

Concept Paper

Not peer-reviewed version

---

# Prompt-Screenplay: A Multimodal Framework for Human-AI Co-Creation and Authorship Protection in Generative Media

---

[Stanislav Lauk-Dubitskiy](#)\*

Posted Date: 17 October 2025

doi: 10.20944/preprints202510.1371.v1

Keywords: prompt engineering; generative AI; human-AI co-creation; screenplay format; copyright protection; multimodal content generation; authorship verification; computational creativity



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

*Concept Paper*

# Prompt-Screenplay: A Multimodal Framework for Human-AI Co-Creation and Authorship Protection in Generative Media

Stanislav Lauk-Dubitskiy

MIREA Russian Technological University, Russia; laukdubitskiy@mirea.ru

## Abstract

The rapid advancement of generative artificial intelligence has fundamentally transformed content creation, yet existing screenplay formats fail to bridge the gap between human artistic intention and machine interpretation while ensuring legal protection of authorship. This paper introduces the prompt-screenplay—a novel hybrid format that simultaneously functions as human-readable narrative, machine-executable instruction set, and legally verifiable proof of authorial intent. We propose a five-layer semiotic architecture integrating: immutable authorial vision through master prompts, visual semantic node-based narrative structure, adaptive presentation for target audiences, personalized viewer experience with biometric feedback, and comprehensive analytics for content evaluation. The system employs morphological modifiers embedded in natural text, inline commands for AI instruction, and cryptographic consistency tags for preventive copyright protection through blockchain registration. A pilot transformation of the science fiction screenplay “SKOR” demonstrates practical applicability for multimodal generation across video, audio, comics, games, and interactive media, with particular efficacy for vertical mobile formats dominating contemporary consumption. Unlike existing approaches, prompt-screenplays preserve artistic expressiveness while encoding machine-readable metadata at the conception stage, addressing the critical legal gap in AI-generated content authorship.

**Keywords:** prompt engineering; generative AI; human-AI co-creation; screenplay format; copyright protection; multimodal content generation; authorship verification; computational creativity

---

## 1. Introduction

Contemporary generative artificial intelligence models—from text-based LLMs to multimodal text-to-video systems such as OpenAI’s GPT-4, Google’s Gemini, Sora, and Runway Gen-3—are fundamentally transforming the nature of authorship in digital media. While traditional screenplays served exclusively as blueprints for human performers, today they increasingly function as input signals for AI agents capable of directly generating images, sound, motion, and interactive narratives [1,2]. However, existing screenplay formats fail to account for machine perception specifics, multimodal generation requirements, and legal documentation of authorial intent under collaborative creativity conditions with AI [3,4].

Recent research demonstrates that effective human-AI collaboration requires structured frameworks that preserve human agency while leveraging computational capabilities [5]. The Creative CRAFT framework shows that modular prompt design significantly improves output quality and creativity in collaborative scenarios [6]. Similarly, role-based agent systems in screenwriting contexts demonstrate enhanced coherence and narrative consistency when human creative input is systematically integrated with AI capabilities [7].

The question of copyright remains particularly acute: when AI generates video from a prompt, authorship attribution becomes ambiguous between the human prompt creator and the generating model [8]. The U.S. Copyright Office’s denial of registration for Dr. Stephen Thaler’s AI-generated

artwork “A Recent Entrance to Paradise” established that works “autonomously generated by an AI” lack “traditional human authorship,” underscoring the urgent need for mechanisms documenting human creative input at the conceptual stage [9].

Digital humanities research has explored prompts as creative control units, yet most work focuses either on text prompt poetics or automatic storyboard generation, neglecting shot dynamics, vertical mobile cinema formats, and metadata embedding at conceptual stages [10,11]. The proliferation of short-form vertical video content demands new narrative structures optimized for mobile consumption patterns while maintaining artistic integrity [12].

This study addresses three critical gaps: (1) the absence of screenplay formats designed for AI interpretation while preserving human readability, (2) the lack of authorship protection mechanisms embedded in creative documents at conception, and (3) the need for adaptive content structures supporting diverse media formats from single source documents.

### 1.1. Research Questions

This investigation addresses four primary research questions:

1. How can screenplay formats integrate human artistic expression with machine-readable instructions while maintaining narrative coherence?
2. What technical and legal mechanisms enable authorship protection for AI-assisted creative content at the conception stage?
3. How can morphological modifiers embedded in natural text provide precise AI instructions without disrupting human readability?
4. What evaluation frameworks assess the effectiveness of prompt-screenplay approaches across different media formats?

### 1.2. Contributions

Our contributions include: (1) a novel prompt-screenplay format enabling simultaneous human-AI collaboration and authorship protection, (2) a five-layer semiotic architecture for multimodal content generation, (3) morphological modifier systems preserving natural language readability while encoding machine instructions, and (4) empirical validation through the “SKOR” screenplay transformation demonstrating practical applicability.

## 2. Related Work

### 2.1. Prompt Engineering and Creative Collaboration

Prompt engineering has emerged as a critical discipline for effective human-AI collaboration in creative contexts. Recent frameworks demonstrate that structured prompt design significantly impacts creative output quality and user satisfaction [6]. The Creative CRAFT structure, incorporating context, role, creative direction, and constraints, shows measurable improvements in creativity metrics when applied to narrative generation tasks.

Research on human-AI collaborative writing reveals that collaboration design choices fundamentally affect output quality, diversity, and creator satisfaction [5]. Systems that preserve human creative input while providing AI augmentation achieve better user acceptance and perceived authorship compared to fully automated approaches [13]. The GhostWriter system demonstrates that interfaces enabling style teaching and explicit control increase personalization and user satisfaction in collaborative writing scenarios [14].

Role-based agent systems show particular promise for screenplay applications. Research demonstrates that role assignment and actor role-playing improve generated screenplay coherence and narrative interest through structured decomposition of creative tasks [7]. These findings suggest that prompt-screenplay architectures should incorporate explicit role definitions for different creative functions.

## 2.2. Multimodal AI Content Generation

Contemporary multimodal AI systems enable unprecedented cross-modal content generation from textual prompts. Comprehensive surveys of text-to-image, text-to-video, and text-to-audio pipelines reveal sophisticated tokenization approaches and tool-augmented agents capable of converting narrative descriptions into visual and auditory assets [15]. These capabilities create opportunities for screenplays to function as comprehensive production documents rather than merely textual blueprints.

Large language models demonstrate effectiveness in human-AI co-creation scenarios across diverse creative domains, from robotic choreography to interactive storytelling [16]. Iterative, example-based prompting with human feedback loops consistently improves artistic impact and creator satisfaction compared to single-pass generation approaches.

## 2.3. Authorship Protection and Copyright in AI-Generated Content

Legal frameworks for AI-generated content authorship remain underdeveloped, creating significant uncertainty for creative professionals [8,9]. Current approaches focus primarily on post-generation attribution rather than preventive authorship documentation during the creative process. Blockchain-based systems offer potential solutions for immutable authorship records, but require integration into creative workflows at the conception stage [17].

Watermarking taxonomies for AI-generated content provide technical foundations for provenance tracking and authenticity verification [18]. However, most approaches focus on post-generation marking rather than embedding authorial intent within source documents. Recent proposals for AI royalty frameworks suggest treating licensed AI models as compensable assets, creating new economic models for creative collaboration [19].

# 3. Methods

## 3.1. Research Design

This study employs a design science research (DSR) approach, combining theoretical modeling with practical artifact development. The primary objective is to create a new form of creative document capable of functioning simultaneously as a work for human perception and as machine-readable instructions for generative AI systems. The research is grounded in principles of human-AI co-creativity and structured prompt engineering.

The methodology comprises three main stages:

1. **Conceptualization:** development of a theoretical model of the prompt-screenplay based on analysis of existing formats (classic screenplay, Fountain, game scripts) and requirements of contemporary generative models.
2. **Formalization:** creation of a system of structural elements, morphological modifiers, and commands ensuring balance between human readability and precision of machine interpretation.
3. **Validation:** pilot application of the developed system for content creation and evaluation of protective mechanism functionality.

### 3.1.1. Prompt-Screenplay Creation Tools

Prompt-screenplay creation is possible both in a specialized editor and in any text editor with mandatory use of tags and special commands.

### 3.1.2. Typology of Prompt-Screenplays

The developed typology includes 4 main types of prompt-screenplay, distinguished by creation method and purpose:

**Self-Written Prompt-Screenplay.** Created by a human author with manual addition of markers and tags for the AI agent. Includes classic directorial screenplay blocks with additional structural elements for machine interpretation.

**Prompt-Screenplay Co-Written with AI Agent.** Created through human-machine collaboration. The AI agent assists the author in real-time, marking keywords, forming metadata for video and audio generation, and suggesting visualization options.

**Generated Prompt-Screenplay.** Reconstructed from already-created content through reverse engineering and used for copyright protection when working with AI. The system collects prompts and metadata used at all production stages and compiles them into a unified document.

**Prompt-Screenplay with Mapped Authorial Intent.** Created in a specialized visual semantic editor for precise manual transformation by a professional team into a game, film, or book. Such a screenplay is distinguished by inclusion of a system of semantic nodes located on the storyline with different types of connections, reflecting both their key information and plot details within each node.

### 3.2. Interconnected elements System:

The prompt-screenplay consists of the following interconnected elements:

#### 3.2.1. Classic Screenplay Blocks

Traditional elements of screenplay format are preserved:

- **Scene Heading (Slug Line):** location, time of day, shooting type
- **Action:** description of visible actions and sounds
- **Character:** name of speaking character
- **Parenthetical:** dialogue annotations
- **Dialogue:** character speech
- **Transition:** transitions between scenes

#### 3.2.2. Technical Location and Shooting Markers

An extended system of spatial markup: **[PV]** — pavilion (controlled studio environment); **[IR]** — interior (indoor setting); **[EX]** — exterior (outdoor shooting); **[CS]** — combined shooting (combination of location and pavilion elements); **[LS]** — location shooting (real location); **[VS]** — visual generation style.

These markers serve as indications for the AI generator about environment type and lighting control level, which is critical for text-to-video models.

#### 3.2.3. Shot Type and Timing Markers

Shot markers determine framing for AI image and video generators, ensuring visual variety and directorial control: **[ES]** — establishing shot (wide shot); **[MS]** — medium shot; **[CU]** — close-up; **[EU]** — extreme close-up (detail shot); **[PS]** — panoramic shot; **[TS]** — tracking shot; **[TI]** — timing (scene duration in minutes and seconds)

#### 3.2.4. Audio Markers

Markers determine framing for AI sound and music generators: **[SO]** — description of sounds, noises, sound effects; **[MO]** — monologue; **[DI]** — dialogue; **[VO]** — thought, voice-over; **[OS]** — off-screen sound or speech; **[MU]** — music (at end of sentence)

#### 3.2.5. Final Scene Markers

Placed after scene completion and document its key characteristics: **[EM]** — dominant scene emotion; **[K]** — key idea or meaning; **[I]** — presence of interactive elements

### 3.3. Morphological Modifier System:

The framework employs morphological modifiers embedded within natural language text to provide machine-readable instructions without disrupting human comprehension.

#### 3.3.1. Letter Modifiers

##### **Doubling the final letter (word → wordd):**

- *Function:* adding visual emphasis to an object
- *Application:* AI generator increases visual significance of the object in frame
- *Example:* "On the tablee lies a letter" → table becomes compositional center

##### **Tripling the final letter (word → worddd):**

- *Function:* adding sound from an object
- *Application:* audio generator creates corresponding sound effect
- *Example:* "Doorrr creaked" → creaking door sound is added

##### **Removing the final letter (word → wor):**

- *Function:* marking inessential word
- *Application:* can be replaced by synonym or removed without loss of meaning
- *Example:* "Person quickl walked" → "quickly" may be replaced by "swiftly"

##### **Capitalizing word (mid-sentence):**

- *Function:* semantic emphasis
- *Application:* AI accounts for heightened concept significance
- *Example:* "He felt Fear" → emotion becomes central theme of moment

##### **Abbreviated writing (e.g., "swift." instead of "swiftly"):**

- *Function:* atmospheric word for detailing
- *Application:* word influences overall mood but not specific objects
- *Example:* "Rain drizzled quietly, gloom." → "gloomily" sets generation tonality

##### **Adding number after final letter:**

- *Function:* denoting word weight in terms of priority and value
- *Application:* allows independent prioritization of words for use in generation

##### **Writing word with final capital letter:**

- *Function:* marker that object should be animated
- *Application:* AI considers object animation options on scene according to context

#### 3.3.2. Spatial Modifiers

##### **Extra space after word (word\_):**

- *Function:* pause or blackout
- *Application:* creating temporal interval or fade-out effect
- *Example:* "He left \_ room emptied" → pause between actions

##### **Removing space between words (word1word2):**

- *Function:* visual connection between objects
- *Application:* AI generates objects in close compositional proximity
- *Example:* "hand-gun" → hand and gun perceived as single visual unit

##### **Adding nearest keyboard layout letter (word → wordr):**

- *Function:* hidden emphasis with subsequent mention
- *Application:* visual element is embedded that will manifest later
- *Example:* "mirrore hung" → mirror will be important in following scenes

### 3.4. Inline Command System

Inline commands are single letters inserted between words or at beginning/end of sentences that transmit meta-instructions to the AI agent.

### 3.4.1. Commands Between Words

Inline commands between words include: **V** – important element, must be considered in generation; **C** – critical for consistency visual element (cannot change between scenes); **A** – accent (highlight compositionally or with light); **X** – mystery (element for creating intrigue, incomplete information)

*Example: “He took K red K book from shelf” → red book must look identical in all scenes.*

### 3.4.2. Commands at Sentence Beginning

Define character of entire sentence: **I** – interactive element (potential audience interaction point); **K** – key scene idea (central meaning for AI interpretation); **A** – atmosphere (mood dominates over action); **F** – appearance from blackout (fade-in)

*Example: “A Fog enveloped street, hiding lanterns.” → priority on creating atmosphere.*

### 3.4.3. Commands at Sentence End

Define post-processing or additional actions: **I** – AI improvisation possible (parameters not rigid); **M** – add musical accompaniment; **H** – request for help (AI suggests improvement options); **F** – fade-out or blackout; **P** – pause (can be filled with ambient sound); **N** – add to notes (for future use); **Q** – quote or cultural reference (AI accounts for reference); **A** – animate entire scene described in sentence

*Example: “He turned and left. M Z” → scene concludes with music and fade-out.*

## 3.5. Creation and Interaction Processes

### 3.5.1. Interactive Author Interviewing

During screenplay writing, a voice AI agent conducts structured interviews with the author, asking questions about visual references and stylistics, key emotional moments, character consistency requirements, musical preferences for scenes, and interactive possibilities. Responses are automatically converted into markers and embedded into corresponding screenplay sections. This approach minimizes cognitive load on the author and allows maintaining focus on the narrative.

### 3.5.2. Gamification of the Writing Process

The system offers a game-based approach to screenplay creation:

- **Points:** author receives scores for using markers, shot variety, creating visual accents
- **Tasks:** AI sets mini-challenges (“add detail for consistency,” “create sound accent”)
- **Levels:** advanced markup capabilities unlock as writing progresses
- **Achievements:** rewards for completed scenes, balanced timing, effective use of morphological modifiers

Gamification increases engagement and trains authors in effective system use.

### 3.5.3. Interactive Screenplay Reading

A specialized prompt-screenplay reading mode has been developed with sequential word appearance, pauses, reader reactions, reader commands, and improvisational branches. This mode transforms reading into an interactive experience and collects audience perception data.

## 3.6. Cryptographic Consistency Framework

Authorship protection employs cryptographic consistency tags that create immutable records of creative decisions throughout the collaborative process. The framework generates cryptographic hashes for all creative elements, enabling verification of authorial contribution and detection of unauthorized modifications.

**Blockchain Registration** creates permanent records of creative intent and authorial decisions at each stage.

**Provenance Tracking** maintains comprehensive audit trails of all creative modifications.

**Consistency Verification** enables automatic detection of unauthorized modifications through cryptographic comparison.

## 2.9. Tools and Technological Architecture

### 2.9.1. Prompt-Screenplay Editor: Five-Layer System Architecture

A specialized editor is proposed, based on node-oriented system logic (blueprint architectures) and video editors, which enables creation, editing, and analysis of prompt-screenplays through multi-level semantic structure. The editor can function as a standalone application or as a plugin for existing screenplay tools (Final Draft, Celtx), integrating AI agents for automated semantic analysis and quality, meaning, and content originality assessment.

#### Conceptual Foundation of the Editor

The editor implements a five-layer architecture where each layer represents a separate level of semantic analysis and content interaction. This structure allows:<sup>[16]</sup>

1. Documenting authorial intention independent of technical implementation
2. Structuring narrative through visual semantic nodes
3. Adapting content for different audiences and perception contexts
4. Personalizing experience based on individual viewer characteristics
5. Synthesizing analytics for comprehensive work evaluation

The architecture is grounded in Roland Barthes' semiotic principles (denotation/connotation), Hans-Robert Jauss's reception theory, and contemporary computational narrative analysis approaches.

#### First Layer: Static Authorial Intent (Master Prompt)

The first layer represents the conceptual core of the work—immutable intentional semantics that defines the author's original vision and serves as reference point for all subsequent transformations.

#### Structure of Master Screenplay Prompt

The master prompt functions as a meta-document describing not specific scenes but the overall logic of the work and principles of its development. It includes two main components:

**Descriptive part** contains:

- Genre and typological characteristics: genre definition (drama, comedy, thriller), subgenres, modality (linear/nonlinear structure, interactivity)
- Type and form of realization: target format (short-form vertical video, feature film, interactive game, comic, novel)
- Complexity parameters: narrative complexity levels (number of plot lines, temporal layers, degree of metaphoricality)
- Target audience: demographic, psychographic, and cultural characteristics of intended viewers

**Instructional part** defines:

- Number of nodes: total number of semantic blocks, their hierarchy and grouping
- Work size: timing, volume (for text), number of scenes
- Form of expression: dominant media (visual, audio, textual), showing/telling ratio
- Implementation tools: technical requirements for generative models

#### Contextual Elements and Verification

The master prompt includes author-verifiable elements ensuring authorial vision consistency:

- Personal meaning: unique authorial interpretations that should not be changed by AI
- References and allusions: cultural codes, allusions, citations with explicit marking
- Factual data: historical, geographical, scientific facts requiring accuracy
- Homages: conscious references to other works
- "World bible": compendium of fictional universe rules (for sci-fi/fantasy)

Each element undergoes mandatory author verification: the AI agent cannot independently modify these parameters without explicit author permission.

### Format Representations of Master Prompt

The system supports multiple input formats:

**Text format:** structured document with tags

[GENRE: psychological thriller]

[FORM: vertical video, 15 episodes × 90 sec]

[TARGET\_AUDIENCE: 18-35 years, active TikTok users]

[KEY\_IDEA: nature of memory and identity]

[TONALITY: tense, claustrophobic]

**Visual schema:** interactive mind-map with parameter nodes and connections between them, created through drag-and-drop interface.

**Interview with AI editor:** voice or text dialogue where AI sequentially asks questions to identify all master prompt parameters. The system uses Socratic dialogue methodology to extract implicit author knowledge.

#### Semiotic Principles of First Layer

The first layer implements intentional semantics—a level where authorial intent exists independent of specific form of expression. This corresponds to Saussure’s concept of “signified” in semiotics: meaning preceding the sign.

#### Second Layer: Visual System of Semantic Nodes with Screenplay Description

The second layer represents the main workspace for creating prompt-screenplay—an extended text-visual editor organized according to node-based system principles.

#### Node Architecture

Each node (semantic unit) represents a logically complete narrative unit—scene, episode, or micro-episode. Nodes have different sizes corresponding to their timing and semantic density.<sup>[16]</sup>

#### Visual Node Representation:

- Color coding: by emotional tonality, genre, intensity
- Block size: proportional to timing or element quantity
- Shape: rectangle for linear scenes, diamond for choice points, circle for cyclical elements
- Status icons: completion indicators, AI-generated content presence, connections

#### Node Content

Each node includes the following components:

1. **Scene description with prompts:** Textual description of actions, dialogue, and visual elements with embedded prompts and markers for AI generation, as previously described in prompt-screenplay structure.
2. **Connections with preceding and subsequent nodes:** Linear connections (chronological), parallel montage, flashback/flashforward with explicit marking; Semantic connections (cause-effect, thematic, symbolic).
3. **Off-chain semantic connections:** Special type of connections not visible in linear viewing but existing at the meaning level—hidden parallels, leitmotifs, symbolic arcs.
4. **Nested hidden nodes:** Within large nodes, nested micro-nodes can exist—detailing visible only when “expanding” parent node. This allows working at different detailing levels: macro-view (20-50 nodes), meso-view (100-200 nodes), micro-view (500+ nodes).
5. **Self-reflexive references:** Elements commenting on the work’s own nature—metafictional devices, fourth wall breaking, references to creation process.

#### Second Layer Editor Interface

- **Timeline mode:** horizontal time scale with nodes positioned proportionally to timing
- **Graph mode:** network visualization emphasizing semantic connections
- **Text mode:** classic screenplay format with one-click switching to node visualization
- **Split mode:** simultaneous text and graph display for navigation convenience

#### Third Layer: Dynamic Field of Viewer Perception

The third layer is a superstructure over the second layer on the timeline, determining how semantic nodes will be presented to specific audiences under conditions of maximum attention focus.

#### **Layer Concept**

This layer proceeds from understanding that the same content is perceived differently depending on consumption method, cultural context, education level, and aesthetic experience. The third layer represents reference perception conditions—a model of ideal viewer with maximum attention and appropriate cultural preparation.

#### **Temporal Node Deployment**

The layer defines presentation timeline—how nodes are shown over time: linear sequence, parallel montage, branching (choice points), adaptive sequence where AI dynamically selects next node based on viewer reactions.

#### **Means of Meaning Expression**

For each node, the third layer forms comprehensive instructions/prompts accounting for visual means (composition, color palette, lighting), audio means (music, sound effects, dialogue), textual means (for books/comics), and interactive means (for games).

#### **Fourth Layer: Personalized Semantic Block**

The fourth layer activates during work demonstration and represents a pluggable module to video player, game engine, or reading app.

#### **Layer Functions**

**Data collection during consumption:** behavioral data (pauses, rewinds, repeat views, reading speed), biometric data (eye tracking, heart rate, skin conductance), interactive data (choices at branching points, comments, ratings).

**Personal reaction:** emotional markers, comments (text or voice notes), questions, alternative suggestions.

**Node editing:** In personal version, viewer receives limited editing tools: changing order of non-critical scenes, choosing alternative visual/audio variants, creating own plot branches. All changes save in personal fork of work without affecting original.

#### **Cognitive and Neuropsychological Characteristics**

The system accounts for individual perception characteristics: information processing speed, dominant perception type (visual, auditory, kinesthetic), attention features (ADHD-friendly mode, meditative mode).

#### **Biological Feedback**

With biometric sensors, the system analyzes attention peaks, emotional reactions, fatigue. Based on biofeedback, AI can suggest breaks, re-show missed important moments, adapt narrative pace in real-time.

#### **Fifth Layer: Summarizing Synthetic Layer**

The fifth layer is a hidden analytical level synthesizing data from all previous layers for comprehensive work evaluation.

#### **Layer Architecture**

The layer functions as a cloud analytics platform that collects anonymized data from all editor instances, aggregates information about the work, applies ML algorithms to identify patterns, and generates final metrics and visualizations.

#### **Machine Learning Algorithms**

The system applies various ML approaches:

- **Natural Language Processing:** sentiment analysis of dialogue and descriptions, topic modeling for thematic cluster identification, stylometric analysis for authorial style assessment
- **Computer Vision** (for visual content): composition and color palette analysis, visual leitmotif detection, visual consistency evaluation between scenes
- **Recommender systems:** collaborative filtering based on viewer data, content-based recommendations by structure similarity, hybrid approaches for personal suggestions

- **Predictive analytics:** predicting work success before release, identifying problematic scenes, optimization recommendations

### Visualizations and Reports

The fifth layer generates interactive dashboards for authors (attention graph by scenes, emotional map, comparison of authorial intent with real perception), producers (commercial success forecast, target audience portrait, optimal distribution channels), and researchers (network visualizations of semantic connections, perception chronology by demographic groups, data export for academic analysis).

This editor represents a comprehensive system that not only helps create prompt-screenplays but accompanies the work through all lifecycle stages—from initial idea to long-term audience impact analysis.

### 2.9.2. Generative Pipeline

The architecture for transforming prompt-screenplay into multimodal content includes document parsing, marker extraction, prompt generation, parallel multimodal generation, and final composition. The system uses state-of-the-art models (Sora, Runway, MusicGen, Stable Diffusion) with custom prompting layer.

### 2.9.3. Consistency Enforcement System

To maintain visual consistency between scenes, character bank, location bank, object tracking, and style anchoring are used. When generating each scene, the system extracts K-tags and applies saved parameters.

### 2.10. Quality Assessment Criteria and Metrics

To validate prompt-screenplay effectiveness, the following metrics are used:

#### 2.10.1. Human-Oriented Metrics

- **Readability:** expert and general reader screenplay readability surveys
- **Artistic value:** expert narrative quality assessment
- **Ease of writing:** screenplay creation time and author cognitive load

#### 2.10.2. Machine-Oriented Metrics

- **Parse success rate:** percentage of correctly interpreted markers
- **Generation quality:** automatic quality assessment of generated content (FID, CLIP score)
- **Consistency score:** visual consistency metric for characters and locations between scenes

#### 2.10.3. Legal Metrics

- **Watermark detection rate:** percentage of successful watermark detection
- **Blockchain verification time:** authorship verification speed through blockchain
- **Legal admissibility:** expert assessment of protective mechanism legal significance

### 2.11. Ethical Considerations

Prompt-screenplay development accounts for the following ethical principles: transparency, attribution, privacy, bias mitigation, and fair use.

**Note:** The presented methodology is a theoretical-practical framework. Full technical implementation of all components requires iterative development with participation of screenplay community, AI system developers, and intellectual property lawyers.

## 4. Results

A pilot example demonstrated the feasibility of creating multimodal scenarios using the developed system. The generated screenplays were successfully used for video, audio, and text generation, confirming the model's functionality and protective mechanisms.

### 4.1. Pilot Application: Transformation of "SKOR" Screenplay into Prompt-Screenplay

To validate the developed methodology, the original screenplay of the first episode of science fiction series "SKOR" (Episode 1 "We're Stuck, Mom!") was selected. This screenplay presents particular interest due to its characteristics:

**Vertical format:** 6:20 minute runtime corresponds to short-form drama format

**Visually saturated structure:** complex multi-layered visualization (underground world, virtual reality, futuristic city)

**Subjective perspective:** first-person narration through protagonist's interface

**Techno-aesthetics:** cyberpunk elements, AR/VR interfaces, robots

**Non-linear narrative structure:** episode built around "SKOR" acronym, where each letter unfolds through separate segment

#### 4.2. Analysis of Original Screenplay Structure

##### 4.2.1. Structural Features

The original screenplay already contains elements close to prompt-format:

**Tabular organization:** use of tables with "shot," "description," "plan" columns creates structure convenient for machine processing

**Visual markers:** text color codes action intensity, serving as prototype of our morphological modifiers

**Detailed shot types:** precise indication of shooting type for each frame (medium close-up, extreme long shot, detail, extreme close-up) facilitates interpretation by AI generators

**Internal monologues:** protagonist Anton's thoughts highlighted in special format ("We hear only Anton's thoughts:"), corresponding to [ZG] markers in our system

**Timing:** each segment has precise timing, critical for vertical format

##### 4.2.2. Narrative Structure

The screenplay follows five-act structure with acronym-dramaturgy :

- Beginning (S) — 1:16 min — world introduction, protagonist and his task
- Inciting incident (K) — 1:30 min — conflict with "Omniron" system, rescuing girl
- Complication (I) — 0:85 min — immersion into virtual reality
- Development (E) — 2:05 min — mass consciousness update, explosion
- Cliffhanger (R) — 0:45 min — revelation and elevator stuck

Each segment reveals one letter of "SKOR" title, creating meta-narrative layer.[9,10]

#### 4.3. Transformation into Prompt-Screenplay

##### 4.3.1. Master Prompt Application (Layer 1)

For SKOR screenplay, the following master prompt was formed:

[GENRE: cyberpunk, sci-fi thriller, dystopia]

[SUBGENRE: techno-noir, social fantasy]

[FORM: vertical video, episode 1 of series, 6:20 min]

[STRUCTURE: five-part acronym dramaturgy SKOR]

[TARGET\_AUDIENCE: 18-35 years, tech-savvy audience familiar with cyberpunk]

[MODALITY: linear with interactive elements in future versions]

[KEY\_IDEA: technological consciousness control vs. individual freedom]

[TONALITY: anxious, claustrophobic, with irony elements]

[TEMPO: high, dynamic editing]

[VISUAL\_STYLE: neon palette, contrast dark dungeons and bright AR interfaces, rusty techno]

[COLOR\_PALETTE: dominant black, orange, scarlet, neon blue]

[WORLD\_BIBLE:

- Two-level city: upper (rich) and lower (underground market)
- “Omniron” system: corporation controlling consciousness through VR/AR
- Robot-courier: autonomous agent with mysterious mission
- Protagonist Anton: hacker with AR implants, invisible in system
- Elevator as metaphor for vertical mobility and social stratification]

[CHARACTER\_CONSISTENCY:

- Anton: not shown physically until finale, only POV and thoughts
- Red-haired girl: short haircut, mole on neck, black dress
- Robot-courier: rusty, |0|0| symbol on face, C-shaped container
- Person in yellow cloak: sun-helmet, mask of many faces]

[NODE\_COUNT: 27 main shots]

[AVERAGE\_NODE\_DURATION: 14 seconds]

[IMPLEMENTATION\_TOOLS: text-to-video (RunwayML, Pika), Midjourney for concept art, AI voice for internal monologue, generative music]

This master prompt documents authorial vision and becomes reference point for all subsequent transformations.

#### 4.3.2. Node Transformation (Layer 2)

Each frame of original screenplay was transformed into node with added prompt elements. Example transformation of frame 5 from “Beginning (S)” segment demonstrates the added structural elements including doubled letters for visual accents, K-markers for consistency-critical elements, V-markers for important elements, A-markers for atmospheric accents, Z-commands for fade/pauses, M-commands for musical accompaniment, [LINK] for connected nodes, and [OFF-CHAIN\_LINK] for leitmotifs passing through episode.

##### **Example transformation of frame 5 from “Beginning (S)” segment:**

###### **Original screenplay:**

Below on the storefronts there are many vintage and homemade goods, people pass by in 20th century clothing, but nobody pays attention to the elevator. In the doorway appears and freezes someone in baggy dark clothing with a hood. He holds in his hands an oval container shaped like the letter “C” with a black heart inside.

###### **Prompt-screenplay with markers:**

[NODE\_05] [TIME: 00:18]

[INT] [MS→CU]

Beloww on storefrontss many vintagee and homemadее goodss, K peoplee pass in 20th centuryyy clothingg, but nobody payss attentionn to elevatorr.

A Atmosphere of oblivion and ordinariness of the fantastic.

In doorway appearss and freezess someone in baggyy darkk clothingg with hoodd. K He holds in hands oval containerr shaped like letter “C”

with K black heartt inside. V

[SOUND: underground market hum, indistinct voices, rattling mechanisms]

[VO: "Robot-courier delivers trouble? Interesting..."]

Behind him suddenlly visiblee crimsonn flashh and soundss sirenn. F

On stranger's facee ignites symbolll |0|0|, he throwss off clothingg

and we see rustyy robott with tracess from bulletss and blows. A

He raisess hand before curtain, it disappearss, container with heart hides

in robot's body, and he runss into elevator and presses button. M F

[EM: anxiety, curiosity]

[KEY: robot-courier carries something important related to "C" (first letter SKOR)]

[LINK: NODE\_04 (cause-effect), NODE\_06 (direct continuation)]

[OFF-CHAIN\_LINK: NODE\_23 (heart appears again in final segment)]

#### Decoding of added elements:

- **Letter doubling** (storefrontss, goodss) — visual accents for AI
- **K-markers** — elements critical for consistency (people in 20th century clothing, black heart, robot)
- **V-markers** — important elements for plot understanding
- **A-markers** — atmospheric accents
- **F-command** — fade-out/pauses at action ends
- **M-command** — adding musical accompaniment
- **[LINK]** — indication of connected nodes
- **[OFF-CHAIN\_LINK]** — "heart" leitmotif passes through entire episode

#### 4.4. Metadata Generation for AI

##### 4.4.1. Key Element Weight Coefficients

For each node, weight coefficients of visual elements were calculated. For NODE\_05 (example):

- Robot: 1.8 (main object)
- C-shaped container: 1.7 (plot-important item)
- Black heart: 1.6 (leitmotif)
- |0|0| symbol: 1.5 (robot identifier)
- 20th century clothing: 1.2 (atmospheric detail)
- Storefronts: 0.9 (background)
- Market people: 0.7 (extras)

These coefficients are transmitted in prompt for Stable Diffusion/Midjourney as word weights.

## 5. Discussion

### 5.1. Implications for Creative Industries

The prompt-screenplay framework addresses critical challenges in contemporary creative production by enabling efficient adaptation across media formats while preserving human creative control. The system's ability to generate multiple presentation formats from single source documents significantly reduces production costs and timeline requirements while maintaining artistic integrity.

Integration of authorship protection mechanisms at the conception stage provides unprecedented legal clarity for AI-assisted creative work. This approach addresses industry concerns about intellectual property protection in human-AI collaborative scenarios while enabling creators to leverage advanced AI capabilities without sacrificing creative ownership.

### 5.1.1. Comparison of Screenplay Writing Formats in Generative AI Context

Criterion	Traditional Screenplay	Fountain	Executable Screenplay	Dramatron / AI Co-Writing	Prompt-Screenplay
<b>Primary Purpose</b>	Guide for film crew	Human-readable, tool-independent format	Game engine control	Automatic script/storyboard generation from text	Multimodal generation + authorship documentation
<b>AI Generation Support</b>	✗ No	✗ No	△ Indirect (via engine scripts)	☑ Yes (text, static images only)	☑ Yes (text, video, audio, interactive)
<b>Dynamic Shot</b> (camera movement, duration)	Described in prose	Described in prose	△ Partial (via events)	✗ Only static frames	☑ Explicit timing, shot types, movement markup
<b>Vertical Format</b> (mobile cinema)	Not considered	Not considered	Limited	✗ Not supported	☑ Explicit support (shots, composition, timing)
<b>Machine Readability</b>	Low (PDF/FDX)	Medium (text with syntax)	High (code)	High (prompts + tags)	☑ High (structured tags + JSON/YAML export)
<b>Embedded Authorship Metadata</b>	✗ No	✗ No	✗ No	✗ Only visual tags	☑ Consistency tags, watermark annotations at conception stage
<b>Multimodal Support</b> (audio, text, video, game)	✗ No	✗ No	△ Game only	△ Image only	☑ Complete: video, audio, comic, novel, interactive
<b>Human-AI Co-Creation Flexibility</b>	✗ No	✗ No	△ Engine-limited	△ Partial (one-way generation)	☑ Bidirectional: AI suggests, author controls priorities

<b>Legal Significance</b>	External docs	External docs	Code repository	<input checked="" type="checkbox"/> No	<input checked="" type="checkbox"/> Embedded markers as proof of authorial intent before generation
---------------------------	---------------	---------------	-----------------	--	---

### Key Insights from Comparison

- Existing formats are either **human-centric** (traditional screenplay) or **techno-centric** (executable scripts), but do not account for **hybrid nature of AI-assisted authorship**.
- Approaches like Dramatron make important steps toward automation but **remain in static image paradigm** and do not address copyright issues.
- Prompt-screenplay is the first format** that simultaneously:
  - Preserves artistic expressiveness for humans
  - Provides structured instructions for AI
  - Embeds legally significant metadata **at conception stage**, not post-factum

### 5.2. Technical Innovation and Morphological Modifiers

The morphological modifier system demonstrates effective integration of human creativity with machine interpretation capabilities. The approach of embedding machine instructions within natural language preserves creative workflow while enabling precise AI control.

The letter doubling technique (storefrontss, goodss) provides visual emphasis for AI systems without significantly disrupting human reading comprehension. Consistency markers ([K], [V], [A]) enable precise narrative control while maintaining document readability.

Command markers (F, M, [LINK]) provide technical specifications for generation systems while preserving the natural flow of creative writing. This balance addresses the fundamental challenge of creating documents simultaneously readable by humans and interpretable by machines.

### 5.3. Legal and Authorship Considerations

The prompt-screenplay framework addresses the critical legal gap in AI-generated content authorship by documenting human creative intent at the conception stage. The cryptographic consistency mechanisms provide legally admissible evidence of human authorship that addresses recent copyright office requirements for demonstrable human creative input.

The blockchain registration approach creates permanent, tamper-evident records of creative decisions throughout the collaborative process. This documentation supports legal claims of authorship while enabling the benefits of AI-assisted creative production.

### 5.4. Limitations and Future Research

Current implementation focuses primarily on narrative content with demonstrated effectiveness for the SKOR screenplay transformation. Future research should investigate prompt-screenplay applicability to documentary, experimental, and abstract creative forms.

Technical limitations include dependency on AI model capabilities and potential quality variations across different generation systems. The framework requires ongoing adaptation to emerging AI technologies and generation capabilities.

Legal validation remains primarily theoretical, requiring comprehensive testing through actual copyright disputes and industry implementation. Cross-cultural applicability needs investigation for international deployment across diverse legal and creative contexts.

### 5.5. Democratization and Industry Impact

The prompt-screenplay system significantly lowers barriers to professional content production. Independent creators without access to traditional film crews and budgets can write prompt-screenplays, generate content through AI systems, automatically protect copyright through blockchain registration, and distribute content on digital platforms.

This democratization is particularly relevant for emerging markets where traditional film industry access is limited. The framework enables new business models including screenplay-as-a-service, licensing across multiple studios, generation royalties, and NFT tokenization of creative assets.

## 6. Conclusion

This study presents the prompt-screenplay framework as a novel approach to human-AI creative collaboration that addresses critical challenges in authorship protection, multimodal content generation, and adaptive media presentation. The five-layer semiotic architecture successfully integrates human creative expression with machine-readable instructions while maintaining narrative coherence and artistic integrity.

Key contributions include: (1) a hybrid format enabling simultaneous human creativity and AI interpretation, (2) morphological modifier systems preserving natural language readability while encoding machine instructions, (3) cryptographic consistency mechanisms providing robust authorship protection, and (4) empirical validation through the SKOR screenplay transformation demonstrating practical applicability across media formats.

The SKOR transformation demonstrates the framework's effectiveness in converting traditional screenplays to prompt-screenplay format while maintaining narrative consistency and enabling multi-format generation. The morphological modifier system successfully embeds machine instructions within natural language without disrupting human comprehension.

Legal analysis confirms that embedded authorship protection mechanisms address current copyright requirements for demonstrable human creative input. The framework provides technical solutions to the legal challenges of AI-assisted creative work while preserving human creative agency.

Future research should focus on expanded domain applications, comprehensive legal validation through industry partnerships, and integration with emerging multimodal AI technologies. The prompt-screenplay framework provides foundation for next-generation creative tools that preserve human artistic vision while maximizing the benefits of human-AI collaboration.

The implications extend beyond creative industries to any domain requiring collaborative human-AI work with clear attribution requirements. As AI capabilities continue advancing, frameworks like prompt-screenplay become increasingly critical for ensuring technology serves human creative potential while protecting intellectual property rights and creative ownership.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable for theoretical framework and format development research.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The "SKOR" screenplay transformation examples and technical specifications are available upon request. Generated content samples will be made available in compliance with copyright and privacy requirements.

**Acknowledgments:** The author acknowledges the creative community whose feedback and collaboration informed the development of the prompt-screenplay format, and the researchers whose work on human-AI collaboration provided the theoretical foundation for this innovation.

**Conflicts of Interest:** The author declares no conflicts of interest.

## References

1. Kong, B. H. (2024). Research on the publication utilization of prompt engineering technology. Han'gug chulpanhag yeon'gu. <https://doi.org/10.21732/skps.2024.117.5>
2. He, Y., Liu, Z., Chen, J., et al. (2024). LLMs Meet Multimodal Generation and Editing: A Survey. arXiv preprint arXiv:2405.19334.
3. Lin, C. (2025). Prompt Engineering as Mediation: Investigating AI Chatbot-Assisted Writing Process From an Activity Theory Perspective. IEEE ICAIE. <https://doi.org/10.1109/icaie64856.2025.11158182>
4. Yeh, C., Ramos, G., Ng, R. S., et al. (2024). GhostWriter: Augmenting Collaborative Human-AI Writing Experiences Through Personalization and Agency. arXiv preprint arXiv:2402.08855.
5. Hosanagar, K., & Ahn, D. (2024). Designing Human and Generative AI Collaboration. arXiv preprint arXiv:2412.14199.
6. Esteban, A. P. (2025). Creative CRAFT: A structured framework for creativity-driven prompt engineering in generative AI. *International Journal of Innovative Research and Scientific Studies*, 8(5). <https://doi.org/10.53894/ijirss.v8i5.9229>
7. Chen, M., Rau, P. P., & Ma, L. (2025). LLM Asks, You Write: Enhancing Human-AI Collaborative Writing Experience through Flipped Interaction. *AHFE International*. <https://doi.org/10.54941/ahfe1006247>
8. Ducru, P., Raiman, J., Lemos, R., et al. (2024). AI Royalties – an IP Framework to Compensate Artists & IP Holders for AI-Generated Content. arXiv preprint arXiv:2406.11857.
9. U.S. Copyright Office. (2022). Copyright Registration Guidance: Works Produced by a Machine or Mere Mechanical Process. *Federal Register*, 87(37), 10388-10391.
10. Yang, J., & Zhang, H. (2024). Development And Challenges of Generative Artificial Intelligence in Education and Art. *Highlights in Science Engineering and Technology*. <https://doi.org/10.54097/vaeav407>
11. De Filippo, A., & Milano, M. (2024). Large Language Models for Human-AI Co-Creation of Robotic Dance Performances. *IJCAI 2024*. <https://doi.org/10.24963/ijcai.2024/844>
12. Ranade, N., Saravia, M., & Johri, A. (2024). Using rhetorical strategies to design prompts: a human-in-the-loop approach to make AI useful. *AI & Society*. <https://doi.org/10.1007/s00146-024-01905-3>
13. Luther, T., Kimmerle, J., & Creß, U. (2024). Teaming up with an AI: Exploring human-AI collaboration in a writing scenario with ChatGPT. *OSF Preprints*. <https://doi.org/10.31219/osf.io/extmc>
14. Böhm, K., & Durst, S. (2025). Collaborative GenAI – Humanized Interaction Fields for Knowledge Creation. *Proceedings of the European Conference on Knowledge Management*, 26(1). <https://doi.org/10.34190/eckm.26.1.3607>
15. Wu, D., Wei, X., Chen, G., et al. (2025). Generative Multi-Agent Collaboration in Embodied AI: A Systematic Review. *Proceedings of the International Joint Conference on Artificial Intelligence*. <https://doi.org/10.24963/ijcai.2025/1190>
16. Soni, G., Waris, N., & Dalai, A. (2025). Generative AI and Creativity: Enhancing Human Creativity Across Visual Arts, Content Creation, Music, Design, and Education. *International Journal for Science Technology and Engineering*. <https://doi.org/10.22214/ijraset.2025.74042>
17. Odhiambo, J.W., & Ondimu, K. (2025). A Framework for Ethical AI-Generated Content Governance. *Preprints*. <https://doi.org/10.20944/preprints202509.0271.v1>
18. Nithiya, C., Revathy, G., Menaha, R., et al. (2025). Decentralized Digital Media Generation. *Advances in Computational Intelligence and Robotics Book Series*. <https://doi.org/10.4018/979-8-3373-1504-1.ch005>
19. Lamprou, S., Dekoulou, P., & Kalliris, G. (2025). The Critical Impact and Socio-Ethical Implications of AI on Content Generation Practices in Media Organizations. *Societies*, 15(8), 214. <https://doi.org/10.3390/soc15080214>
20. Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly*, 28(1), 75-105.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.