SHIFT

SHIFT TECHNOLOGY **INSURANCE PERSPECTIVES**

THE FRAUDULENT IMAGE ANALYSIS EDITION

www.shift-technology.com/en-gb

From the editor

In the fight against insurance fraud, generative AI (Gen AI) has emerged as somewhat of a double-edged sword. It has proven to be a powerful tool that can help insurers accurately extract and classify information from many types of insurance documents and analyse the results to identify anomalies and inconsistencies that may indicate fraud. Gen AI can present human investigators with document summaries, best next steps, and automated reporting and audit trails. Gen AI can automate many aspects of fraud detection, identify both complex fraud rings and opportunistic bad actors targeting an insurer, and accurately analyse both structured and unstructured data—including text, images, and audio.

But the ubiquity of Gen AI tools has also made it significantly easier for those with ill intent to attempt insurance fraud. With a simple prompt made to freely available online resources, a fraudster can create legitimate-looking accident reports, estimates and invoices, supporting documentation, and even images.

The industry finds itself in a classic "fighting fire with fire" scenario. So, how can insurers use Gen AI to actually stay ahead of those that wish them economic harm? In this edition of *Shift Insurance Perspectives* we take a closer look at how image analysis is already being used to help insurers spot the deepfakes and avoid paying out for illegitimate claims.

The current situation

Is it fraud, or not? For insurers facing a suspicious claim, that decision can only be as certain as the veracity of the sources used to make it. Unfortunately for the industry, producing fake “evidence” for insurance claims is becoming not only scarily photorealistic but ever more simple to produce for the average person. We have seen dents, dings, scrapes and scratches added to images submitted with auto claims. Fraudsters manipulate photos to show holes in walls, and smoke or water damage to justify property claims. There are cases where items such as televisions, stereo equipment, and other electronics are inserted into photos of living spaces to bolster theft claims—often submitted alongside fabricated police reports. We have reached a point where human claim handlers and investigators can no longer trust their own eyes when confronted by these AI-generated images.

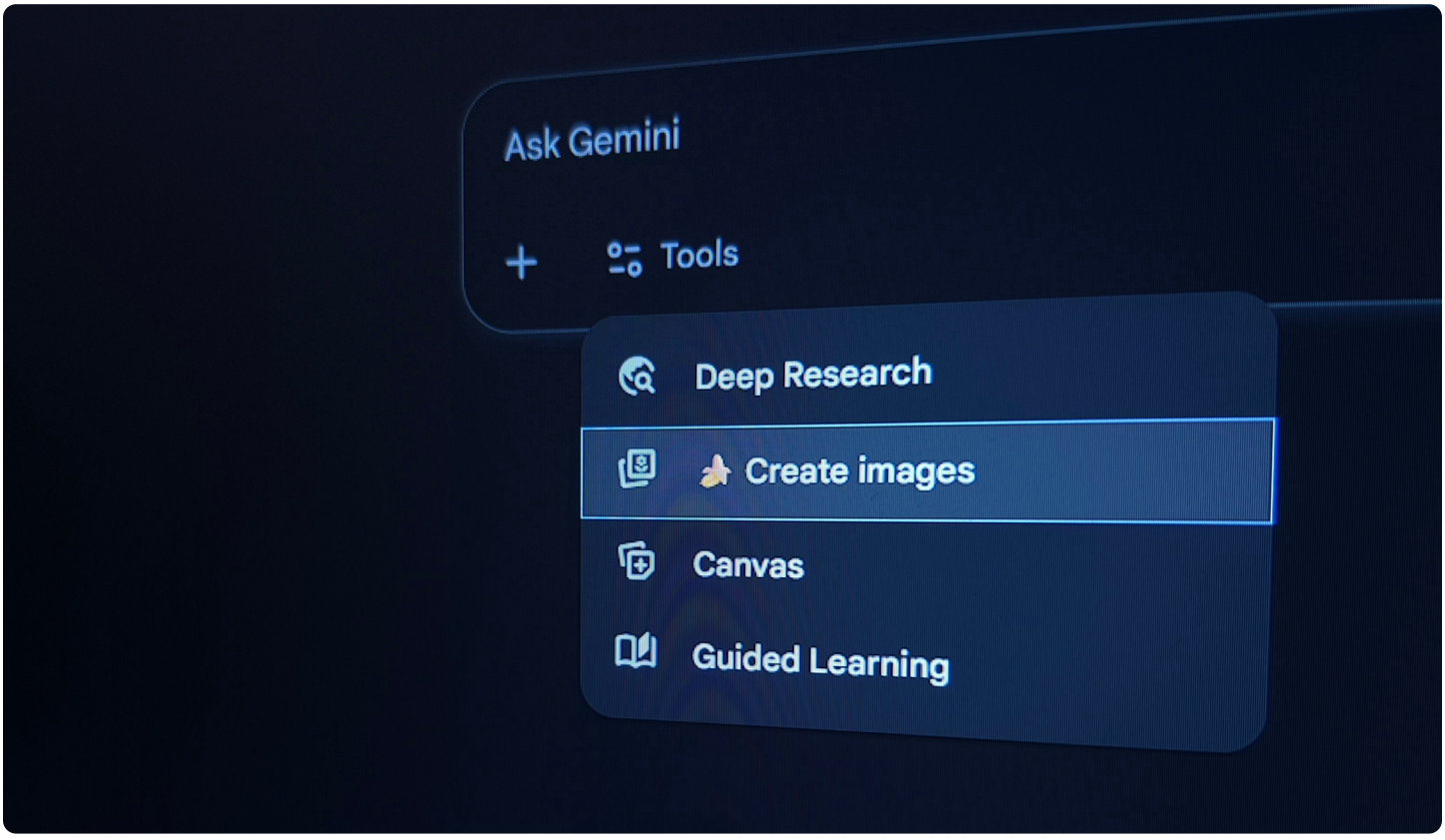
At the same time, we no longer have to rely solely on our human senses to make these critical determinations. What AI has helped create, it can also help detect. How should insurers be thinking about the use of AI to analyse images as part of the claims process?

Analysing the underlying metadata

The first step to understanding whether or not an image is legitimate is analysing its associated metadata. Does the timestamp align the timing of the reported incident? Does geolocation indicate the photo was taken where the loss occurred or in some far-flung destination? Does associated device information show that the image was produced on a camera/phone/tablet reported as part of the claim, or highlight other abnormalities? Does the metadata reveal that the image was downloaded from the internet or otherwise manipulated? Any of these scenarios provide a good indication that a manipulated image is being used to support a fraudulent claim.

When you cannot rely on metadata

Unfortunately, the reality of image analysis is that metadata can be manipulated almost as easily as the image itself. Images can be converted into different formats than the original, obfuscating the underlying metadata. Images can be cropped or even have their associated metadata removed completely. And while these actions may be indications of fraudulent intentions, they also may not. Policyholders have been known to submit screenshots instead of original photos. We have seen cropped images being used to highlight damage. As with any element of a claim that could raise suspicion, the industry must be diligent in avoiding false positives in relation to image analysis.



Leveraging image analysis in the fight against fraud

There are several ways one can identify illegitimate images being used in an attempt to perpetrate fraud. Error-Level Analysis (ELA) is an established technique for understanding the veracity of photographic and other digital evidence. Although not specifically GenAI-based, ELA can help determine whether specific elements of a photo have been cleverly modified or spliced together. It can also be used to determine if the elements of a digital document are consistent. This approach has been put to good use identifying Deep Fakes, spotting when new elements have been added to photos, and highlighting when information has been "cut and pasted" from one document to another. However, with the advent of fully AI-generated images, such as those being produced by widely available models, ELA has become less effective. Images created by LLMs are a single, cohesive unit. As such, there is no splicing or "cutting and pasting to be discovered.

Now that highly available models, like those offered by OpenAI, Anthropic, and others, are capable of receiving image inputs alongside text, Gen AI tools can be used to more easily weed out clearly fictional creations, such as drawings. And although the ability to identify an AI-created drawing in the context of fraud detection may seem unnecessary, a drawing is a genuine image. In simple claims, especially those identified as suitable for straight through process (STP), without this check in place this type of image could be enough for a fraudulent claim to be successful. However, this level of check would generally operate alongside a more complex image analysis model, and act more like a human admin sorting evidence types.

Image analysis becomes truly interesting now that Transformers (a deep learning architecture) are becoming more widely available. In Shift's own testing, these models are delivering the best, most consistent results when tasked with detecting AI-generated images. Transformers are derived from Gen AI models such as GPT, but can be modified into smaller, specialised versions designed to pay attention to other types of input—such as images. By training Vision Transformers on "real" and "fake" images, some really impressive results have been achieved.

Image analysis becomes truly interesting now that Transformers (a deep learning architecture) are becoming more widely available.

Conclusion

When modern smartphones automatically use some form of AI to enhance/upscale photos by default, one could argue that all photo evidence submitted to justify an insurance claim could be considered AI-generated. It becomes even more important than ever to accurately discern between images that illustrate the truth from those that support a lie. Despite Gen AI giving bad actors new tools in their attempts to defraud insurers, it is also giving insurers the weapons needed to fight back.

SHIFT

About Shift Technology

Shift Technology is the leading AI platform for insurance. Shift combines generative, agentic, and predictive AI to transform underwriting, claims, and fraud & risk—driving operational efficiency, exceptional customer experiences, and measurable business impact. Trusted by the world's leading insurers, Shift delivers AI when and where it matters most, at scale and with proven results.

Learn more at www.shift-technology.com/en-gb.