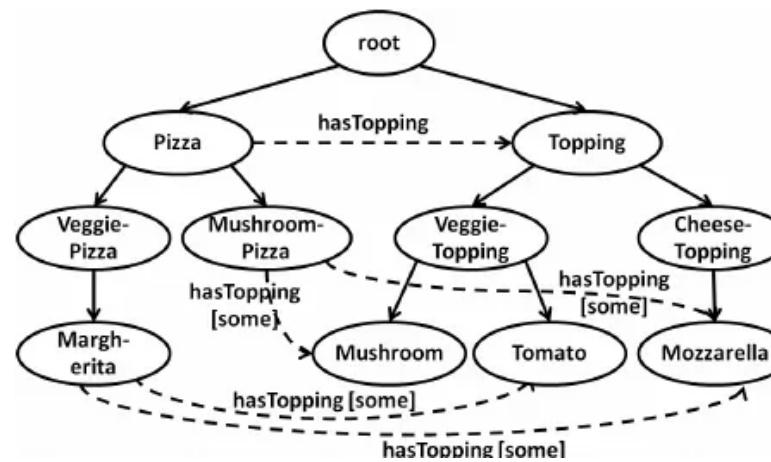


Understanding Ontologies

Before we delve into advanced image-searching techniques, let's briefly explore what ontologies are and why they matter in the context of images.

Ontology refers to a formal representation of knowledge, often in the form of a structured hierarchy of concepts. These concepts are interconnected with defined relationships, allowing for a rich and nuanced understanding of a particular domain.

In the realm of image searching, ontologies can be used to define and organize concepts related to images, making it easier to search, filter, and retrieve images based on their content, context, and meaning. The example below on pizza is the most used example to teach about ontologies in computer science knowledge base development.



Building an Image Ontology

Creating an image ontology is the first step in using ontological techniques for advanced image searching. Here's how to get started:

1. Define Your Domain: Determine the specific domain or subject matter for which you want to improve image searching. This could be anything from wildlife photography to medical imaging.

2. Identify Key Concepts: Within your chosen domain, identify key concepts or categories. For example, in wildlife photography, categories might include different species of animals, habitats, or behaviors.

3. Establish Relationships: Define relationships between concepts. For instance, you might specify that “Tiger” is a subclass of “Big Cats” or that “Rainforest” is part of the “Habitat” category.

4. Add Metadata: Attach metadata to each concept and image. This metadata can include keywords, descriptions, and relevant attributes. For images, metadata could encompass resolution, color profile, or location.

Ontological Image Searching

Once your image ontology is in place, you can leverage it for more advanced image searching. Ontological image searching mainly consists of two components.

1. Storing the Ontology

XML, RDF, RDF Schema, OIL, and OIL+DAML are the earliest web ontology languages, while OWL is the current W3C recommendation. In many

researches, it has been observed that the combination of RDF and OWL can accurately describe the instances and constraints in an ontology.

- **RDF** – Resource Description Framework
- **OIL** – Ontology Interchange Language
- **DAML** – Digital Asset Modeling Language
- **OWL** – Web Ontology Language

2. Querying the Ontology

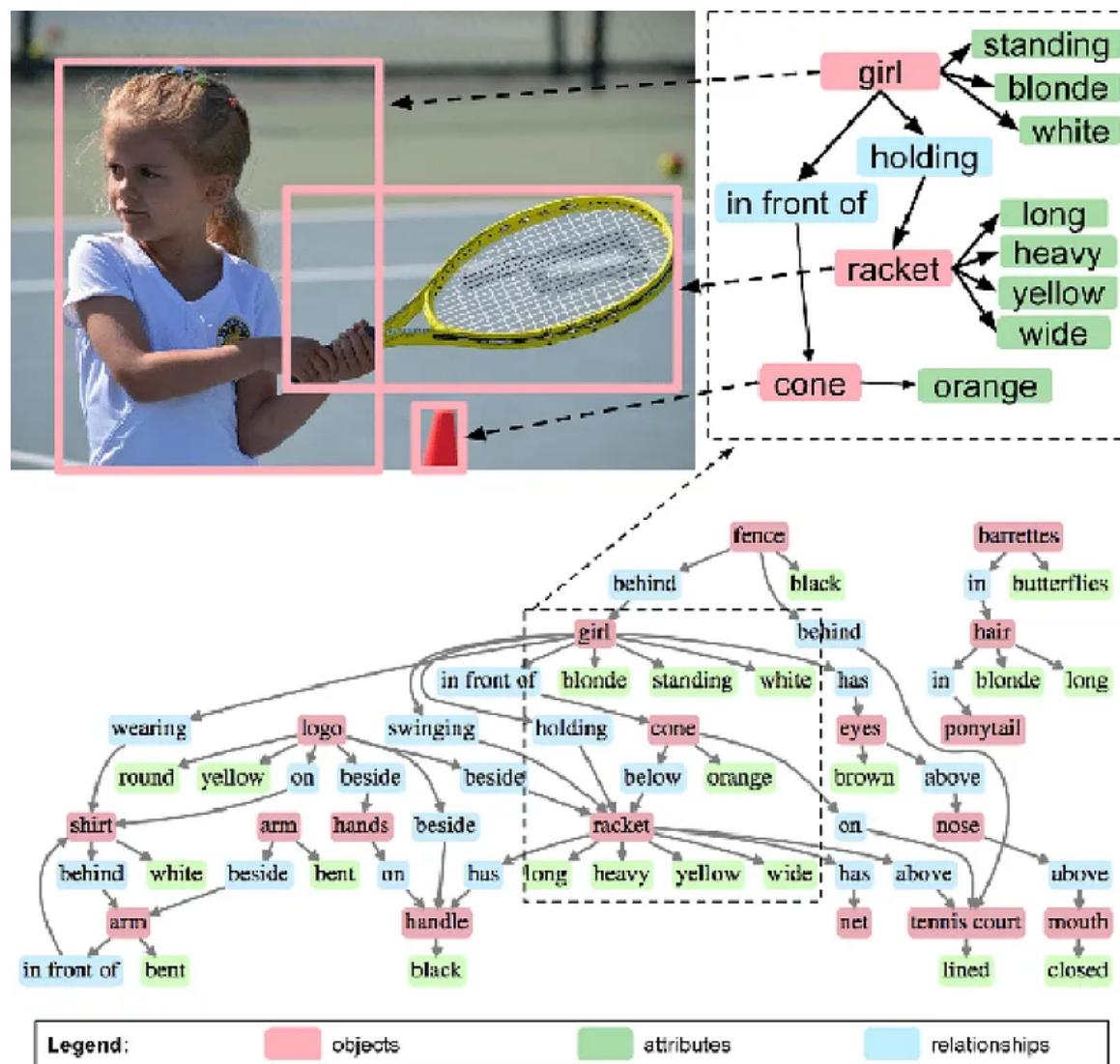
Ontology query languages allow expressions to be written that can be evaluated against an ontology. The queries can be used by knowledge management applications as a basis for inference actions. Existing ontology query languages include OntoQL, SPARQL, SERQL, TRIPLE, and Versa. The SPARQL query language has been adopted by W3C as the means to query ontologies built using RDF2 and has been extended to support OWL format. SPARQL is based on SQL and has the capabilities for querying visual graph patterns along with their conjunctions and disjunctions.

The challenge exists in these 2 steps. Research is ongoing for automating these 2 steps, making these 2 steps more accurate and efficient.

Scene Graphs

Image scene graphs are a data representation and analysis technique in computer vision that captures the spatial and hierarchical relationships between objects within an image. Similar to traditional scene graphs, which describe the structure of a scene, image scene graphs provide a detailed breakdown of the objects in an image and their interconnections. This structured representation enhances the understanding of image content, enabling applications like object localization, image segmentation, and image-based reasoning. Image scene graphs are a powerful tool for advancing computer vision tasks by providing a rich, context-aware description of the visual elements within an image, contributing to improved object recognition and scene understanding.

Below is a more advanced scene graph developed by the Visual Genome open project which will be discussed later in the article.



Ontology VS Scene Graph

Differences

1. Representation Focus:

- **Image Ontology:** Primarily focuses on representing high-level semantic concepts and relationships within images. It aims to capture the meaning and context of objects in images.
- **Scene Graph:** Concentrates on the structural and spatial relationships between objects in a scene. It emphasizes the physical arrangement and organization of objects.

2. Abstraction Level:

- **Image Ontology:** Operates at a higher level of abstraction, providing a conceptual understanding of image content. It may capture concepts like “dog,” “beach,” or “sunset.”
- **Scene Graph:** Works at a lower level of abstraction, dealing with the concrete spatial relationships between objects, such as “dog is on the beach.”

3. Data Structure:

- **Image Ontology:** Typically represented as a hierarchy or network of concepts, often utilizing semantic web technologies like RDF (Resource Description Framework).
- **Scene Graph:** Represented as a graph data structure, with nodes representing objects and edges denoting relationships (e.g., “is-a,” “part-of,” “near”).

Similarities

1. **Semantic Understanding:** Both image ontology and scene graphs contribute to the semantic understanding of image content. They help

computers comprehend the meaning and relationships of objects within images.

2. Enhanced Search and Retrieval: Both concepts improve the efficiency and accuracy of image search and retrieval by providing a structured representation of image content. Users can search for images based on concepts or spatial relationships.

3. Facilitation of Computer Vision Tasks: Image ontology and scene graphs serve as valuable tools in various computer vision tasks, aiding in object recognition, image segmentation, and scene understanding.

Image ontology and scene graphs are distinct but complementary approaches to understanding image content. Image ontology emphasizes semantic concepts and relationships, while scene graphs focus on the spatial and structural aspects of objects in a scene.

Visual Genome Open Project

Visual Genome is a remarkable open-source project at the intersection of computer vision and natural language processing (NLP). It aims to create a comprehensive knowledge base for visual understanding by connecting images with rich textual descriptions. Developed collaboratively by researchers and the wider community, Visual Genome provides a vast dataset of images, each accompanied by detailed annotations that describe objects, relationships, and attributes within the image. This valuable

resource has fueled advancements in a wide range of applications, including image captioning, scene understanding, and visual question-answering systems. By making this data freely available, Visual Genome has played a pivotal role in fostering innovation and research in the field of computer vision, enabling the development of more intelligent and context-aware visual recognition systems.

Link to the project:-

<https://homes.cs.washington.edu/~ranjay/visualgenome/index.html>

Visual Genome is a more advanced open-source project in which datasets and the code is available for experimentation. Visual Genome project can be used for 2 main purposes;

1. Advanced image search
2. Image description generation

Semantic Image retrieval is still an ongoing research area with a lot of attention. In this article, we presented basic concepts on how ontologies and scene graphs are useful to solve semantic gap problems. Combining these two approaches can enhance the interpretation of images for a wide range of applications in computer vision and image processing.

Advanced image description technology For Warehouse Operations

Advanced image description technology can significantly enhance various aspects of warehouse operations. Here are some use cases for advanced image description in warehouse management:

1. Inventory Management:

Automated Inventory Tracking: Cameras can capture images of items in the warehouse, and advanced image description algorithms can identify and describe these items. This allows for real-time tracking of inventory, including its location, condition, and quantity.

2. Quality Control:

Product Inspection: Image analysis can be used to inspect products for defects, ensuring that only quality items are shipped to customers.

Label Verification: Automated reading and verification of product labels and barcodes can reduce errors and improve inventory accuracy.

3. Packing and Shipping:

Package Dimension Verification: Advanced image description can measure the dimensions of packages to ensure they match the shipping labels, helping to reduce shipping errors and costs.

Load Optimization: Images of packed pallets or containers can be analyzed to optimize load distribution, maximizing space utilization and minimizing damage during transit.

4. Order Picking:

Picking Verification: Warehouse employees can use image-based systems to verify that they are picking the correct items and quantities, reducing errors.

Visual Guidance: Workers can receive visual instructions through augmented reality or smart glasses to locate and pick items more efficiently.

5. Safety and Security:

Intrusion Detection: Cameras with advanced image description can detect unauthorized personnel or vehicles in restricted areas, enhancing security.

Safety Compliance: Monitoring for safety compliance, such as checking for the proper use of safety equipment or identifying potential hazards.

6. Route Optimization:

Forklift Guidance: Cameras and image description technology can guide forklift operators by providing real-time information on optimal routes and avoiding collisions.

Traffic Management: Monitoring and managing the flow of vehicles and personnel in the warehouse to minimize congestion and enhance efficiency.

7. Returns Processing:

Item Condition Assessment: When a returned product is received, images can be analyzed to determine its condition and whether it can be restocked or needs refurbishment.

8. Space Utilization:

Rack and Shelf Optimization: Image-based analysis can help warehouse managers optimize the placement of products on racks and shelves to maximize storage capacity.

9. Documentation and Compliance:

Record Keeping: Images can be used to document the condition of items during shipping and receiving, serving as visual evidence in case of disputes or damage claims.

Regulatory Compliance: Advanced image description can help ensure compliance with regulations related to labeling, storage, and handling of certain products.

10. Training and Maintenance:

Training Materials: Advanced image description can be used to create training materials and documentation for warehouse staff.

Equipment Maintenance: Images of warehouse equipment and machinery can be used to identify maintenance needs and schedule repairs proactively.

Below is a possible use case scenario with an image description that can be used for warehouse safety. If any danger-related text is generated then it could trigger an alert.

Safe warehouse operation



Unsafe warehouse operation



These use cases highlight how advanced image description technology can improve accuracy, efficiency, and safety in warehouse operations while reducing human error and enhancing overall productivity. Integrating image-based systems with warehouse management software can provide a comprehensive solution for modern warehouse management.

AI

Image Processing

Computer Vision

Ontology

Image Retrieval



Written by Tharindu Vitharana

[Follow](#)

181 Followers · Writer for Sysco LABS Sri Lanka

On a mission to inspire others through writing about tech insights and self-improvement tips on Medium.

More from Tharindu Vitharana and Sysco LABS Sri Lanka



 Tharindu Vitharana

The Gemini AI Revolution: Reshaping the Landscape of...

A shock ripples through the silicon gorge, a whisper of mutiny manage operations the...

8 min read · Dec 25, 2023

👏 12



...



 Kaveesha Dinamidu in Sysco LABS Sri Lanka

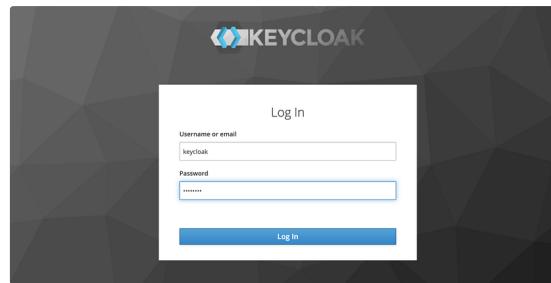
State management in Micro-Frontends

8 min read · Jun 6, 2023

👏 120



...



Arkam ahamed in Sysco LABS Sri Lanka

Setting up Keycloak and Securing Spring Boot Rest APIs

Overview

12 min read · Apr 18, 2023



 Tharindu Vitharana

10 Strategies to Enhance Work Consistency and Achieve...

Introduction:

9 min read · Dec 3, 2023



4



1



...



123



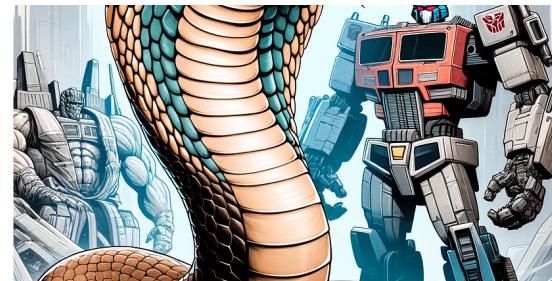
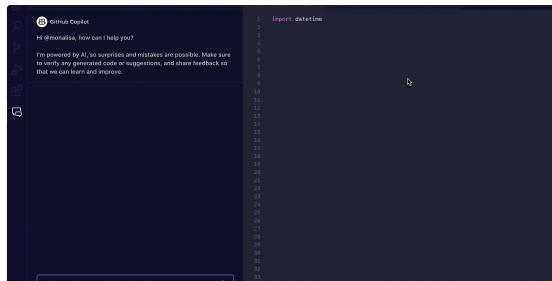
1



...

[See all from Tharindu Vitharana](#)[See all from Sysco LABS Sri Lanka](#)

Recommended from Medium



Jacob Bennett in Level Up Coding

The 5 paid subscriptions I actually use in 2024 as a software engineer

Tools I use that are cheaper than Netflix

◆ · 5 min read · Jan 4



5.7K



67



...



Joe El khoury

Vision Mamba

Mamba is a new design for large language models (LLMs) that deals with long...

5 min read · Jan 23



169



2



...

Lists



Generative AI Recommended Reading

52 stories · 670 saves



What is ChatGPT?

9 stories · 283 saves



The New Chatbots: ChatGPT, Bard, and Beyond

12 stories · 289 saves



Natural Language Processing

1147 stories · 617 saves



The Pareto Investor

ChatGPT has Just Been Dethroned by French Geniuses!

These Three Individuals, a Former Researcher at DeepMind and Two Others...

◆ · 5 min read · Jan 20

4K

37



The Tenyks Blogger

The Foundation Models Reshaping Computer Vision

Learn about the Foundation Models—for object classification, object detection, and...

10 min read · Oct 26, 2023

329

4

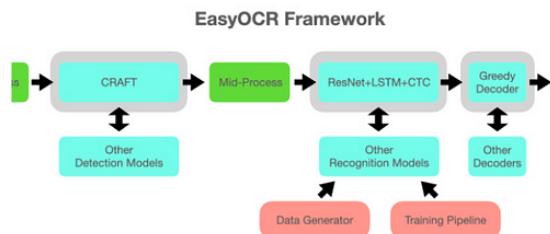


JeremyK in AI monks.io

Image similarity with DINOv2 and FAISS

Explore the robust performance of DINOv2 in image similarity tasks and gain valuable...

9 min read · Sep 27, 2023



Aditya Mahajan

EasyOCR: A Comprehensive Guide

Detailed explanation of EasyOCR with usage examples

11 min read · Oct 28, 2023

285

4



...

26



...

[See more recommendations](#)