

Press Release

Title: JKL.IO Introduces Innovative Automated Image Collage Feature

JKL.IO, the leading digital news aggregation platform, is proud to announce the launch of its groundbreaking Automated Image Collage Feature. This cutting-edge technology enhances the visual appeal of news articles by creating composite images from high-quality, Creative Commons-licensed sources.

The new feature automatically generates visually striking collages by expertly extracting and combining relevant images from trusted sources like Wikipedia. These composite images provide readers with a richer, more engaging visual context for news stories, improving comprehension and retention of information.

"Our Automated Image Collage Feature represents a significant leap forward in how we present news to our readers," said Jane Smith, CEO of JKL.IO. "By offering visually compelling, factually accurate images alongside our curated articles, we're enhancing the user experience and making complex news stories more accessible and engaging."

Key benefits for JKL.IO users include:

1. Enhanced visual storytelling: Complex news stories come to life through carefully crafted composite images.
2. Improved information retention: Visual aids help readers better understand and remember key points from articles.

3. Trustworthy imagery: All components are sourced from reputable, Creative Commons-licensed repositories.
4. Consistent style: Professional-looking collages maintain the serious tone of our news aggregation service.

The Automated Image Collage Feature is now live on the JKL.IO platform, offering users a more immersive and informative news reading experience.

Frequently Asked Questions:

How does the system select appropriate images for each news article?

Using NLP we extract entities, key phrases, and other information from the article to figure out what to search our repositories for as basic building blocks of the image creation task.

What image processing techniques are used to extract entities and remove backgrounds?

We have OpenCV and other libraries to do this with models for face detection and facial vector extraction and clustering.

Similarly there are background removal libraries.

Finally we have techniques for estimating image quality.

A very basic result from some of the current code:



How does the system ensure that the final composite image is visually appealing and coherent?

That is a key reason we should be building and prototyping straight away instead of over-engineering and seeking to deliver “at the end”. This is because aesthetics is subjective and seeing initial compositions immediately will inform the design of compositional rules.

What measures are in place to respect copyright and ensure proper attribution for used images?

We need to make sure we store the API response and only use images that are public domain or creative commons licences that allow commercial use.

How does the system handle articles covering multiple topics or entities?

We should support 1-3 main PERSONS and other types of entity at first.

What is the average processing time for creating a composite image?

It is not in the UI/UX loop where a user clicks a button and we need to do it quickly, so we have some time to do it. We should try to only process entities once though, so “Donald Trump” is done once and the images are stored for use later.

How does the system adapt to different image sizes and aspect ratios?

We are only targeting one size 1920x1080 for the moment, the techniques should be adaptable to other sizes so functions should try to operate on heights and widths instead of hardcoding 1920 and 1080.

What fallback mechanisms are in place if suitable images cannot be found or processed?

Textures, generic images for different themes (“US Politics”), and text can be used. Text with entity names especially would be useful as a “baseline” to cover instances where a person or entity is not found.

How does the system ensure that the created collages maintain a serious and professional appearance?

This is where using real photos from reference material is far more serious than generative AI. That’s the point of doing it this way.

What kind of scalability considerations should we keep in mind for this feature?

We can scale up the compute on this, so scaling is more about image delivery. If the student project can deliver images at 1920x1080 in PNG format the other scaling challenges (a CDN, conversion to efficient image formats) can be done elsewhere.

How will the system handle updates or changes to the original source images?

This doesn’t matter as we will cache a copy, and the news cycle is 24 hours or shorter so news will cycle off before the source images are changed.

Later on we will want the capacity to “edit the cache” but it’s far more important to get initial images generated than deal with this problem.

What quality control measures are in place to prevent inappropriate or misleading image combinations?

The feature will only use trusted sources, and will be in beta for some time, so there is some time to see if this is an issue before trying to devote engineering resources to stopping it.

How can we optimise the feature for different devices and screen sizes?

We should target 1920x1080 at first, and let CSS on the front-end handle scaling and the conversion to mobile views. This is outside the scope of the student work.

What kind of user controls or customization options should we consider implementing?

This would be later on after the feature is built. E.g. the capacity to have black and white images rather than colour ones.

How will we measure the success and impact of this feature on user engagement?

Users have requested it, we know it’s a high value feature. We can show it to a sample of users and get feedback.

The front-end will have the capacity to turn it off and on. We can track the percentage of users who turn it on potentially, or the percentage of users

who have it on for some time and then turn it off. However the system has minimal logging as user privacy is another feature and we try to store everything in the user cookie and not track everything.

What are some more examples of the types of images we are pursuing?

These kinds of thumbnails from <https://www.youtube.com/@abcnewsaustralia> are good examples of the styles of composition.

The screenshot displays the ABC News Australia YouTube channel interface. It features a grid of video thumbnails with titles, view counts, and upload times. The thumbnails include news reports, market updates, and breaking news. The layout is organized into sections: a top row of six videos, a 'Videos' section with a 'Play all' button, and an 'ABC News updates' section with a 'Play all' button and a description: 'Take a look at the top stories making headlines across Australia and around the world.'

Top Row Videos:

- Travel through US airports to be simplified for Australian...** (2:45, 6.1K views • 7 hours ago)
- US troops injured in suspected rocket attack on...** (3:49, 23K views • 7 hours ago)
- Wall St plunges for second day in row amid recession...** (3:57, 24K views • 10 hours ago)
- Kamala Harris officially secures enough votes to...** (3:48, 8K views • 3 days ago)
- Three September 11 suspects agree to plead...** (8:16, 7.1K views • 3 days ago)
- Emotional welcome on US soil in prisoner swap betwe...** (2:53, 3.1K views • 3 days ago)

Videos Play all

ABC News updates Play all
Take a look at the top stories making headlines across Australia and around the world.

ABC News updates Videos:

- Middle East travel warning amid concerns of serious...** (8:15, 7.9K views • 5 days ago)
- Sharemarket sell-off continues + RBA decision o...** (12:14, 22K views • 7 hours ago)
- Legionnaires' outbreak claims second life; NSW...** (7:29, 8.2K views • 1 day ago)
- 24 released in Russia-US prisoner swap & Hezbollah...** (9:19, 4.9K views • 4 days ago)
- Flights between major cities cancelled as Rex Airlines...** (10:39, 6.6K views • 6 days ago)
- Legionnaires' disease outbreak in Melbourne +...** (5:48, 22K views • 7 days ago)

Many examples cut for file size limit in this version of the document.

Where are we sourcing free or copyleft images?

Wikipedia and importantly OpenVerse which encompasses Wikimedia and other sources with a good API.

Can we formulate it into a series of named styles?

1. Centred Full Image

- Name: "Full Frame"
- Description: A single image, centred and framed appropriately using super-resolution and entropy-based cropping techniques. Includes an option for black and white conversion.

2. Half Picture with Text

- Name: "Split Frame: Image-Text"
- Description: Left half contains a cropped image (either in colour or a "genre" picture), while the right half displays text. Various text options can be explored.

3. Gradient with Person

- Name: "Gradient Portrait"
- Description: Features a person (entity) prominently displayed in front of a gradient background. Includes text overlay.

4. Half Picture with Person and Gradient

- Name: "Split Frame: Scene-Portrait"
- Description: Left side shows a background scene with text overlay, right side displays a gradient background with a person (entity) in focus.

5. Two-Person Juxtaposition

- Name: "Dual Portrait: Simple"

- Description: Two images placed side by side, each occupying half the frame. No background extraction, just two pictures juxtaposed.

6. Two-Person Gradient

- Name: "Dual Portrait: Gradient"
- Description: Similar to the simple dual portrait, but with gradient backgrounds added to both halves.

7. Diagonal Two-Person

- Name: "Dual Portrait: Diagonal"
- Description: Two images cropped and placed diagonally, with a border separating them. This layout is more dynamic and visually interesting.

8. Tripartite Layout (New)

- Name: "Triple Frame"
- Description: Divides the image into three vertical sections, allowing for three different entities or scenes to be displayed side by side.

9. Circular Cutout (New)

- Name: "Circular Focus"
- Description: Main subject is presented in a circular cutout, with a relevant background or texture filling the rest of the frame.

10. Quad Layout (New)

- Name: "Quad Frame"
- Description: Divides the image into four equal quadrants, each containing a different entity, scene, or text element.

Can you expand on the type of shared and specific processing that might be required for each style?

Shared Functions:

- **Image Acquisition:** Fetch and cache images from approved sources.
- **Entity Recognition:** Identify and extract key entities from the article.
- **Image Quality Assessment:** Evaluate and rank images based on quality metrics.
- **Background Removal:** Extract subjects from their original backgrounds when needed.
- **Colour Manipulation:** Adjust colour schemes, including conversion to black and white.
- **Text Overlay:** Generate and position text elements on the image.
- **Composition:** Arrange visual elements according to the chosen layout.
- **Output Generation:** Produce the final composite image at the target resolution.

Specific Functionality for Each Image Type:

1. Full Frame

The Full Frame composition presents a single, impactful image that encapsulates the essence of the news article. This function will first identify the most relevant and highest quality image from our cached repository based on the article's main entities and themes.

The selected image undergoes a series of enhancements to maximize its impact. We employ super-resolution techniques to upscale the image if necessary, ensuring it meets our 1920x1080 target resolution without loss of quality. An entropy-based cropping algorithm then analyzes the image to identify the most information-rich area, which guides our framing decisions.

The function offers a black and white conversion option, which can lend a timeless or serious tone to certain news stories. This conversion isn't a simple desaturation but a carefully calibrated process that enhances contrast and preserves detail.

Finally, we overlay the article headline or a key quote, positioned to complement the image's composition. The text's colour and backdrop are dynamically adjusted to ensure readability against the varying background.

2. Split Frame: Image-Text

This composition divides our canvas into two distinct halves, marrying visual and textual elements. The left half is dedicated to an image, while the right presents text information.

For the image selection, our function first attempts to find a directly relevant photograph. Failing that, it falls back to a "genre" picture that represents the article's broader theme (e.g., a stock image of a courtroom for a legal story). The chosen image is then cropped to fit the left half of our frame, with care taken to preserve the most salient visual information.

The right half is reserved for text. This isn't merely the article's opening paragraph, but a carefully curated extract that complements the image. Our function employs natural language processing to identify key sentences that encapsulate the article's main points. These are then formatted and laid out in a visually appealing manner, with attention paid to typography, line spacing, and margin.

A subtle gradient or texture may be applied to the text background to enhance readability and aesthetic appeal. The function also ensures a

cohesive colour scheme between the image and text portions, possibly drawing colours from the image to inform the text styling.

3. Gradient Portrait

The Gradient Portrait function creates a visually striking composition that places a key figure from the news story in the foreground, set against a dynamically generated gradient background.

First, the function identifies the primary person entity from the article. It then searches our image repository for high-quality photographs of this individual, preferring forward-facing, clear shots. The chosen image undergoes background removal, isolating the figure.

Simultaneously, we generate a gradient background. This isn't a random colour selection, but a process informed by several factors: the tone of the article (cool colours for serious news, warmer for positive stories), colours associated with the subject matter (e.g., greens for environmental news), and colours that complement the subject's clothing or features.

The isolated figure is then placed onto this gradient background. The function carefully positions the figure, typically off-centre to create visual interest, and scales it to achieve the right balance with the background.

Lastly, we overlay text - usually the person's name and a short, impactful quote or headline. The text colour and positioning are calculated to ensure maximum readability while maintaining the composition's aesthetic balance.

4. Split Frame: Scene-Portrait

This composition creates a narrative juxtaposition between a scene and a key figure, divided across the frame. The function begins by identifying two crucial elements: a relevant background scene and a key person entity.

For the left side, we search for an image that captures the context of the story - perhaps a location, an event, or a symbolic representation of the issue at hand. This image is carefully cropped to fit the left half of the frame, with attention paid to retaining the most informative and visually appealing portion of the scene.

The right side focuses on the key person. Similar to the Gradient Portrait, we isolate the figure from its original background. However, instead of a full gradient, we create a subtle, blurred gradient that transitions from the scene on the left to a complementary color on the right. This creates a sense of continuity across the image while still distinctly separating the two elements.

The person is then placed on this gradient background, typically facing inward towards the scene. This creates a visual connection between the individual and the context.

Text elements are incorporated on both sides. On the scene side, we overlay a brief description or headline that sets the stage. On the portrait side, we include the person's name and possibly a relevant quote. The function carefully balances the text placement and styling to ensure readability without obscuring crucial visual elements.

5. Dual Portrait: Simple

The Dual Portrait: Simple function creates a straightforward yet powerful visual comparison between two key figures in a news story. This is

particularly useful for articles discussing conflicts, competitions, or collaborations between two individuals.

The function begins by identifying the two most prominent person entities in the article. It then searches for high-quality images of each, prioritizing photos where the individuals are facing each other if possible, to create a sense of interaction.

Each selected image is carefully cropped to occupy exactly half of the 1920x1080 frame. The cropping algorithm focuses on centering each person's face and including enough of their upper body to provide context without unnecessary background.

A key challenge here is ensuring visual balance. The function analyzes both images for overall brightness and color tone, and may apply subtle adjustments to create a harmonious pair. However, it's careful not to alter the fundamental character of the original photographs.

A thin border is added between the two halves to clearly delineate them. The color of this border is chosen to complement both images without drawing undue attention.

Finally, text overlays are added to identify each individual. The function places these labels in a way that doesn't obscure important facial features, typically at the bottom of each half.

6. Dual Portrait: Gradient

Building on the Simple Dual Portrait, this function adds an extra layer of sophistication with custom gradient backgrounds. The core process of

selecting and cropping two portraits remains the same, but the treatment of these images is more complex.

After selecting the portraits, the function employs advanced background removal techniques to isolate each figure. It then generates two distinct gradient backgrounds, one for each half of the image. These gradients are not random, but carefully crafted based on several factors:

1. The colour palette of the original images, particularly the subjects' clothing.
2. The tone and topic of the article (e.g., cooler tones for a sombre story, warmer for a positive one).
3. Colours associated with each person (e.g., party colours for politicians).

The isolated figures are then placed onto their respective gradient backgrounds. The function adjusts the positioning to create a balanced composition, typically placing each figure slightly off-centre, facing inward.

To enhance the sense of connection between the two halves, the function creates a subtle blend at the centre line where the two gradients meet. This helps to unify the composition while still maintaining the distinct space for each figure.

Text elements are added last. In addition to identifying labels for each person, this version of the function may include a brief quote or key point associated with each individual. The text is styled to be easily readable against the gradient backgrounds, with dynamic colour selection to ensure contrast.

7. Dual Portrait: Diagonal

The Diagonal Dual Portrait function creates a more dynamic and visually interesting composition by introducing a diagonal split between two figures. This layout can be particularly effective for stories that pit two sides against each other or showcase a strong contrast.

The function begins similarly to other dual portrait types, by selecting high-quality images of the two main person entities. However, the cropping process is more complex. Instead of a straight vertical division, the function calculates a diagonal line across the 1920x1080 frame. This line isn't always a perfect 45-degree angle; the function may adjust it slightly based on the composition of the selected images.

Each portrait is then carefully cropped to fit its diagonal section. This requires more sophisticated image analysis, as the function must ensure that key features (particularly faces) aren't awkwardly cut off by the diagonal line. It may need to zoom, pan, or even select alternative images to achieve the best fit.

A key feature of this composition is the border between the two halves. Unlike the subtle dividers in other layouts, this border is more pronounced. The function generates a bold line or a small gradient band that clearly separates the two sides. The colour and style of this divider are chosen to complement both images while adding to the overall aesthetic.

The diagonal layout presents unique challenges for text placement. The function carefully positions names or titles for each figure in a way that follows the diagonal flow of the image. It may use angled text or strategically placed text boxes that align with the diagonal split.

To further enhance the composition, the function may apply subtle post-processing effects. This could include adding a gentle vignette to each

half to draw focus to the central figures, or applying a unified colour grading across both halves to tie the composition together.

8. Triple Frame

The Triple Frame function creates a composition divided into three vertical sections, allowing for a more complex visual narrative. This layout is particularly useful for stories with three key elements, whether they're people, places, or concepts.

The function begins by identifying the three most significant elements from the article. These could be three people, or a mix of people and relevant imagery. It then searches for appropriate high-quality images for each element.

Each selected image is cropped to fit precisely one-third of the 1920x1080 frame. The cropping algorithm is sophisticated, ensuring that the most relevant part of each image is centred in its section. For portraits, this means focusing on the face and upper body. For location or concept images, it means capturing the most informative or iconic elements.

A key challenge of this layout is creating visual harmony across three potentially disparate images. The function employs colour analysis and adjustment techniques to ensure a cohesive look. This might involve subtly adjusting the colour temperature or contrast of individual sections to create a balanced overall appearance.

Thin borders are added between the sections to clearly delineate them. The colour of these borders is carefully chosen to complement all three images without drawing undue attention.

Text overlays are crucial in this layout to clarify the significance of each section. The function places descriptive labels or titles at the top or bottom of each third. In cases where the three elements represent a progression or process, it may also add small numbers or arrows to indicate the sequence.

Finally, the function may add a unifying element across the bottom of the entire image, such as a headline or summary statement. This helps tie the three sections together into a coherent whole, reinforcing the overall narrative of the article.

9. Circular Focus

The Circular Focus function creates a composition that draws immediate attention to a central element, surrounded by contextual imagery. This layout is effective for stories that have a clear main focus but require additional visual context.

The function starts by identifying the primary subject of the article - typically a person, but it could also be an object or symbol. It then searches for a high-quality image of this subject and uses advanced image processing techniques to isolate it from its background.

Next, the function creates a circular mask in the centre of the 1920x1080 frame. The size of this circle is calculated to be large enough to prominently feature the subject, but not so large as to dominate the entire composition. The isolated main subject is then placed within this circular area.

For the surrounding area, the function has several options:

1. It can search for a relevant background image that provides context to the main subject. This could be a location, an event scene, or a thematic image.
2. It can generate a texture or pattern that relates to the article's topic.
3. It can create a complex gradient or abstract design that complements the central image.

Whichever option is chosen, the function applies a blur effect to this background. This blur is strongest near the edges of the circular cutout, creating a subtle vignette effect that further draws the eye to the central subject.

The transition between the circular focus area and the surrounding background is crucial. The function creates a soft feathered edge, possibly with a subtle glow effect, to seamlessly blend the two elements.

Text elements are then added to the composition. The subject's name or a short title is typically placed below the circular area. Additional text, such as a quote or brief description, can be overlaid on the blurred background area, carefully positioned to be readable without detracting from the central focus.

10. Quad Frame

The Quad Frame function divides the image into four equal quadrants, allowing for a multi-faceted visual representation of a complex story. This layout is particularly useful for articles that deal with multiple related topics or require the presentation of several key figures or concepts.

The function begins by identifying four key elements from the article. These could be any combination of people, places, objects, or symbolic

representations of concepts. It then searches for appropriate high-quality images for each element.

Each selected image is cropped to fit precisely one-quarter of the 1920x1080 frame. The cropping algorithm is designed to focus on the most important aspect of each image, whether that's a person's face, a landmark in a landscape, or a specific part of an object.

One of the main challenges of this layout is ensuring visual coherence across four potentially diverse images. The function employs sophisticated colour analysis and adjustment techniques to create a harmonious overall appearance. This might involve subtly altering the colour balance or applying a unified colour grade across all four quadrants.

Thin borders are added between the quadrants to clearly delineate them. The colour and thickness of these borders are carefully chosen to provide structure without overwhelming the imagery.

The function then adds text overlays to each quadrant. These could be simple labels identifying each element, or more detailed captions providing context. The text positioning and styling are dynamically adjusted to ensure readability against each specific background.

To tie the four quadrants together conceptually, the function may add a central element that overlaps all four sections. This could be a circular logo, a key headline, or a symbolic image that encapsulates the overall theme of the article.

Finally, the function considers the overall balance of the composition. It may slightly adjust the sizing or positioning of elements within each

quadrant to create a visually pleasing and balanced final image that effectively communicates the multi-faceted nature of the story.

What might a program of work look like for that?

Here's an agile-oriented hierarchical statement of work with outputs and milestones for a group of 7 students working on this project for 7 weeks at 10 hours per week. This plan emphasises producing and demonstrating results regularly.

Week 1: Foundation and Basic Image Processing

Goal: Set up the project and implement basic image processing functions.

Tasks:

1. Set up project repository and development environment
2. Implement image acquisition and caching from approved sources
3. Develop basic image quality assessment function
4. Create a simple background removal function
5. Implement basic color manipulation (including B&W conversion)

Output: Demo of acquired images and basic image manipulations

Milestone: Project setup complete, basic image processing pipeline functional

Week 2: Entity Recognition and Text Overlay

Goal: Implement entity recognition and text overlay capabilities

Tasks:

1. Develop entity recognition function to extract key entities from articles
2. Create text generation function for overlays
3. Implement text positioning and styling function
4. Integrate text overlay with existing image processing pipeline
5. Begin work on the "Full Frame" composition type

Output: Demo of entity recognition and text overlay on single images

Milestone: Entity recognition and text overlay functional

Week 3: Full Frame and Split Frame: Image-Text

Goal: Complete "Full Frame" and implement "Split Frame: Image-Text"

Tasks:

1. Finalise "Full Frame" composition type
2. Implement advanced cropping algorithm for "Split Frame: Image-Text"
3. Develop text extraction and formatting for the text portion
4. Create cohesive colour scheme function for image and text portions
5. Integrate all components for "Split Frame: Image-Text"

Output: Demo of completed "Full Frame" and "Split Frame: Image-Text" compositions

Milestone: Two composition types fully functional

Week 4: Gradient Portraits and Split Frame: Scene-Portrait

Goal: Implement "Gradient Portrait" and "Split Frame: Scene-Portrait"

Tasks:

1. Develop gradient generation function
2. Implement figure isolation and placement for "Gradient Portrait"
3. Create scene selection and cropping for "Split Frame: Scene-Portrait"
4. Develop blended gradient background for portrait in split frame
5. Integrate components for both composition types

Output: Demo of "Gradient Portrait" and "Split Frame: Scene-Portrait" compositions

Milestone: Four composition types fully functional

Week 5: Dual Portraits (Simple, Gradient, Diagonal)

Goal: Implement all three Dual Portrait variations

Tasks:

1. Develop face detection and optimal cropping for dual portraits
2. Implement "Dual Portrait: Simple" with balancing adjustments
3. Create gradient background generation for "Dual Portrait: Gradient"
4. Develop diagonal splitting and cropping for "Dual Portrait: Diagonal"
5. Implement dynamic text placement for diagonal layout

Output: Demo of all three Dual Portrait variations

Milestone: Seven composition types fully functional

Week 6: Triple Frame and Circular Focus

Goal: Implement "Triple Frame" and "Circular Focus" compositions

Tasks:

1. Develop three-part image selection and cropping for "Triple Frame"
2. Implement colour harmonisation across three images
3. Create circular masking and feathering for "Circular Focus"
4. Develop background blur and vignette effect for "Circular Focus"
5. Implement text placement for both composition types

Output: Demo of "Triple Frame" and "Circular Focus" compositions

Milestone: Nine composition types fully functional

Week 7: Quad Frame and Integration

Goal: Implement "Quad Frame" and integrate all components into a cohesive system

Tasks:

1. Develop four-part image selection and cropping for "Quad Frame"
2. Implement colour harmonisation across four images
3. Create central overlapping element for "Quad Frame"

4. Develop main interface for selecting composition types based on article content
5. Integrate all components into a unified system
6. Conduct thorough testing and bug fixing

Output: Final demo of all ten composition types and the integrated system

Milestone: Complete functional prototype of the Automated Image Collage Feature

Throughout each week:

- Daily stand-ups to discuss progress and obstacles
- End-of-week retrospective to adjust the plan as needed
- Continuous integration and testing of new components
- Documentation of code and processes

This plan allows for regular demonstrations of progress, with each week building upon the previous work. It's designed to be flexible, allowing for adjustments based on the team's progress and any challenges encountered. The team should be prepared to reprioritize tasks or shift focus as needed to ensure the most critical functionalities are completed by the end of the 7-week period.