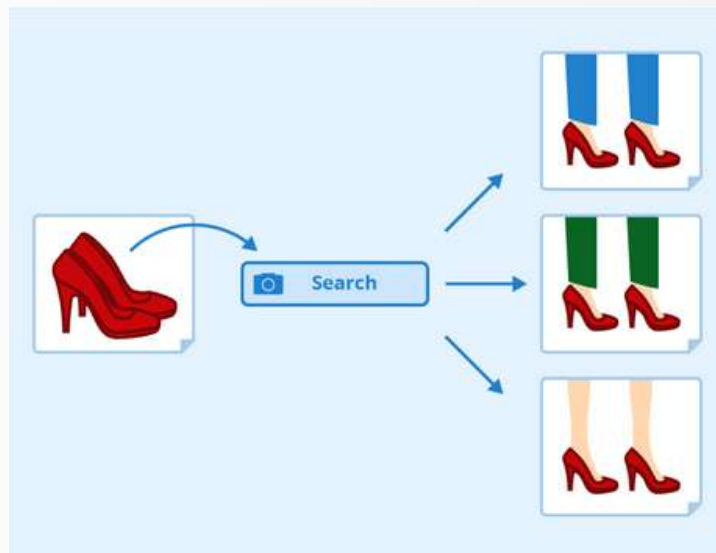


# **AEM VISUAL & SEMANTIC SEARCH**

**Extending Adobe Experience Manager with  
the power of search by similar images**

## EXECUTIVE SUMMARY

In the ever-evolving digital landscape, the ability to efficiently search and manage vast collections of images and media assets is becoming increasingly crucial for businesses. Traditional text-based search methods often fall short in providing accurate and intuitive results, leading to frustration and inefficiencies. Our cutting-edge visual search components offer a transformative solution that allows users to search, discover, and interact with digital assets more effectively, using both image and text-based queries.



Our Visual Search solution is specifically designed for web users and comprises two primary components: the **Image Gallery** and the **Search Field**. The Image Gallery enables users to browse and explore a vast collection of images and discover similar images with a single click. The Search Field provides a powerful tool for web users to find relevant images using either text queries or by uploading or capturing an image. Leveraging advanced AI models for semantic search, our solution ensures accurate and relevant results, enhancing user experience and operational efficiency.

## INTRODUCTION TO VISUAL SEARCH TECHNOLOGY

### The Need for Visual Search

As the volume of digital content continues to grow exponentially, the challenges of managing and retrieving specific assets have become more pronounced. Traditional keyword-based search methods are often limited by the accuracy of metadata and the specificity of user queries. In many cases, users struggle to find the images they need, even if they exist within the systems. This inefficiency can lead to wasted time, lost opportunities, and decreased productivity.

## Benefits of Visual Search

Visual search technology addresses these challenges by allowing users to search using images themselves or by leveraging the power of semantic understanding in text queries. By doing so, it offers several key benefits:

- **Improved Accuracy:** Visual search can identify and retrieve images based on visual similarity, rather than relying solely on text-based metadata.
- **Enhanced User Experience:** Users can find what they need faster and easier, whether they're searching with text or images.
- **Greater Engagement:** Visual search is intuitive and aligns with how users naturally browse and discover content.
- **Operational Efficiency:** Reduces the time and effort required to manage and retrieve digital content, leading to cost savings and increased productivity.

## OVERVIEW OF VISUAL SEARCH COMPONENTS

### 1. Image Gallery Component

Our Image Gallery component is designed to provide an interactive and dynamic browsing experience. Users can explore a curated list of images and, with a simple click, discover visually similar images.



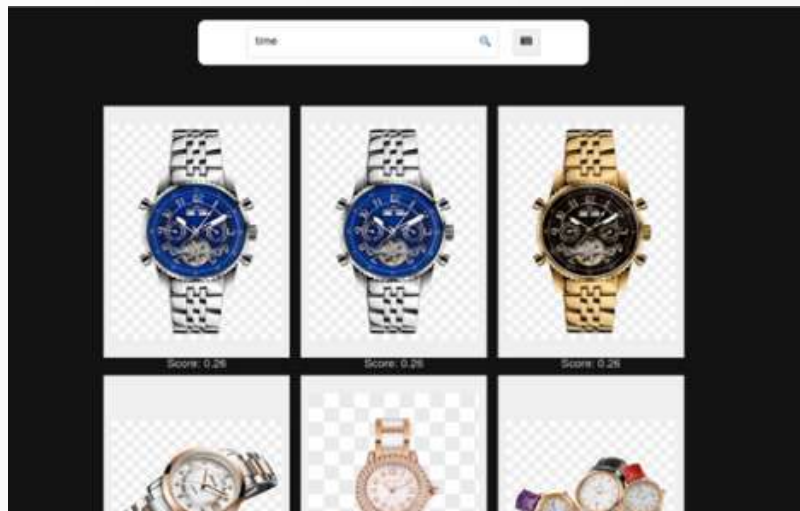
#### Key Features:

- **Visual Similarity Search:** When a user clicks on an image in the gallery, the system automatically retrieves and displays images that are visually similar. This is achieved using advanced image embedding techniques that compare the visual content of images rather than relying on text-based tags.
- **User-Friendly Interface:** The gallery is designed to be intuitive and easy to navigate, ensuring that users can quickly find the images they need.

## 2. Search Field Component

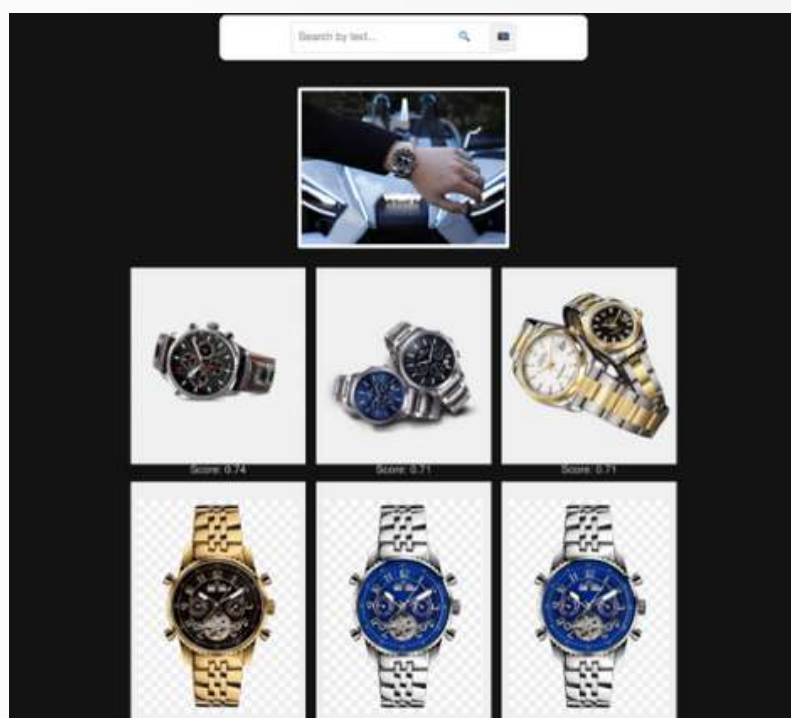
The Search Field component offers versatile search experience by supporting both text-based and image-based queries. Users can enter a text query to search for related images or upload an image (or take a live picture) to find visually similar images.

- **Text Search:**



Our text search functionality leverages state-of-the-art semantic search techniques using CLIP model . This allows users to find images based on the conceptual meaning of their queries rather than just keyword matching. For example, searching for "sunset on the beach" will return images that visually represent this idea, even if those exact words aren't in the metadata. The model's extensive training ensures robust performance and adaptability for various use cases.

- **Image Search:**



- **Upload or Capture:** Users can either upload an existing image or take a live picture using their device's camera. The system then analyzes the visual content of the image and retrieves similar images from the image collections.
- **Efficient Image Processing:** Our solution ensures that images, whether uploaded or captured, are processed efficiently to handle large file sizes without compromising performance.

## CASE STUDIES AND USE CASES

### Real-World Applications

The versatility of our visual search components makes them ideal for a wide range of industries and use cases:

- **E-commerce:** Retailers can use the image search feature to help customers find products visually similar to what they are interested in, enhancing the shopping experience and increasing sales.
- **Content Management:** Marketing teams can easily find and reuse images from their digital library. This helps maintain consistent branding and speeds up the creation of marketing campaigns.
- **Product identification and support:** everyone can take a picture of a product / item and upload in a platform able to identify it and provide technical information.

### Potential impact

Adopting our visual search technology can significantly enhance the efficiency and effectiveness of digital assets. By enabling more intuitive and accurate search capabilities, organizations can improve user satisfaction, increase productivity, and gain a competitive edge in their respective markets.

## TECHNICAL OVERVIEW

### Architecture

Our visual search solution is built on a robust and scalable architecture that integrates seamlessly with AEM. Key components include:

- **Image/Text Embedding Model:** We use CLIP (Contrastive Language-Image Pretraining), an advanced model developed by OpenAI, to generate embeddings for both images and text queries. CLIP excels at mapping images and text into a shared vector space, thanks to its training on millions of image-text pairs. This model enables highly accurate semantic searches by understanding the conceptual meaning of queries and visual content. Available as open-source on platforms like Hugging Face, CLIP enhances our solution's ability to compare and retrieve similar images based on both visual and semantic relevance.

- **Web Service:** The backend service is implemented using Flask, handling the processing of search queries and communicating with the image collections to retrieve and display results.
- **PostgreSQL Database:** A PostgreSQL database stores image embeddings, enabling fast and efficient retrieval of similar images during searches.

### Scalability and Performance

Our solution is designed to handle large volumes of data and high search traffic. Key performance optimizations include:

- **Efficient Image Resizing:** Images are resized before processing to ensure that large files do not impact system performance.

### Security Considerations

We prioritize data security in our implementation:

- **Authentication:** Basic HTTP authentication is used to protect access to the web service.
- **Secure Data Handling:** Images and embeddings are stored and processed securely, ensuring that sensitive information is protected.

## CONCLUSION

Our visual search components represent a significant advancement in digital assets management, offering a powerful, intuitive, and efficient way to search and manage visual content. By combining the latest in AI-driven image and text search technologies with a user-friendly interface, we are empowering organizations to unlock the full potential of their digital assets.