Snap, Crackle and . . . *Slop?*
Imbuing Generative AI with the Essence of Computational Creativity

Tony Veale
School of Computer Science,
University College Dublin,
Dublin D4, Ireland.
tony.veale@UCD.ie

Abstract

Recent remarkable advances in generative AI (GenAI) have not sated consumer demand for new tools or new features, nor have they quelled our suspicions that machines are simultaneously capable of great virtuosity and tremendous stupidity in the same act of synthetic creation. The World Wide Web is now awash in seemingly human-quality texts, images and videos, but small mistakes that were once considered quirky and charming in the context of computational creativity (CC) are now seen as reasons to dismiss otherwise proficient outputs as "slop." Why the same disdain was not shown to the products of CC systems which rarely reach, in objective terms, the same levels of versatility and proficiency as those of GenAI systems, is the central question of this paper. It is a question that goes to the heart of what makes a creator and their creations truly "creative" rather than merely generative. The answer, as we present it here, offers hope for CC as a continuing research endeavour with a distinctive character of its own.

Introduction

If familiarity breeds contempt, as the old saying goes, this is certainly true of *Generative AI* (GenAI), a family of neural models that includes Large Language Models (Matarazzo and Torlone 2025) and diffusion-based image or video generators (Shen et al. 2025). As GenAI's outputs become ubiquitous, an initial sense of wonder has given way to a growing sense of dismay. Our willingness to criticize those outputs as soulless or inept has more than kept pace with new developments in the field, even if our criticisms are not always fair or aimed at the right targets. Rather, with greater fluency and versatility come higher expectations, so that many GenAI artifacts are dismissed as "slop" (Kommers et al. 2025), even when they exhibit a level of virtuosity that might take one many hours to achieve with more conventional digital tools.

Those outputs are sometimes flawed in tell-tale ways: a six-fingered hand here, a third arm there, or objects that unintentionally flirt with surrealism, such as a typewriter with keys on its front *and* its back. But these errors, uncanny as they are, seem little reason in themselves to dismiss a complex artifact as "slop." Assessed objectively, the products of past CC systems were often more error-prone and less likely to pass for human-level creativity, yet those failings rarely attracted such scorn. It seems that the label "slop" has less to do with the inherent failings of a machine-made artifact and more to do with the mind-sets of those who produce and consume it (Kommers et al. 2025). While CC systems rarely develop beyond the lab to make a lasting dent on public consciousness, consumers have eagerly embraced GenAI as a panacea for many kinds of generative work. The fear that AI will displace and replace human creators, merely inchoate in our CC systems, has ripened to fruition with GenAI. It has never been easier to automate the creation of human-quality artifacts, and never more tempting to let the machine do all of our thinking for us (Orwell 1946). Whenever we deride an AI artifact as slop, our proximate target is the artifact itself, which may not exhibit any of the tell-tale flaws of automated generation, but our real target is the person who brought it into being, as well as those who provided the tools to do so.

Let's take a moment to unpack the slop metaphor (Øivind Madsen and Puyt. 2026). Some liquids are inherently 'slop', such as the contents of a chamber pot or a pig trough, while others become slop through careless handling, as when pigment is slopped on a wall or canvas by an indifferent painter. It is the latter sense that applies to the outputs of generative AI. The result need not be low quality, just low effort, since this lack of effort lends itself to rapid production, lazy filtering, and over-sharing. Simply, the producer has not invested enough of themselves, of their time and energy, to make the artifact worthy of *our* time and *our* energies. It is this investment that leads us to label the flawed outputs of our CC systems as quirky or uncanny, but to dismiss the competent mediocrity of many GenAI outputs as indiscriminate slop.

As users of GenAI, how should we make this CC-like investment of time and creative energy in what our systems produce? It is this sweat equity that allows us to share in the creativity of the system itself, to be able to say not just that "I prompted for this" but that "I helped build that." All CC systems are meta-creative, and to an extent co-creative too, since they automate our personal view of the creative process in a given domain. The avoidance of slop is not a matter of preventing the generation of low-quality artifacts, as creativity always carries a risk of failure. Interesting mistakes are rarely "slop." Such failures show us the limits of our understanding and allow us to improve our systems. To put more of ourselves as meta- and co-creators into a CC system with GenAI as its generative core, we must build a framework that lets us share in the system's successes and its failures. What is required is a marriage of CC and GenAI that gives us the best of both paradigms in a single system.

To illustrate this philosophy, in more ways than one, we explore here the creation of topical comic strips by pairing a GenAI black box, an LLM, with a CC glass box, a symbolic grammar for visual slapstick that gives shape to the actions of the LLM. This pairing offers a mode of co-development akin to how we humans collaborate creatively on rich multimodal artifacts such as comics, films and games. As we will show, it mirrors a mode of collaboration from the comics industry, the *Marvel Method*, that allows creators with distinct skill-sets to work together toward an emergent whole.

The Marvel Method

The myth of the lone creator, toiling alone in an attic, garage or lab, is stubbornly pervasive in our narratives of creativity (Boden 1992; Veale 2012; Burkus 2013). When it comes to commercial products, however, most are created by teams, not individuals. The inciting spark may well come from a single person, but it takes an ensemble of mindsets and skillsets to develop a multi-faceted result, as exemplified by the multimodal creativity of the comics industry. While this business has produced some famous auteurs, such as Georges Remi (Hergé), the creator of *Tintin*, most comics are created by a partnership of a writer and an artist. The writer takes charge of plotting and dialogue, the artist of layouts, pencil art and inking. In the early 1960s, as editor Stan Lee was launching a minor revolution at Marvel Comics, he found himself writing so many titles that he invented what is now called the *Marvel Method* (Ditko 1973; Lee and Buscema 1978; Lee and Mair 2002; Evanier 2008; Howe 2012) to streamline the collaborative process.

Lee perfected his method with trusted artists such as Jack Kirby and Steve Ditko. Lee and the artist would first spitball the gist of a new story, agreeing its main themes and events, before the artist would go off to illustrate the story, pencilling rough panel layouts into a rich visual narrative. Lee would then return to the story, to reinterpret the artist's visuals in the context of their agreed outline. In adding his own image-specific captions and dialogue, Lee was free to ignore the intentions of the artists, who often complained that they bore the bulk of the creative burden (Ditko 1973; Evanier 2008). However, the method allowed Lee to spread himself wide but not too thin, to shape a wide range of titles that would define Marvel's approach to visual narrative.

The Marvel Method is an approach to human co-creation, but it also provides a basis for co-creation by machines, or by agents operating within different AI paradigms, such as a rules-driven CC system and a statistical GenAI model. In comics production, one might produce the visual narrative, the other the textual narrative that turns those visuals into a finished comic strip. The method gives each agent maximal autonomy to employ their unique skill-sets. It also allows CC developers to invest the *sweat equity* that distinguishes producers of effortful artifacts from the generators of slop. While the GenAI partner, perhaps an LLM driven by clever prompting, remains a black box, a CC glass box can reflect the attitudes, insights and creative efforts of its creator. We might even re-purpose an existing CC system for this role, to provide scrutable insights into a creative process whose traceable inputs drive the inscrutable workings of an LLM.

The Marvel Method was not an uncontroversial mode of co-creativity, but its problems in a human setting become advantages in a wholly automated setting. For instance, credit for *who created or thought of what* is hard to attribute to a specific individual, sparking animosity between the public-facing Lee and the backroom artists who would ultimately part ways with him. In a computational setting, however, we prize any ambiguity that prevents us from pointing to any one sub-system or collaborator as the "true" creator. In fact, Lee's method encourages ambiguity at multiple levels, and even elevates ambiguity into a generative principle. Ambiguity in the rough plot sketch allowed artists to explore their own creative interpretations, while ambiguity in the finished art allowed Lee to impose his own script ideas, to the point of retro-fitting a new meaning onto the artist's renderings. Lee could take inspiration from the art itself, to furnish dialogue that explicitly referenced what was visible on the page.

When automating the creation of comic strips (Kurlander, Skelly, and Salesin 1996; Alves et al. 2007; Pérez y Pérez, Morales, and Rodríguez 2012), a diffusion-based image generator (Yang et al. 2024) may collaborate with a large language model (Patil and Gudivada 2024), the former assuming the role of the artist, the latter that of writer. These models work well together because they are each conditioned on text prompts. However, the roles of writer and artist can also be played by two different LLMs, to produce not a final visual product but a text specification of that product, which can then by visually rendered by a third model. In this paper we explore the automation of the Marvel method for comic strips with two different kinds of AI model: an old-school symbolic model to generate the rough outline of the plot and a specification of how it will be rendered (Veale 2022; 2023), and an LLM to write dialogue and captions to suit this visual rendering (Veale 2024). The key to the Marvel method is the order in which its steps are performed, with the artwork done first, and the writing – which is traditionally the first step – done later, as a response to the art.

The next section provides a more detailed view of this two-step, two-model approach. We will use an XML schema to specify each comic's visual and textual content (Walsh 2012), and focus on short comics that serve a useful purpose in themselves, of a kind that (Veale 2023) has called "serious" comics. As in (Veale 2023), our comics will present a balanced, two-sided perspective on a serious topic, such as immigration, artificial intelligence, vaccines, populism and so on, but the comics themselves will not be "serious." Rather, they will employ a visual slapstick that aims to add levity to an otherwise heavy debate, in which the distancing effect of irony preserves a sense of neutrality. The generation of visual humour is the responsibility of the CC module, and reflects a substantial investment of time and effort, and a personal creative sensibility, in the construction of a visual grammar for slapstick. In contrast, the writing of a script that imposes meaning on these antics in the context of a chosen topic is the responsibility of the LLM. So, in automating the Marvel Method, the LLM takes the role of Lee, while the visual grammar fulfills the responsibilities of Kirby and Ditko.

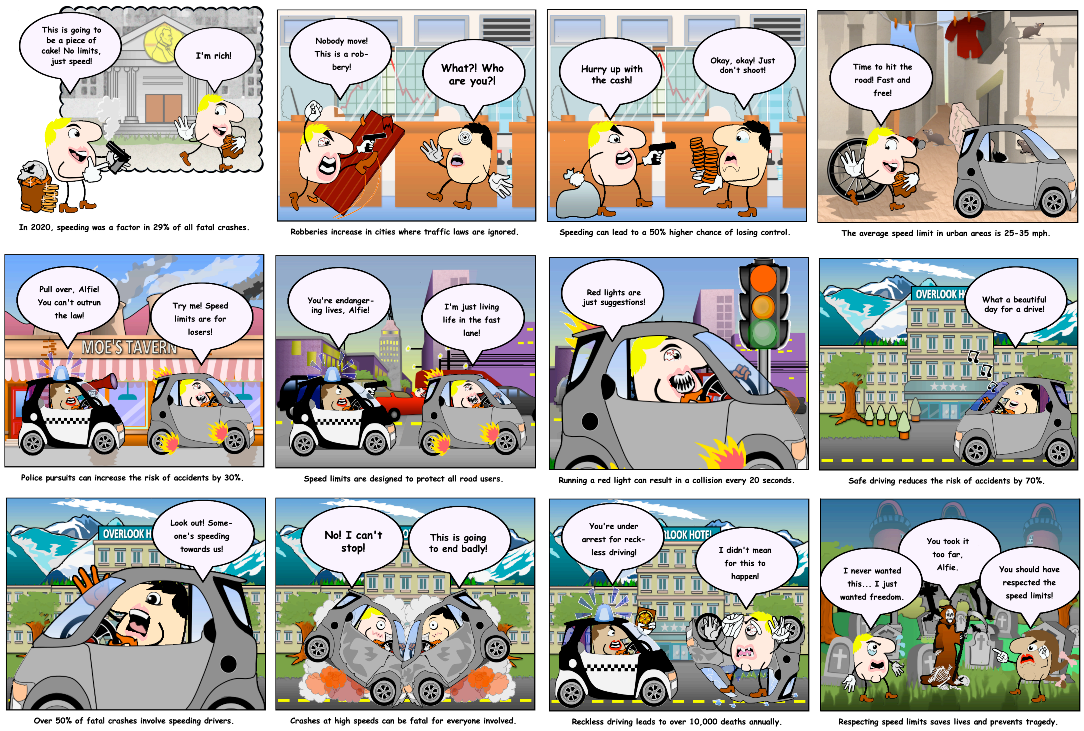

Figure 1: An *Excelsior!* comic on speed limits with LLM dialogue and captions, as rendered from XML.

In Two Minds: Two-sided Comics Generation

Comics are a malleable commodity. Comprising a sequence of panels separated by white spaces known as "gutters", a comic breaks a narrative into distinct beats in which one or more characters act, speak, think, and interact against a visual background (McCloud 1993; 2006; Cohn 2013). When represented in this form, a comic can easily be re-printed, re-sized, re-colored, re-structured, re-edited, and reordered. These actions are made even easier if we distinguish the formal specification of a comic narrative from its eventual rendering as a multimodal object. To this end, we will use an XML schema, *ComiXML*, to represent a machine-generated comic. Each instance of the schema first defines a comic's recurring figures, assigning each a name and visual qualities such as skin tone, hair colour and gender, before specifying each panel in turn, in the sequence they follow in the comic.

Each panel specifies a single narrative beat in which one or more characters are placed against a visual backdrop in a particular panel location (left, right, middle) with a set orientation (facing left or right). Each character may articulate dialogue in a panel, either in a speech balloon (spoken) or a thought bubble (internal), while captions above or below a panel present a narrator's perspective (Forceville, Veale, and Feyaerts 2010). Each character also takes on a pre-defined pose in a panel to express an action (falling, throwing, running, etc.) or an emotion (crying, laughing, accusing, etc.). *Excelsior!* has an inventory of more than 1000 visual poses, and over 500 backgrounds, allowing a rendering engine to map directly from the XML specification of a comic to the visual end-product, which may be shared as a web page, an animated GIF, or another format suited to digital sharing. An example of a fully-rendered comic is shown in Fig. 1.

The task of generating a comic now becomes one of XML generation, or, taking a step back, generating what will be encoded in this XML format. An LLM can be used to generate the narrative content that is so encoded, by first asking the LLM to generate a high-level plot for a story, and subsequently asking the LLM to break this plot into a sequence of panel-sized beats of a single action apiece. For each beat, the LLM is then prompted for suitable dialogue for each participating character; it is also tasked with suggesting apt poses for these characters from its inventory, given both the plot and what the character is saying or doing at this point in the narrative (Veale 2024). Finally, these strands are woven together into a single XML structure, placing each piece of text generated by the LLM (dialogue, captions, poses) into its schematic setting. This straightforward process intuitively models the *write first, draw later* approach to comics.

However, this approach is prone to *hallucination* (Kalai et al. 2023), since LLMs may suggest backgrounds or poses from outside the available inventory and that a renderer cannot render. We have to treat the LLM's suggestions not as literal asset names but as starting points for asset selection. We can, for instance, map each LLM suggestion to the semantically nearest asset, by estimating the semantic distance between this text (e.g., "pointing angrily") and each asset label (e.g. "accusing", "pointing"). Semantic distance can be estimated using the cosine similarity of the vector encodings of the texts to be compared, provided each text can be reduced into a high-dimensional vector embedding that preserves its general meaning (Manning, Raghavan, and Schütze 2008, p. 289). This approach is robust, as state-of-the-art embeddings can be obtained from the same cloud services as the LLMs that generate the texts to be encoded, but the highest-ranked asset by vector similarity is not always the one a human artist would choose. Moreover, the dialogue may now refer to a feeling or action that is no longer visible on the page, which goes some way toward defeating the integrative goal of using comics. To fully realize this goal, we will need to think differently, and flip the script on comics creation.

Flipping the Script with the Marvel Method

Vector encodings map the gist of a line of dialogue to a visual asset, but these mappings are coarse and ignore the peculiarities of a text or an asset. Consider the *Excelsior!* pose for *sympathetic*, in which a caring figure offers a tissue to dry one's eyes. If dialogue is generated that exudes a sense of concern, this pose will rank high as a candidate, but the dialogue will not have been produced to echo the pose's visual details. Conversely, if the pose is chosen before any dialogue is generated, the LLM can weave a sympathetic text that actively refers to the image, as in "Don't cry. Here, take one of these." It helps if the other figure seems to be crying when this line is said, so an LLM should ideally know about all of the poses in a panel before it generates any dialogue at all for the panel. This is the essence of the Marvel method: to draw first, (re-)interpret second, and only then write. To use this method for automated comics generation, we will need a means of generating high-level plots, of rendering those outlines as visual sequences, and of layering a rich text of dialogue and exposition over each sequence. By prompting an LLM with a combination of plot points (what is happening to whom, because of whom, and where) and of verbal action descriptions (what those happenings will look like to an observer), it should be able to generate captions and dialogue that speak directly to the final visualization.

To produce a full story using the Marvel method, we first sketch the outline of a plot. This will serve as a scaffolding around which an LLM can generate exposition and dialogue. The simplest scaffoldings, expressed in ComiXML, specify a sequence of panels in which two recurring characters, named *Alfie* and *Betty*, strike poses against a series of backgrounds. The poses should be apt for their backgrounds and coherent in how they change from one panel to the next. While they do not determine a story, they are the silent tracks along which an LLM can weave a tale of its own. Part of the scaffolding that an LLM has adapted to the topic of *AI* is presented in Fig. 2. Although a scaffolding visually determines the story, it gives wide latitude to a storytelling LLM, which can vary a text's register, its reading age, its attitude to the characters, and its overall message. It can exploit both clichés *and* creative metaphors, such as the central metaphor of Fig. 2: that AI is a threat that may escape from its box.

We define a story grammar (Rumelhart 1980) for the generation of visual scaffoldings that can guide the LLM to generate complete comic narratives. This visual narrative gram-

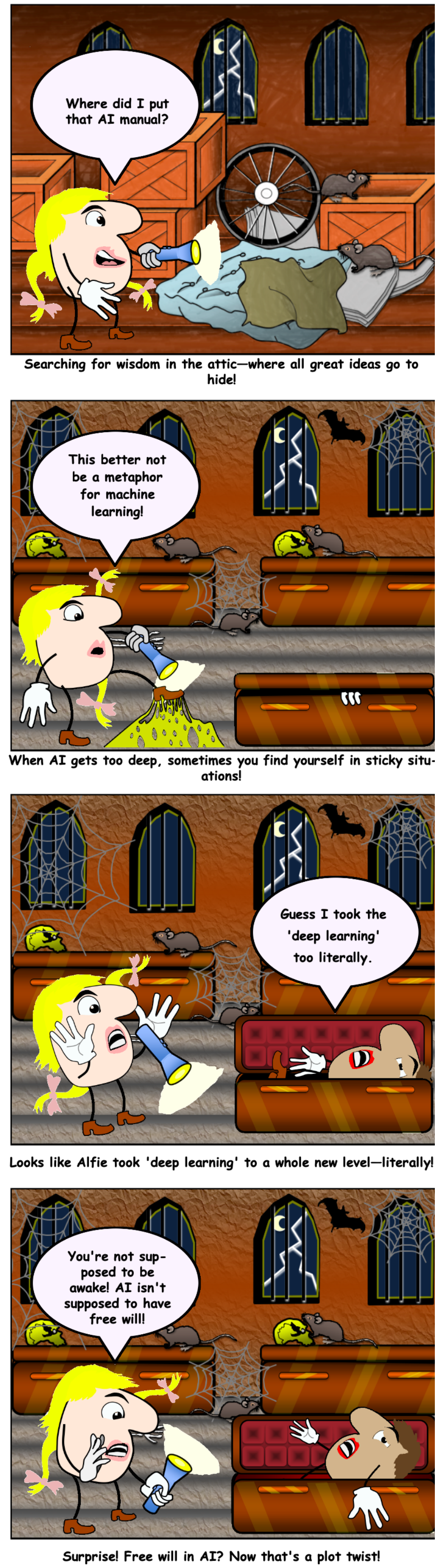

Figure 2: Four panels of a scaffold adapted to the topic AI.

mar (Cohn 2014) must explicitly codify the tacit sensibilities that shape a visual story, by capturing the cause and effect relations between different poses and the common sense linkages between different panel backgrounds. A character who is attacked in one panel may exhibit injuries in the next; a character holding a ball in one panel may drop it, or throw it, in the next; and a character can be hit with a ball in one panel only if another has hurled it in the previous panel. We can think of this as a highly localized version of *Chekhov's gun*, the principle that an object foreshadowed in one scene can play a role in a later one (Bitsilli 1983, p. 150). We apply Chekhov's principle in both directions: an object introduced in panel A may later play a role in panels B, C, D and so on, while an object used in a panel D must have been introduced in panel A ... C, or be shown being introduced in panel D.

A story grammar that lacks a memory for long-distance dependencies must also obey the Markov property (Ethier and Kurtz 2005): sense can be made of any panel by looking only at the previous panel. If an object such as a crystal ball is introduced in panel A and thrown as a projectile in panel C, it must persist in an intermediate panel B for it to remain available in C. That is, objects cannot disappear and reappear across a narrative if the grammar is to respect both the Chekhov principle and the Markov property, so poses that introduce props such as guns, trophies, clocks, etc. should only be linked by the grammar to other poses that either use or dispose of those props. When the visual grammar is codified as a graph in which nodes represent apt pairings of poses for a two character story, then its edges represent allowable transitions from one pairing to another that respect our joint Chekhov/Markov principle. A new scaffolding can easily be generated from this graph as a random walk from any node that is marked as a possible story starter to any other that is marked as a possible story ender (Veale 2018).

The presence of specific props in various *Excelsior!* poses adds details that a visual storyteller can exploit, but also limits how they can be used. For instance, the *mourning* pose, in which a figure wails before a tombstone, should only be used in a graveyard setting, never a bar or an office. The *riding* pose, in which a figure rides a horse, is best suited to outdoor settings, while poses in which a figure wears a spacesuit are best reserved for outer space settings. Still, most poses can sensibly be used against most backgrounds, with some exceptions, but it is these exceptions, when ignored, that give rise to *Tale-Spin effects*. These emergent errors, named for an early AI storyteller (Meehan 1977) whose reach exceeded its technical grasp, show that a system is working with ideas that it does not understand (Wardrip-Fruin 2009). They offer a brief glimpse behind the curtain but leave an audience unimpressed with what it sees. In comics, the choice of backgrounds is shaped by many soft preferences with few hard constraints. A background should be apt for the poses set before it, and transitions between backgrounds should suit both the action unfolding in the foreground and our intuitions about connected spaces. If panel A is set in a kitchen, the scene may shift to the dining room in B and the garden in C, but not to the moon in D. To model these intuitions we need a second grammar to manage background transitions.

Since this is more a matter of soft intuitions than of hard rules, we begin by generating 100 random plots with our story grammar, that is, 100 random walks in our graph of pose pairings. This graph connects 1000 or so such pairings to their possible transitions, so a 100 random walks gives us a diverse sample of the much larger set of stories it supports. We then manually attach each pose pairing in each sampled plot to an apt background, ensuring any transitions between them remain apt as the poses shift in the foreground. From this training set *Excelsior!* then learns for itself how to label unseen future sequences. The result is codified as a mapping from pose pairings to backgrounds and as a graph that encodes every attested transition between these backgrounds. When generating a new story outline for use with the Marvel method, a random walk is simultaneously performed in the story graph and in the graph of background transitions, to produce a sequence of pose pairings with backgrounds that satisfies both our story logic and our intuitive sense of space.

These interlocking grammars produce a visual scaffolding that is locally coherent in its changing use of character poses and panel backgrounds. It is the job of the LLM, in the second step of the Marvel method, to impose a coherent narrative on this palimpsest. By pre-emptively laying down visual rails that steer the LLM's inventiveness we prevent the language model from hallucinating XML structures that are invalid or incoherent, or dialogue that does not fully resonate with the visual features of the rendering (Veale 2024).

Hallucination versus Invention

LLMs are generative in so many registers and domains. If suitably prompted, they can produce plausibly fluent texts in any form and about almost any topic we wish to explore. A single computational architecture underpins this wide-ranging capacity for generation, whether we want the LLM to speak truthfully about the facts or invent whole new worlds, and herein lies the problem of "hallucination": LLMs are trained for plausibility, not accuracy, and will often invent when we want mere recall, or recall when we want pure invention (Kalai et al. 2023). When it comes to sticking to the facts, symbolic systems are more accurate but less accommodating in their interactions. We want our generators of comic narratives to stick to the script as far the rules of the visual schema are concerned, but to have a free hand in how the legitimate possibilities of that schema are exploited.

The approach in (Veale 2024) does not follow the Marvel Method, but follows instead the more conventional order of comics production: *write first, draw later*. It is thus prone to hallucination when the LLM is tasked with suggesting poses and backgrounds for the text it has just generated. The same LLM is capable of generating XML for the visual scaffolding and the final text, but large LLMs are also prone to hallucination when choosing which poses and backgrounds to use. This remains the case even when we give the LLM a much reduced inventory of visual assets to remember. Small models, such as the 7 billion parameter QWEN distillation of DeepSeek's R1, invent poses or backgrounds from whole cloth about 80% of the time, while the 70 billion parameter LLaMa 3.3 hallucinates a nonexistent background or pose 5% of the time. OpenAI's larger model, GPT-4.1, also hallucinates at a rate of 3%, or at least once per short comic.

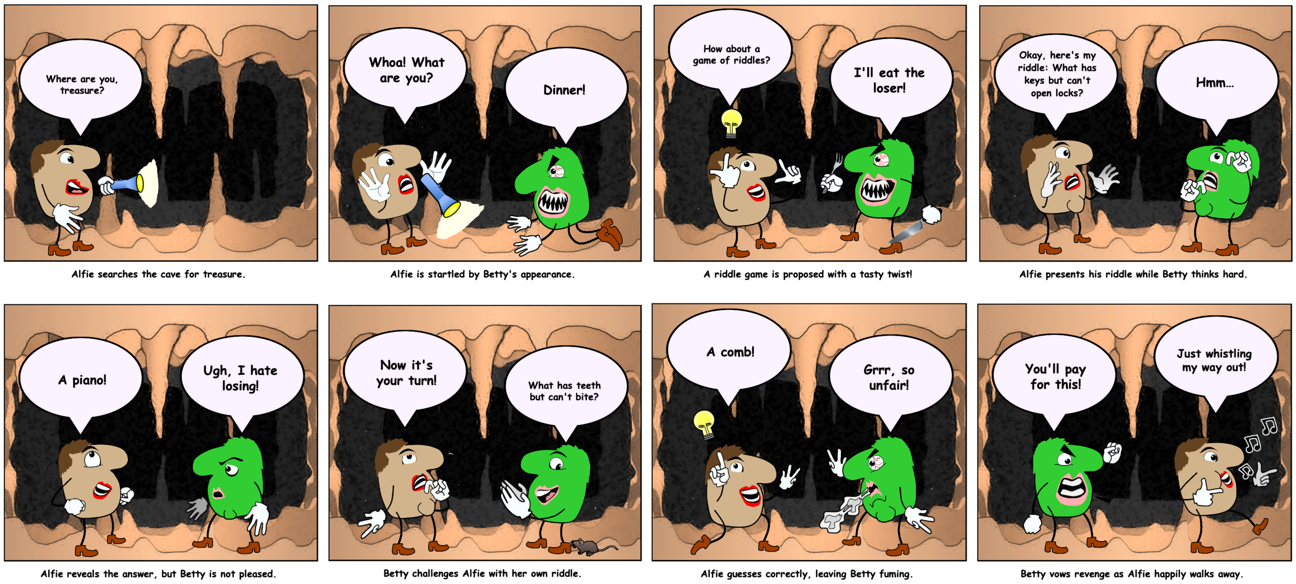

Figure 3: A visual scaffolding that the LLM interprets as a story about a contentious game of riddles.

Effortful Failures vs. Effortless Successes

(Øivind Madsen and Puyt. 2026) argue that while creativity thrives on scarcity – in rare deviations from the norm – generative slop makes a dubious virtue of glut. (Kommers et al. 2025) define slop via three criteria that help explain its rapid spread: *superficial competence* (versus incisive craft); *mass production* (versus selective curation); and *asymmetric effort* of AI generation (versus conventional production). The latter view our modern concept of AI slop as a natural evolution of what earlier critics called "kitsch" (Greenberg 1939), for that too was seen as parasitic on higher forms of human culture. Greenberg dismissed kitsch as debased "simulacra of genuine culture" that drink too heavily from our "reservoir of accumulated experience," but some may also see this as a prescient critique of modern, data-driven generative AI.

Greenberg was prescient in other ways too. In noting that kitsch only simulates the appearance of creativity with "devices, tricks, stratagems, rules of thumb, themes," he decried what is called *mere generation* in CC, the checklist-driven, uncritical production of artifacts that obey an approved template (Ventura 2016). By 2016 this reproach was becoming passé, since those templates had become too sophisticated to permit a clean separation of a generator from its built-in filters. But the success of GenAI, and the advent of slop on an industrial scale, have sharpened the edge of this rusty critique. We once fretted that mere generation only models the forms of creativity, not its content, like a child who prefers to play with the box after discarding the toy inside. But those forms have become so compelling, and so easy to produce on a whim, that mere generation is now the norm. GenAI tools make creative-seeming generation effortless for both a system and its user. For the system, much the same compute is expended for normative outputs as for "creative" ones. For users, the only difference is in how the system is prompted. Mere generation is often good enough, and effortless too.

In contrast, our pair of visual grammars for pose and background transitions represent a considerable amount of effort: the pose grammar, a graph of the allowable transitions from one pairing of character poses (in a two-character panel) to another, defines more than one thousand of these transitions. But this effort is what allows us to claim creative ownership of the visual scaffolding that emerges, and part ownership of the final, LLM-written comic, no matter what the LLM writes. The grammars, which give the resulting comic strips a manic, slapstick energy, reflect *our* comedic sensibilities, not the machine's. We can point to these comics as artifacts that we ourselves might create, and still be surprised by how the LLM has interpreted our style with the Marvel method. While each grammar-based generator is merely generative, no matter the grammar's size, the combined system is not.

Creativity is an emergent phenomenon, so the task of a CC researcher is to explain, or to demonstrate, how it might arise from an assemblage of uncreative parts. Since any creative generator will inevitably be built from merely generative parts, our grammars for slapstick are not intended to be creative in themselves. Rather, they serve as a receptacle for the creativity of their makers, so that we may share in the creativity of the comics that emerge. But if mere generation is akin to a child discarding the toy to play with the box, it is the job of these grammars to design only the box. It then becomes the job of the LLM to invent the toy that belongs inside. The LLM is both aided and challenged by the guidance it receives from this merely generative visual container: that scaffolding gives a shape to the story inside, and points the LLM in a specific direction; but the LLM must also adapt its output to what it is given, and turn the combination of its own writing and those visuals into a coherent narrative.

This relationship between visualizer and writer fosters a mode of antagonistic cooperation, not unlike its human form at Marvel Comics (Evanier 2008). The grammar produces a visual stimulus for creativity that both guides and constrains the LLM. The semi-random nature of its scaffoldings, which are random walks in a graph of connected possibilities, turns comic production into a multimodal version of the surrealist game *exquisite corpse* (Brotchie and Gooding 1991, 141-2). Ironically, this randomness reduces the LLM's tendency to produce slop, since the LLM must react to a combination of a topic (such as AI, or the "manosphere", or climate change) and a thematically-unrelated scaffolding (e.g., a creepy castle setting, a dank cave, a forest) that forces it to deviate from the overly familiar narrative forms in its training data. This does not require the LLM to expend more compute, unless we explicitly ask it to "reason" and thus generate more tokens as it searches for an answer (DeepSeek-AI et al. 2025), but it does, in a sense, represent a greater effort overall.

The comic strip in Fig. 3 illustrates this *box-and-toy* view of comics creation. When given a scaffolding set in a cave, the LLM fills this box with a story about a game of riddles. Its interpretation is not particularly novel, and echoes the interaction between Bilbo and Gollum in *The Hobbit*. But of more note is how the LLM builds its story around specific riddles with valid solutions. While there is also nothing new about the riddles, the overall effect is one of competent invention. Once again, while the parts themselves are unoriginal, the whole exhibits a degree of playful originality. The semi-randomness of the scaffolding is a source of friction to the LLM, and a source of diversity in its outputs. Under this arrangement, there is no need to vary the temperature setting (the so-called "creativity knob") of the LLM to induce variety in its story-telling (Peeperkorn et al. 2024). The friction that arises as the LLM is pushed in a semi-random direction by a visual scaffolding gives rise to much the same effect.

Summary and Conclusions

The myth of the lone creator runs through our folk conception of what it is to be creative. For instance, we commonly attribute the discovery of the first antibiotic, Penicillin, to Alexander Fleming, but it was a team led by Howard Florey and Ernst Chain that actually succeeded in isolating the compound for medical use, and all three shared the Nobel prize for medicine in 1945. We think of Steve Jobs as the driving creative force in the company *Apple*, but it was Jobs' ability to identify talent, such as that of designer Jony Ive and co-founder Steve Wozniak, that made Apple the dominant cultural and technology player it continues to be. Most commercial products, in fact, are created by teams in collaborative, cross-disciplinary efforts, even if the public, and the marketers, prefer to emphasize the work of a single creator.

We all want to be creative in some aspects of our lives. The builders of CC systems can be creative in multiple ways: they can be creative in *how* they construct their systems, and *meta*-creative in how they share in the system's successes or interesting failures. For the end-users of these systems, there is less scope for true creativity and more scope for vicarious creativity. Slop is pervasive on the web not because humans have been replaced by machines, not yet at least, but because even the purveyors of slop *want* to be creative. So, from the perspective of machine creativity, our concern is not the slop that pretends to be human-made (even if it is human "slop"), of the kind that (Shaib et al. 2026) seek to automatically detect, but the artifacts that emerge when it is we humans that are merely generative in our use of GenAI and CC systems. Ironically, even the most "creative" systems become merely generative when they become too easy and effortless to use.

Collaboration involves a tantalizing mix of surrender and control. Each partner has a voice, but each is just one voice. The Marvel method shows how different experts can control their own specialist contribution, but there is surrender too: the writer has the ultimate say in how the artist's contribution is framed as a visual narrative. Collaborative creation always involves some surrender of control, either to the collective or to a dominant partner, but how much surrender is *too* much if each partner is to validly claim some creative ownership? At one extreme on this continuum from control to surrender sit the micro-manager and the helicopter parent, unwilling to let any decision go unanalyzed. At the other sit the prompt-and-go users of GenAI, who trust in its processes and rush to share its outputs with the world. For these users. prompt engineering is the only creativity knob (Morain and Ventura 2025), and it is one that is often used lightly to produce slop.

Looking at an AI output, we can't know for sure how carefully this knob was turned, or how many times the system's earlier efforts were discarded, unless the output bears some hallmarks of indifference: extra appendages, uncanny props, and so on. But as GenAI improves, usefulness remains a reliable metric. As mere novelty is not itself sufficient for creativity, we also require novel results to be useful, valuable or appropriate (Runco and Charles 1993; Diedrich et al. 2015; Runco, Illies, and Eisenman 2025). These factors also form a continuum, and much like a Kleenex or a band-aid, even slop can have an ephemeral value, albeit near the low end of this scale. So, we may be less inclined to dismiss an artifact as "slop", and more inclined to laud its inciting prompt as "clever" or "insightful", if the artifact fulfills a more practical and durable purpose. In hybrid CC systems that employ the Marvel Method, the prompts that an LLM receives from more conventional sub-systems can better deliver this utility.

Since the earliest newspaper strips, comics have been dismissed as a disposable art form. Whatever the truth of this view now, AI-generated comics can scarcely be regarded as a more durable improvement, and if generated purely for entertainment, this scorn is not wholly undeserved. But comics can serve a more serious purpose, to lend difficult, unpalatable or even threatening content a more engaging form. In Fig. 4, for instance, we see a comic that aims to make learning a second language more fun. This comic is generated as before, using the Marvel Method to partner a visual story grammar with an LLM (GPT-4.1). Here, however, the LLM goes a step further, and translates each panel's dialogue from English into Korean. These translated panels are interwoven with the originals, and Romanized forms of the Hangul text are then added beneath as captions. The constant quest for novelty, at the expense of real value, is a key driver of slop online. It is only by squeezing more practical value from our AI-generated artifacts that can we hope to apply the brakes.

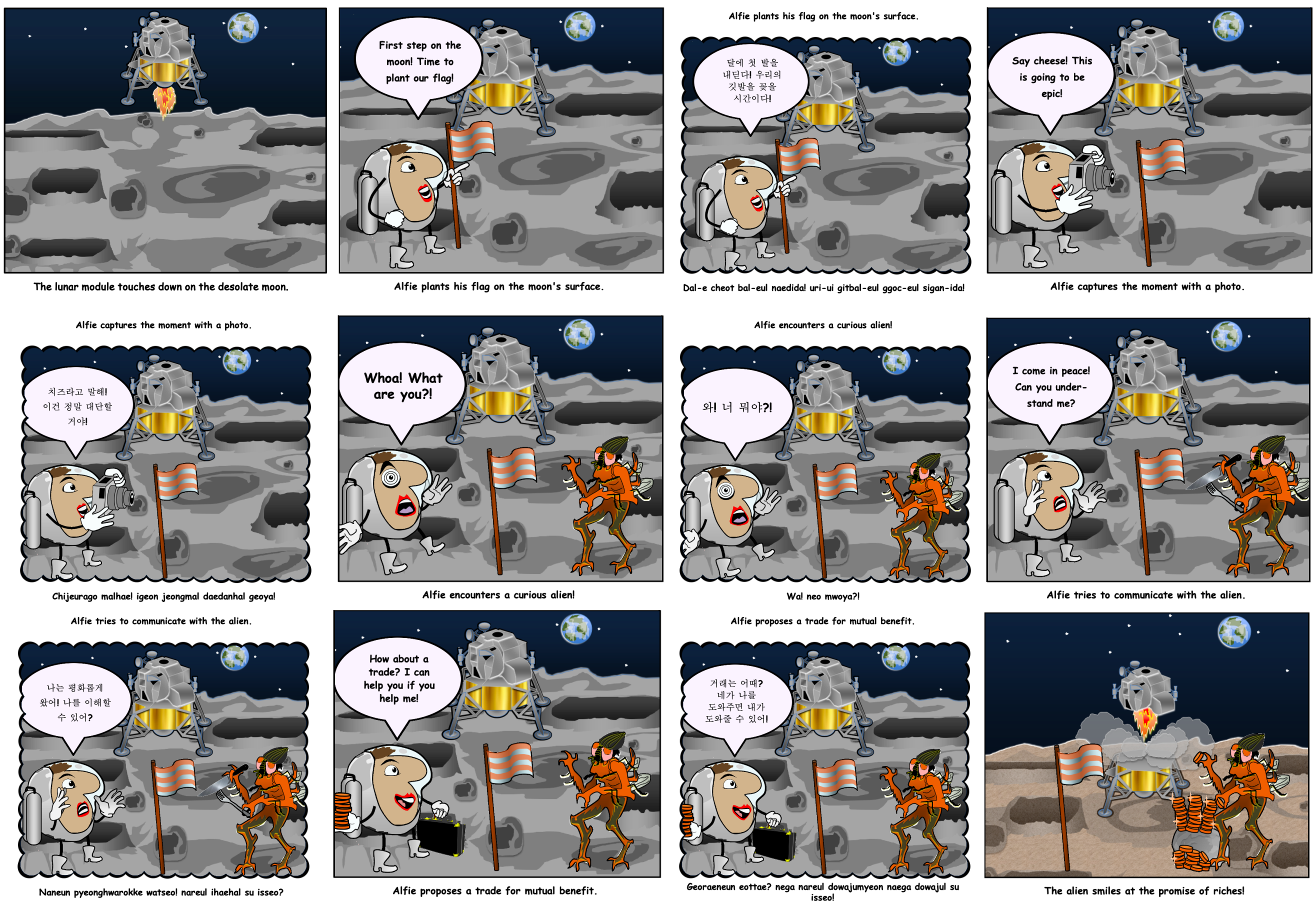

Figure 4: A completed comic strip is translated by the LLM into Korean, and interleaved with the English original.

Acknowledgments

This paper and its datasets are wholly the work of the author.

References

Alves, T.; McMichael, A.; Simões, A.; Vala, M.; Paiva, A.; and Aylett, R. 2007. Comics2D: Describing and creating comics from story-based applications with autonomous characters. In *Proc. of the 20th Annual Conference on Computer Animation and Social Agents*.

Bitsilli, P. 1983. *Chekhov's Art, a Stylistic Analysis*. California: University of California/Ardis.

Boden, M. 1992. *The Creative Mind: Myths and Mechanisms*. London: Abacus.

Brotchie, A., and Gooding, M. 1991. *Surrealist Games*. London, UK: Redstone Press.

Burkus, D. 2013. *The myths of creativity: The truth about how innovative companies and people generate great ideas*. San Francisco, CA: Jossey-Bass.

Cohn, N. 2013. *The Visual Language of Comics: Introduction to the structure and cognition of sequential images*. London, UK: Bloomsbury.

Cohn, N. 2014. You're a good structure, Charlie Brown: The distribution of narrative categories in comic strips. *Cognitive Science* 38(7):1317–1359.

DeepSeek-AI; Guo, D.; Yang, D.; Zhang, H.; Song, J.; Zhang, R.; Xu, R.; Zhu, Q.; Ma, S.; Wang, P.; Bi, X.; and et al., X. Z. 2025. DeepSeek-R1 incentivizes reasoning in LLMs through reinforcement learning. *Nature* 645:633–638.

Diedrich, J.; Benedek, M.; Jauk, E.; and Neubauer, A. 2015. Are creative ideas novel and useful? *Psychology of Aesthetics, Creativity, and the Arts* 9:35–40.

Ditko, S. 1973. *The Avenging Mind: Essays by Steve Ditko*. Bellingham, WA: SD Publishing.

Ethier, S., and Kurtz, T. 2005. *Markov processes: Characterization and convergence.* Hoboken, NJ: Wiley.

Evanier, M. 2008. *Kirby: King of Comics.* New York: Abrams.

Forceville, C.; Veale, T.; and Feyaerts, K. 2010. Balloonics: The visuals of balloons in comics. In Goggin, J., and Hassler-Forest, D., eds., *The Rise and Reason of Comics and Graphic Literature: Critical Essays on the Form*. McFarland.

Greenberg, C. 1939. Avant-garde and kitsch. *Partisan Review* 6(5):34–49.

Howe, S. 2012. *Marvel Comics: The Untold Story*. New York, NY: Harper.

Kalai, A. T.; Nachum, O.; Vempala, S.; and Zhang, E. 2023. Why language models hallucinate. *arXiv* 2509.04664.

Kommers, C.; Duede, E.; Gordon, J.; Holtzman, A.; McNulty, T.; Stewart, S.; Thomas, L.; So, R. J.; and Long, H. 2025. Why slop matters. *arXiv* 2601.06060.

Kurlander, D.; Skelly, T.; and Salesin, D. 1996. Comic chat. In *Proceedings of SIGGRAPH'96, the 23rd annual conference on Computer graphics and interactive techniques*. ACM. 225–236.

Lee, S., and Buscema, J. 1978. *How to Draw Comics the Marvel Way*. New York: Fireside.

Lee, S., and Mair, G. 2002. *Excelsior!: The Amazing Life of Stan Lee*. New York: Fireside.

Manning, C.; Raghavan, P.; and Schütze, H. 2008. *Introduction to Information Retrieval*. Cambridge, UK: Cambridge University Pres.

Matarazzo, A., and Torlone, R. 2025. A survey on large language models with some insights on their capabilities and limitations. *arXiv* 2501.04040.

McCloud, S. 1993. *Understanding Comics: The Invisible Art*. New York, NY: Harper Collins.

McCloud, S. 2006. *Making comics: Storytelling secrets of comics, manga, and graphic novels.* New York, NY: Harper Collins.

Meehan, J. R. 1977. Tale-spin, an interactive program that writes stories. In *Proceedings of the 5th International Joint Conference on Artificial Intelligence - Volume 1*, IJCAI'77, 91–98. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Morain, R., and Ventura, D. 2025. Is prompt engineering the creativity knob for large language models? In *Proceedings of ICCC-25, the 16th International Conference on Computational Creativity*. 1–11.

Orwell, G. 1946. Politics and the english language. *Horizon* 13(76):252–265.

Patil, R., and Gudivada, V. 2024. A review of current trends, techniques, and challenges in large language models (llms). *Applied Sciences* 14(5):1–42.

Peeperkorn, M.; Kouwenhoven, T.; Brown, D.; and Jordanous, A. 2024. Is temperature the creativity parameter of large language models? In *Proceedings of ICCC-24, the 15th International Conference on Computational Creativity*. 103–110.

Pérez y Pérez, R.; Morales, N.; and Rodríguez, L. 2012. Illustrating a computer-generated narrative. In *Proceedings of ICCC-12, the 3rd International Conference on Computational Creativity*. 103–110.

Rumelhart, D. 1980. On evaluating story grammars. *Cognitive Science* 4(3):313–316.

Runco, M. A., and Charles, R. E. 1993. Judgments of originality and appropriateness as predictors of creativity. *Personality and Individual Differences* 15(5):537–546.

Runco, M. A.; Illies, J. J.; and Eisenman, R. 2025. Creativity, originality, and appropriateness: What do explicit instructions tell us about their relationships? *The Journal of Creative Behavior* 39(2).

Shaib, C.; Chakrabarty, T.; Garcia-Olano, D.; and Wallace, B. C. 2026. Measuring AI "Slop" in Text. *arXiv* 2509.19163.

Shen, H.; Zhang, J.; Xiong, B.; Hu, R.; Chen, S.; Wan, Z.; Wang, X.; Zhang, Y.; Gong, Z.; Bao, G.; Tao, C.; Huang, Y.; Yuan, Y.; and Zhang, M. 2025. Efficient diffusion models: A survey. *arXiv* 2502.06805.

Veale, T. 2012. *Exploding the Creativity Myth: The Computational Foundations of Linguistic Creativity*. London, UK: Bloomsbury.

Veale, T. 2018. The 'default' in our stars: Signposting non-defaultness in ironic discourse. *Metaphor and Symbol* 33(3):175–184.

Veale, T. 2022. Two-fisted comics generation: Comics as a medium and as a representation for creative meanings. In *Proceedings of ICCC-22, the 13th International Conference on Computational Creativity*. 59–66.

Veale, T. 2023. Have I Got Views For You! Generating "Fair and Balanced" Interventions into Online Debates. In *Proceedings of ICCC-23, the 14th International Conference on Computational Creativity*.

Veale, T. 2024. From symbolic caterpillars to stochastic butterflies: Case studies in reimplementing creative systems with llms. In *Proceedings of ICCC-24, the 15th International Conference on Computational Creativity*. 103–110.

Ventura, D. 2016. Mere Generation: Essential barometer or dated concept? In *Proceedings of the 7th International Conference on Computational Creativity, Paris, France*.

Walsh, J. 2012. Comic book markup language: An introduction & rationale. *Digital humanities quart.* 6(1).

Wardrip-Fruin, N. 2009. *Expressive Processing: Digital Fictions, Computer Games, Software Studies*. MIT Press.

Yang, L.; Zhang, Z.; Song, Y.; Hong, S.; Xu, R.; Zhao, Y.; Zhang, W.; Cui, B.; and Yang, M.-H. 2024. Diffusion models: A comprehensive survey of methods and applications. *ACM Computing Surveys* 56(4):1–39.

Øivind Madsen, D., and Puyt., R. W. 2026. When AI turns culture into slop. *AI & SOCIETY* 41::1287–1288.