Computational Tools for Table-Top Role-Playing Games: A Scoping Review

Fiona Shyne
Khoury College of Computer Sciences
Northeastern University
Boston, MA, USA
shyne.f@northeastern.edu

Seth Cooper
Khoury College of Computer Sciences
Northeastern University
Boston, MA, USA
se.cooper@northeastern.edu

Abstract

Table-top role-playing games (TTRPGs) are a form of gameplay that often requires a variety of complex tasks to be completed both in preparation and throughout gameplay: from tracking game state to the creation of fictional worlds. This has presented an opportunity for computational assistance in TTRPG sessions, both in the creation of artifacts and throughout the gameplay. We investigate the current research in computational tools for TTRPGs through a scoping review of academic works and present the major trends and opportunities from these works. We screened over one thousand works sourced from three different academic databases: ACM Digital Library, IEEE-Xplore, and Google Scholar. Papers were included based on relevance to TTRPGs, computational interface, and academic venue. In total, we evaluated 46 works in terms of produced artifacts, computational methods, evaluation, and outcomes. These papers include a diverse set of produced artifacts and computational methods, with an emphasis on tangible interfaces and generative AI systems. However, we found an opportunity for future work in terms of long-term studies, mixed-initiative methods, and different aspects of gameplay.

CCS Concepts

• **Human-centered computing** → HCI design and evaluation methods; *User interface design*; Participatory design; Information visualization; • **Computing methodologies** → *Natural language processing*; *Machine learning approaches*; Genetic algorithms; • **Software and its engineering** → **Interactive games**.

Keywords

TTRPGs, Literature Review, Procedural Content Generation, Tangible Interface

ACM Reference Format:
Fiona Shyne and Seth Cooper. 2025. Computational Tools for Table-Top Role-Playing Games: A Scoping Review. In *International Conference on the Foundations of Digital Games (FDG '25), April 15–18, 2025, Graz, Austria.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3723498.3723816

FDG '25, Graz, Austria

ACM ISBN 979-8-4007-1856-4/25/04
https://doi.org/10.1145/3723498.3723816

1 Introduction

Table-top role-playing games (TTRPGs) are a popular variety of table-top games, where one player, the Game Manager [1] (GM), runs the game for other players who take on the "role" of a character. GMs are also often, at least partially, in charge of game design in terms of world-building and designing challenges. TTRPGs often run through multiple game playing sessions and follow one storyline throughout a campaign. TTRPGs encompass a wide variety of gameplay, from storytelling to combat and puzzle-solving. The most popular system for playing TTRPGs, Dungeons and Dragons (D&D), brings players into a high-fantasy setting to tackle quests and fight fantasy monsters.

The continued popularity of TTRPGs in the modern age has come alongside an increase in digital opportunities for gameplay. The role of a GM can be a taxing endeavor, with them being given the task of world-building, game design, and information tracking. Computational tools have the ability to assist GMs and players both in planning a campaign and during gameplay.

This work investigates the current academic work in creating tools to assist GMs and players for TTRPGs. Through this review we attempted to address the following questions.

- RQ1: What do academic computation support tools for Table-Top Role-Playing Games look like in terms of computational techniques, study design, types of assistance, outcomes, and limitations?
- RQ2: What areas in this domain remain unexplored?

We systematically identified 46 academic works related to computational tools for TTRPGs, and reported our findings using the PRISMA reporting guidelines for scoping reviews [2]. We examined these papers in terms of produced artifacts, computational methods, evaluations, and outcomes. After examining these papers, we identified major opportunities for future work in this area.

2 Related Work

2.1 Table-Top Role-Playing Games

Table-Top Role-Playing Games (TTRPGs) have their origins in the 1970s, with the first edition of the high fantasy game Dungeons and Dragons [26]. TTRPGs remain a popular form of gameplay and can take on a variety of forms and settings. TTRPG gameplay is unique in a couple of different ways. First, a single game most often takes place over multiple sessions and can take months to years to complete. There are, however, exceptions to this in terms of "one-shots" that take place over one or just a couple of sessions.

[1] Also called the game master.
[2] https://www.prisma-statement.org/scoping

Second, the game rules, called a TTRPG system, often only lay out a variety of ways the game can be played and not specific instructions for how a game looks. While pre-made modules exist, often the task for world-building and game design is at least partially the responsibility of one player called the Game Manager (GM). The GM must prepare for the game, either by familiarizing themselves with pre-authored content or by generating the worlds, non-player characters, situations, and challenges the players encounter (GMs often use a combination of existing and self-authored content). These challenges include navigating social encounters, solving puzzles, or fighting in combat. The system can guide how gameplay looks for this, for example in D&D there is a monster manual [41] that provides a set of monsters GMs can select from to design combat. The GM also runs the game for the other players: narrating the environment and determining the effect of players' actions on the world. Lastly, TTRPGs are unique in that the players will embody the characters in the game through role-play. The TTRPG system can also provide support for the role-play aspect, for example in D&D players attempt to perform actions using a dice roll and adding a skill modifier.

Guzdial et al. [25] argues that TTRPG systems are a form of procedural content generation, and that computer PCG systems can model themselves after TTRPGs. However, the large amount of complex rules and content needed to play a TTRPG provides an opportunity for computational support. This is especially true in a post-COVID time, when many TTRPG campaigns are being played in online or hybrid formats. Several commercial venues provide some type of computational support for TTRPG gameplay (discussed below). Tang et al. [54] interviewed GMs to identify the requirements for computational support and found that the most important features were inspiration, controllability, and ease of use. Our work examines the current academic research into computational tools for TTRPGs.

2.2 Non-Academic Computational Support for TTRPGs

While this work focuses on academic research, there are many publicly available tools to help support TTRPG gameplay. Wizards of the Coast, which owns the Dungeons and Dragons system, hosts D&D Beyond [3], a website where players and GMs can manage character sheets, combat encounters, and homebrewed content including spells, monsters, and items. Several websites, including Roll20 [4] and Foundry [5], host virtual tabletops where parties can play TTRPGs completely remotely. Similarly, Discord extensions such as Avrae [6] provides tools for playing DnD over Discord. Many tools are available for generating content for TTRPGs, including creating/generating maps [7], encounters [8], non-player character stat blocks [9], names [10], and shops [11]. The website Donjon [12] includes generators for all of the above along with information about game rules. Websites like Start Playing [13] and Dungeon Masters Direct [14] allow players to find games led by professional GMs.

[3] https://www.dndbeyond.com/
[4] https://roll20.net/
[5] https://foundryvtt.com/
[6] https://avrae.io/
[7] https://inkarnate.com/maps
[8] https://tools.goblinist.com/5enc
[9] https://rpgtinker.com/
[10] https://www.fantasynamegenerators.com/
[11] https://5emagic.shop/
[12] https://donjon.bin.sh/

2.3 Related Reviews of Literature

Several previous works have examined literature related to computation and games. Informal review approaches were used to review the literature related to procedural puzzle generation [16] and procedural content generation for games [62]. Scoping reviews using a methodology similar to ours have been done to examine literature related to Large Language Models (LLMs) for games [60] and procedural generation of dungeons [56]. A systematic review was done to examine game design tools [23]. While we are the first work to perform a scoping review of computational tools for TTRPGs, previous work has done a scoping review of TTRPGs as a means for psychological interventions [61].

3 Methodology

This scoping review was reported following the Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR).

3.1 Definitions and Edibility

TTRPGs do not have a clear, universal definition, and there is a gray area between what can be considered a TTRPG and other forms of gameplay. For this paper a TTRPG is defined by the following traits:

(1) **Game Manager (GM)**: one (or more) human players is in charge of running the session for the other players in the game.
(2) **Role-playing**: players embody characters via spoken or written communication during at least part of the gameplay.
(3) **Homebrewing**: The system allows for custom game content and storytelling to be developed by the GM, even if there are also pre-written modules available.

Many other elements could appear in TTRPGs, including combat, puzzle-solving, and multi-player collaboration. We also included some papers where a computational system or a pre-written module replaced the GM, as we deemed these tools as something that a human GM could use as well. Using this definition, we developed the following inclusion criteria for papers:

(1) Must directly mention applications to TTRPGs generally or to a TTRPG system (e.g. D&D).
(2) Must have a computational interface that is used either offline (in preparation for the session) or online (during the campaign).
(3) Must be published in an academic venue. Workshop papers can be included, but theses of any kind will be excluded.
(4) We must be able to retrieve the full text of the paper, and it must be available in English.

Computational	Tools	TTRPGs
procedural automatic computational	tool generation assistant	tabletop role-play table top role play table-top role-play tabletop roleplay TTRPG TTRPGS TRPG TRPGS Dungeons and Dragons Dungeons & Dragons

Table 1: Key Terms Used in search. The acronym "DnD" (along with "D&D") was removed as it is commonly used in other fields, such as "Drag and drop" in HCI. The full query text is provided in Appendix B.

3.2 Process

We searched on the ACM Digital Library [15], IEEE Xplore [16], and Google Scholar [17]. The search string contained terms of three categories: computation, tools, and TTRPGs. The terms within each category were joined with the "OR" keyword, and the three sets of keywords were joined with the "AND" keyword. The keywords are given in Table 1. The search terms were identical across the three databases. The databases were queried at the beginning of September 2024. There was no restriction on publication date.

For each query, the resulting citation information was extracted and put in to the literature review software "Catchii" [28]. This software automatically detected potential duplicate citations, which were reviewed manually. Additional duplicate papers were identified later on in the process. All non-duplicate papers were screened (using single screening) first using only the title and abstracts to identify relevance. The papers that were deemed relevant, were then reviewed for eligibility based on the full text of the paper. The reason for exclusion was recorded for each paper in this step.

The resulting papers were read in detail to extract the following information:

(1) **Artifact**: What element or interface was created that can be used in a TTRPG setting?
(2) **System**: What (if any) TTRPG System is this tool designed for or restricted to?
(3) **Method**: What process was used to produce this artifact?
(4) **Evaluation**: How was the artifact tested for viability?
(5) **Outcomes**: What were the major findings of this work?
(6) **Limitations**: What were the major limitations of this work?

The data extraction was also done by one person (the first author). While the papers were being read, the first author wrote a description for each data item in a spreadsheet. After all papers were read, an open coding process was used for the artifact, method, and evaluation to label each data item. For papers where multiple codes could apply, the most relevant code was selected. Outcomes and limitations did not go through a formal coding process, but the major themes are discussed in the results section. This spreadsheet is available on the OSF page.

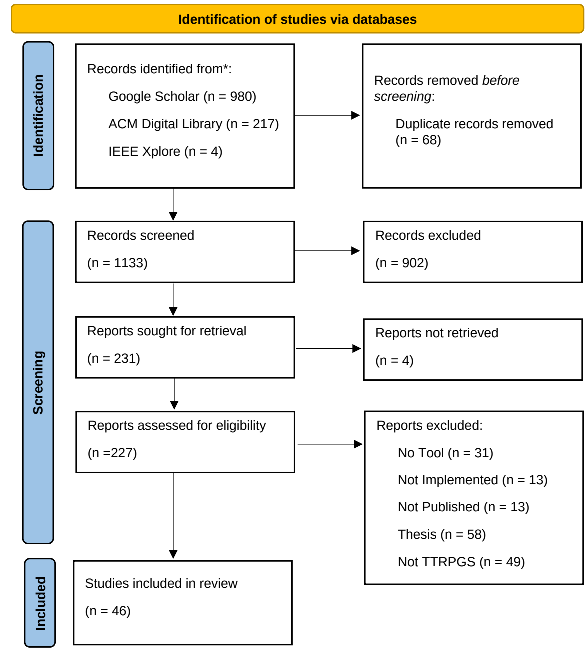

Figure 1: Flow diagram for the collection and screening of papers in this review.

4 Results

4.1 Papers retrieved

The keyword search retrieved 217 papers from ACM, 4 papers from IEEE, and 4,940 papers from Google Scholar. However, only the first 980 papers from Google Scholar could be retrieved, due to the server not returning pages after a certain number. This limitation was replicated on multiple instances, even months apart, and on different computers. However, relevant papers were scarce in the later pages we were able to retrieve, meaning it is unlikely that a substantial amount of relevant papers would be found on the pages we were unable to retrieve. From these, 68 papers were flagged as duplicates by Catchii, which was confirmed manually. Out of the 1,113 papers screened using titles and abstracts, 231 were deemed eligible for full-text reviews. From these papers, 58 were excluded for being a thesis, 49 were not related to TTRPGs, 44 were excluded for not having a tool or not implementing a proposed tool, 13 could not be confirmed to be published, and 4 papers were not able to be accessed. This resulted in 46 papers that are included in this review. A flow diagram is given in Figure 1. A table of all papers in this review is given in Table 2. All data created from this process is available on the open science framework [18].

[13] https://startplaying.games/
[14] https://dungeonmasterdirect.com/
[15] https://dl.acm.org/
[16] https://ieeexplore.ieee.org/Xplore/home.jsp
[17] https://scholar.google.com/
[18] https://osf.io/5fgku/

4.2 Metadata

The year and venue for the included papers is shown in Figure 2. The oldest paper was published in 2008, and the newest paper was published in the same year as this review (2024). There appears to be a slight increase in the amount of papers published, starting around 2020 and continuing until the present day (noting that the search was conducted in September 2024).

The papers were published in a variety of venues. The most popular conferences were artificial intelligence conferences that focused on games (13), including AIIDE: Artificial Intelligence and Interactive Digital Entertainment (7), the ICAART: International Conference on Agents and Artificial Intelligence (2), the Games and NLP workshop (2), and ICC: the International Conference on Computational Creativity (2). The next most common were general game conferences (11) such as FDG: the Foundations of Digital Games (6), COG: Conference on Games (2), SBGames: the Brazilian Symposium on Computer Games and Digital Entertainment (2), and DIGRA: Digital Games Research Association (1). Human-computer interaction conferences were also represented (6), such as CHI: the ACM conference on Human Factors in Computing Systems (3), CHI-Play: the Annual Symposium on Computer-Human Interaction in Play (2), and MHCI: Proceedings of the ACM on Human-Computer Interaction (1).

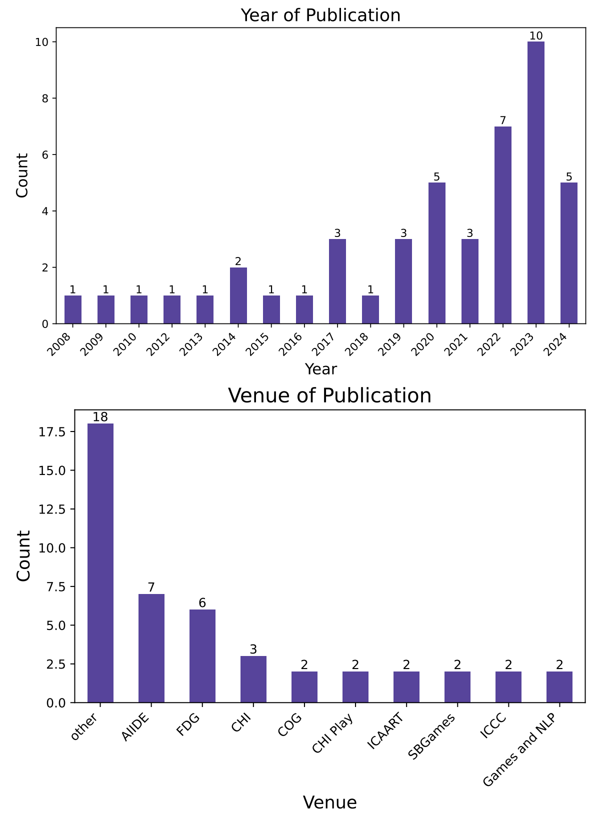

Figure 2: A histogram of the year (top) and venue (bottom) for which the selected papers were published. The venue was labeled as "other" for all venues for which only one paper was included.

4.3 Artifacts

The codes for artifacts are shown in Figure 3. The most common type of artifacts were tangible interfaces (10). These included digital game boards [29, 40, 48, 50] and systems to sense tangible figures on touch screens [39, 51]. These artifacts were designed to allow players to connect to the game on a physical level (which is part of the appeal of playing tabletop games), while their computational elements can assist with data tracking (such as a player's current stats) or visualization (for example showing the possible range of movement). The system proposed in Stemasov et al. [50] enabled new forms of game-play through crafting and distraction. New forms of gameplay were also proposed in the WEARPG series of publications [8–10], which introduced an electronic wristband that enabled motion-based mini-games in place of random skill checks. One paper [52] created a robotic bird capable of engaging audiences during Twitch streams, for the digital audience to feel more engaged in the physical environment of the streamers.

Another popular type of artifact was artifacts related to narrative either from generating narrative text (4) or through visualization of narratives (3). An issue identified in the selected papers, is the amount of dense narrative information GMs must sort through to run TTRPG sessions. To assist with this issue, interactive flowcharts of narrative events [1–3], information highlighting [36], and summarization tools [65] were created. Computational tools were also used to create text elements that are thematically relevant [47, 63], answers to brainstorming prompts [65], or suggestions for character dialog [36].

Five papers related to the creation of maps and dungeons. Three of these papers produced structured dungeons for TTRPGs. These dungeons were created based on the flow of the mission [17], user-specified constraints [6], or the physical limitations of dungeon tiles [24]. Two papers looked at open-ended environments. One produced tile maps based on user-created constraints [11], and the other generated opened worlds based on text prompts [15].

Five papers presented domain-specific languages (given the "language" code) that can be used to assist in different aspects of TTRPGs. These artifacts provide programming languages that are specific to particular aspects of TTRPGs, often significantly more restricted than traditional programming languages. This includes a language that models dice rolls and calculates probability distributions [42], languages to generate textual artifacts such as character or item descriptions [31–33], and one to create social simulations [53]. All these languages aim to present an interface that allows the expressiveness of natural language but with less technical expertise necessary to program than with traditional programming languages.

Four papers presented works that replaced a human-GM with a computational system. The majority of these systems had a similar structure, where players describe actions in natural language and the computer GM describes an outcome [4, 27, 55]. The remaining paper took the approach of creating a game flow state with goals and methods that players can choose from to approach those goals [38]. Notably, all the works that replace a human-GM focus exclusively

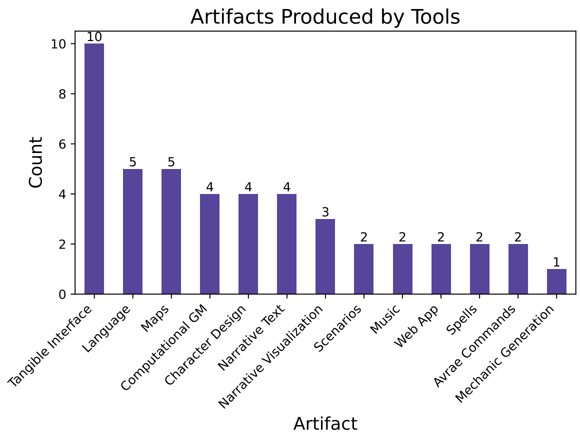

Figure 3: Artifacts produced by the computational Tools

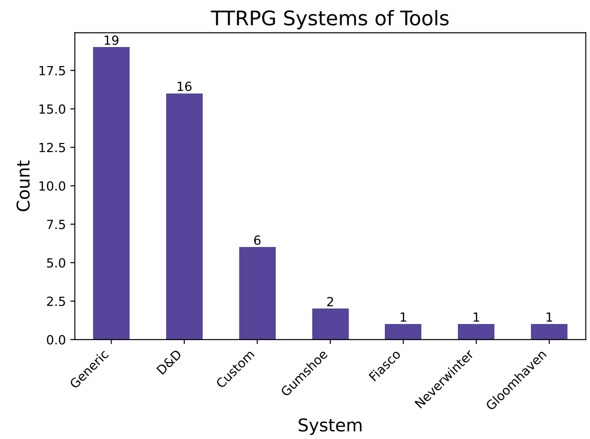

Figure 4: The system codes for included papers

at the narrative aspect of gameplay and do not consider other components such as puzzles or combat.

Four papers related to various aspects of character design within TTRPGs. This includes the generation of stat blocks and character traits [19, 58], converting character data between TTRPG systems [22], and generating image sheets for a character [57].

Several works focused on generating specific aspects of TTRPGs. Two papers looked at background music recommendation [45] and generation [21], based on the emotional tones of the session. Two papers looked at generating spells for D&D [43, 44]. Two works looked at generating commands for the D&D Discord bot "Avrae" [46, 64], to relieve the burden on players to type the commands during gameplay. Two papers aid in the starting of a TTRPG game by generating initial scenarios either in the form of quests [35], or environment setups [30]. One work generated new encounters using a digitized version of an encounter lookup table [5]. One work looked at generating new mechanics for role-playing games [66]. One work created a digitized version of the character sheet for D&D [12].

Overall the artifacts present solutions to several major problems that occur in TTRPGs. Several artifacts address the information overload that can occur in TTRPGs, through visualization, data tracking, or digitalizing paper artifacts. Other artifacts address the burden of content creation through tools to generate narrative text, maps, character information, and more. Similarly, the language artifacts give tools for designers to create their own generators. Several of the artifacts don't seek to address problems, but allow new forms of gameplay including movement based mini-games, gameplay though physical creation, and playing with a completely virtual GM.

4.4 Systems

The system codes for all papers are shown in Figure 4. The most common code was "generic" which represented tools that were not specific to a particular TTRPG system. In total there were 19 papers that did not specify a system. Some of these works included other restrictions such as only being applicable to games live streamed on Twitch [52], or only producing artifacts in the high fantasy style of D&D [57]. Further, one of these papers [21] evaluated their system using only D&D, even though it could be applied to many systems.

The second most popular code, with 16 papers, was tools specific to D&D. This is not particularly surprising as D&D is the most popular system overall, and the only one included by name in our search terms. Most often these papers were specific to the 5th edition of D&D (D&D 5e [14]), or did not specify the edition but were written in a year when D&D 5e was the most popular edition (roughly 2015 to 2024). Only one paper mentions an earlier edition [22], but only in the context of converting between D&D 5e and other editions.

Six of the papers used a system that was custom to the study. These custom systems were designed to accommodate new modes of play [8–10, 50] or to present a simplified version of a system [40, 58]. All but the WEARPG [8–10] custom system, are based in some capacity on D&D. All these custom systems are drastically simplified when compared to D&D 5e. In one case [58], it does not even represent a playable system.

Five papers used an established system different from D&D. One paper [1] used the Gumshoe System, which is a narrative-based detective TTRPG. Two papers used "GM-less" TTRPG systems, like Fiasco [30], and Gloomhaven [24]. One paper used Neverwinter [38], which is a digital roleplaying game. In these last two cases the tools produced can still be used in a TTRPG setting, and so they were still included.

These papers show a clear bias towards the D&D system. Although we did include this in our search terms, even the papers that were system-agonistic or developed custom systems, often incorporate D&D elements or themes. While some papers used smaller systems, many popular systems were not represented at all.

4.5 Computational Methods

The codes for the methods are shown in Figure 5. Fourteen works predominantly used software or hardware engineering techniques to produce the tools. Of these, five papers used only software engineering, including mobile app development [12], interactive visualization [2, 3], wep app development [5], and programming language development [42]. Nine works used various hardware engineering techniques. Four papers used markers or computer vision to sense and locate tangible objects, such as the miniatures used in TTRPGS to represent player characters [29, 39, 40, 50]. Two works used robotics and motors in their tools [39, 52]. The WEARPG series of papers used sensors to detect various motion gestures [8–10]. One paper looked at the application of laser cutters and 3D printing pens for TTRPG gameplay [50].

Thirteen papers used some form of "generative AI". The majority (11) of generative AI papers used a Large Language Model (LLM). These papers most often used GPT models including GPT-2 [4, 21, 43, 44], GPT-3 [36, 46, 64, 65], ChatGPT (which given the year of publication was most likely using GPT-3) [55], and GPT-Neo [47]. However, Gemini [27], Llama-2 [43], and OPT [43] were also used. Several techniques were used to improve the quality of the models. Seven papers fine-tuned the LLMs on additional data [4, 21, 43, 44, 46, 47, 64]. Three papers prompted the LLM model with a representation of the game state [27, 36, 65]. Only one paper used prompt engineering alone [55]. Two papers used text-to-image generative AI, using the Stable Diffusion model [15, 57]. One paper used the unmodified model to generate world images [15], while the other fine-tuned the model with example D&D style character images [57].

Nineteen papers used some form of more "classical AI" techniques. Out of these, the most common technique was constraint solving with eight papers. These papers used Answer Set Programming (ASP) [11, 53, 66], the Prolog Language [1, 30], a "Prolog-like" language they developed [33], Boolean satisfiability (SAT) solvers [31], and Satisfiability Modulo Theories (SMT) solvers [32]. Five papers used some form of machine learning including K-Nearest Neighbor [58], naive Bayes [58], support vector machines (SVM) [22, 63], regression [22], conditional generative adversarial networks (C-GANs) [51], convolutional neural networks (CNNs) [63], and bidirectional neural networks [63]. Additionally, while they predominantly used LLMs, Newman and Liu [44] also used other kinds of language models including long short-term memory (LSTM) networks and k-gram models. Three papers used genetic algorithms (GAs). One paper used a GA with a static fitness [24], while the others included user input either at the beginning [6] or by iteratively collecting user feedback to update fitness [19]. Two papers used game-playing algorithms, including Monte-Carlo Tree Search (MCTS) [35] and belief–desire–intention (BDI) agents [38]. One paper used a grammar, implementing both a graph grammar and a shape grammar [17].

While the most common technique used Large Language Models, the papers in this study still represented a range of computational techniques. Software and hardware engineering methods were well represented, along with "traditional" AI approaches such as constraint solving or genetic algorithms.

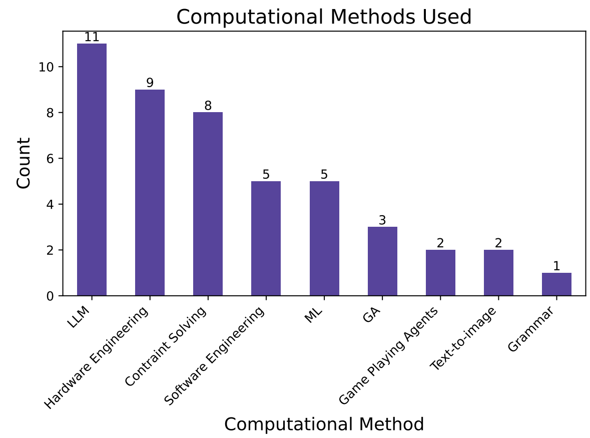

Figure 5: Computational methods code for included papers

4.6 Evaluation

The evaluation codes are shown in Figure 6. Fifteen papers only conducted an informal evaluation of their work. Most often (13 papers) this was a presentation of a "proof of concept" or "prototype." This meant the authors described their interpretation of the artifacts produced by their tool and often included an example for the audience to evaluate themselves. In the other 3 papers [32, 38, 53], an implied play-test or human evaluation of the artifact was conducted, but no procedure or data were presented in the paper. Similar to the proof of concept evaluation, the authors would give their general impression of the outcomes of this user test.

Six papers conducted only non-human based evaluation. Three of these papers used automated tests, including reporting accuracy measures on test data [47, 51] and analyzing the numeric values of generated artifacts [58]. Three papers conducted an evaluation based on comparison either to a framework [11] or to a previously implemented system [33, 39].

Eighteen papers conducted only human-based evaluation. One of these papers included only design interviews with 11 participants [52] and no evaluation of the created artifact. Five studies included both a preliminary user study to guide the design process and a follow-up user study to evaluate the created artifact [9, 10, 15, 50, 65]. The preliminary design studies ranged from 4 to 68 participants (median 16), with all but one of them involving qualitative interviews with participants, with experience in TTRPGs or related tasks. The follow-up studies, in all but one study, collected both qualitative and quantitative (mixed-methods) data. The follow-up studies ranged from 12 to 71 (median of 15) participants. Nine papers only did a user study to evaluate the created artifact. Two of these were mixed-method studies [4, 48] with 8 and 30 participants. Five papers only collected qualitative data [2, 3, 8, 12, 29], with between 5 and 19 participants (median of 8). Three papers collected only quantitative data [19, 22, 55], with 7, 14, and 28 participants.

Six papers included both an automatic and a human evaluation (the "mixed" code). The automatic evaluation in these papers includes text similarity metrics such as BertScore [43, 44, 57, 64], accuracy measures [45, 64], and comparison between baselines [21, 45]. For the human evaluation, four papers included a quantitative analysis [21, 43, 45, 57], one paper used a mixed-methods analysis [64] and one paper used a qualitative only analysis [44].

While some form of human evaluation was common amongst the included papers, many papers included only informal methods of evaluating work. In particular, it was common for papers to evaluate their ideas with a proof of concept that was presented

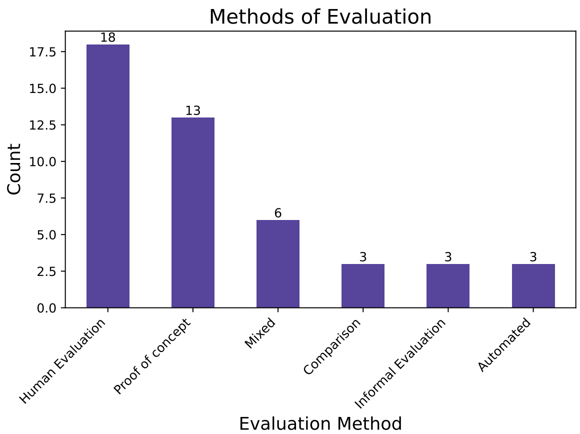

Figure 6: Evaluation codes for the included papers

to readers. Very few papers only included a automated form of evaluation.

4.7 Interactions

We also wanted to examine the intersection between the different data items extracted from the papers. In particular, this was done to identify current gaps in research. Heatmaps of these interactions are provided in Figure 7.

Looking at the interaction between artifact and evaluation, we can notice which artifacts have been evaluated with various methods. For example, nearly all of the tangible artifacts were evaluated with a human evaluation, whereas none of the domain-specific languages had any kind of human evaluation. We can also look at the intersection between artifacts and TTRPG systems. Tangible interfaces were most likely to have a custom interface, although many of them were generic as well. Languages, maps, mechanics, and music were all likely to be generic to the system; while narrative text, narrative visualization, and spells were all specific to a system (most often D&D).

Looking at the intersection between evaluation and computational methods, we can notice which methods have the strongest evaluation. The constraint-solving, game-playing, and grammar methods were only evaluated with informal evaluations and did not have any human evaluations. It should be noted that the system created by Carpenter et al. [11] was later evaluated using human evaluation in a paper that was excluded from this study as it did not implement a new tool [13]. Large language models had the broadest variety of evaluation methods, including informal, automated, human, and mixed evaluation methods.

Finally, we can look at the trends for evaluation methods by publication venue. The general game conferences (FDG, CoG, SbGames) published papers with a wide variety of methods from constraint solving, genetic algorithms, LLMs, ML, game-playing, and grammars. The artificial intelligence conferences (AIIDE, Games and NLP, ICAART, ICCC) contained papers using constraint solving, LLMs, ML, and text-to-image techniques. The HCI-focused (CHI, CHI-Play) conferences included hardware engineering, LLMs, and ML techniques.

We can notice several trends through these interactions. Tangible interfaces, which often were designed through hardware engineering techniques, were most often to be published in HCI specific conferences and include human evaluation. More "classical" computational techniques, such as domain specific languages, constraint solving, game-playing, and grammars, were more likely to be published in general game or AI conferences and include informal evaluations. Modern techniques, particularly LLMs, were published in a variety of conferences and used a variety of evaluation techniques.

4.8 Outcomes and Limitations

Several trends in outcomes can be found in the included papers. The most universal outcome is the willingness and excitement of GMs and TTRPG players to use computational tools in their games. Every paper in this survey that asked players about their perception of computational support generally, or about the specific computational tools, reported a positive reception to this concept, even when also noting the limitations of current implementations. A particular opportunity was identified in this survey for computational support for the tracking and visualization of game data. TTRPGs, especially complex systems like D&D, have a large cognitive burden on players and GMs in terms of the amount of information they need to keep track of. The GMs and players involved in the human evaluations across the studies in this survey appear particularly excited about the opportunity of computational tools to assist in this aspect in game play.

Another trend in the outcomes of these papers was the desire for tangible interactions among GMs and TTRPG players. Even when utilizing digital equipment, players remained attached to the physical components of the game, including dice and miniatures. Further, players demonstrated excitement about exploring new forms of game-play, including gesture/movement-based [8–10] and crafting elements of game-play [50]. It is important to remember this continued desire for tangible elements of gameplay, even with the increase in digital opportunities.

There were also several trends to the limitations presented in the included papers. A major theme that came across in the included papers was the ease-of-use vs. controllability spectrum. On one end of this spectrum, there are domain-specific language artifacts, where the end-user has a very high degree of control over the output of the system but must manually author all content and rules for combining content. On the other end of the spectrum, there are computational GM artifacts, which are very easy for users to use but offer no control over the direction of gameplay. Nearly all the papers presented in this study fall somewhere on this spectrum, where the easier the system is to use the less control is given to the user. This proposes a major challenge to computational systems as Tang et al. [54] highlights ease of use and controllability as vital in tools to support GMs.

A large quantity of papers included very little or only informal evaluation of their work. Even the works that did include some form of human evaluation are limited in scope. None of the papers included in this survey evaluated their system in a long-term campaign that lasted more than one session. While there exist one-shot campaigns that last only one or a couple of sessions, TTRPGS are often played over many sessions. It is also interesting the study was often designed in a one-shot fashion, but didn't mention this as the intent or motivation. Additionally, none of the included studies evaluated the use of a computational tool through the complete process of designing a session: from the planning pre-session through playing the session. In fact, very few studies looked at the processes GMs use to design campaigns.

Finally, there are limitations specific to the computational methods implemented. Any kind of machine learning technique (including generative AI) requires a large amount of training data. This limits the type of content that can be produced to what data can be retrieved. In particular, these techniques are only well suited for popular systems such as D&D (where a large amount of data exists), and ignore smaller systems that may not have the same volume of data. There are particular concerns when it comes to generative AI techniques. These strategies often create false or nonsensical outputs [34], commonly referred to as "hallucinations." There are also ethical concerns with regard to the sourcing of data [37] and the environmental cost of these models [18, 59]. These large models also can perpetrate bias; for example, in Weerasundara and De Silva [57] the generated images of women were extremely sexualized, a fact that was not acknowledged by the authors of the paper.

There is also a concern for the accessibility of some of the proposed tools in the papers. The first concern is the cost and size of equipment needed to use the tools; for example, the tool proposed in Stemasov et al. [50] requires users to have access to a laser cutter. There is also an issue with the amount of technical expertise needed to use some of the tools presented in this survey. This is particularly evident in the text artifact systems, which despite efforts to be easier to use for novice programmers, still require a certain amount of computational thinking ability.

4.9 Funding Sources

The majority (36) of included papers had no funding sources listed. Three papers had funding from the United States government, through the National Science Foundation [5, 64, 65], Defense Advanced Research Projects Agency [64], or Intelligence Advanced Research Projects Activity [64]. Other works had government funding from the Spanish Government [39], Brazilian Government [45], Uruguayan Government [27], German Government [50], and the European Union [43]. One work got funding from "VC Research" [12], which we were unable to identify definitively but suspect refers to the first author whose initials are "VC". Two (2) works got support from commercial venues in the form of products, including D&D Beyond (parented by Wizards of the Coast) [64] and the WarpGate Store [55].

5 Discussion

5.1 Major Trends

The first research question (RQ1) asked what current academic efforts for computational support for TTRPGs looked like in terms of artifacts, evaluation, and computation methods. While we go over these factors in detail in the results, we noticed several trends among the included papers.

The first trend of interest is the high amount of human evaluation, with 52% (24) of papers including some human evaluation in their work. These evaluations took various forms, including quantitative, qualitative, and mixed-methods approaches. A particularly interesting technique that appeared in 5 papers, was including human feedback both to guide the design and to test the implemented systems. This practice, called "participatory design" [49] has shown success across design disciplines [7]. Future research could benefit from this model and can consider more areas where human feedback could be incorporated.

Another trend that appeared in the papers was incorporating digital and tangible elements in the design of tools. The most common artifact was tangible devices, from digital game boards to robotic assistants. These works highlight how part of the desire to play tabletop games, like TTRPGs, is the physical embodiment of gameplay that can be lost in purely digital games. By mixing digital elements with the tangible, the benefits of computation in terms of data tracking and visualizing can be combined with the desire for physical interaction. Future research can further investigate how digital tools can be incorporated into tangible artifacts.

Finally, these papers highlight the continued interest in "traditional" AI techniques even in the modern landscape of generative AI. While large language models (LLMs) were the most popular computational method investigated, older techniques like constraint-solving, machine learning algorithms, and genetic algorithms still have papers published in 2023 (and possibly the later half of 2024). These works highlight the value of continued investigation of older methods of computation alongside researching the most modern technological advancements.

5.2 Opportunities for Future Work

The second research question (RQ2) asked what areas remain unexplored in this domain. Through the examination of the included papers, we have identified several key open areas for future work.

The limitations identified in the included papers present new opportunities for future research in this area. One of the major limitations was the ease-of-use vs. controllability trade-off that was seen in many artifacts in this study. For computational tools to actually be used by GMs and players, they have to both lighten the workload of creation and abide closely by the vision of the designer. Tools, such as the domain-specific languages presented here, require incredible labor on the part of the human designer. These tools are likely to be too daunting or require too much effort for GMs or players to actually consider using. On the other hand, completely computational GMs lose the personalized human-designed aspect of gameplay that makes TTRPGs special and unique. A potential solution to this problem is the use of co-creative and mixed-initiative AI techniques. With this methodology, the computational tool and the human user would both contribute to the content being generated. This would allow the end user to give some of the responsibility of content creation to the computational system, while still being able to refine and guide this process. Only one of the papers presented in this study [11] proposed any kind of mixed-initiative method. Any work going forward that hopes to be used in real TTRPG games will need to consider how to make system both controllable and easy to use.

There are also opportunities in the evaluation of computational tools. None of the studies included in this review were evaluated

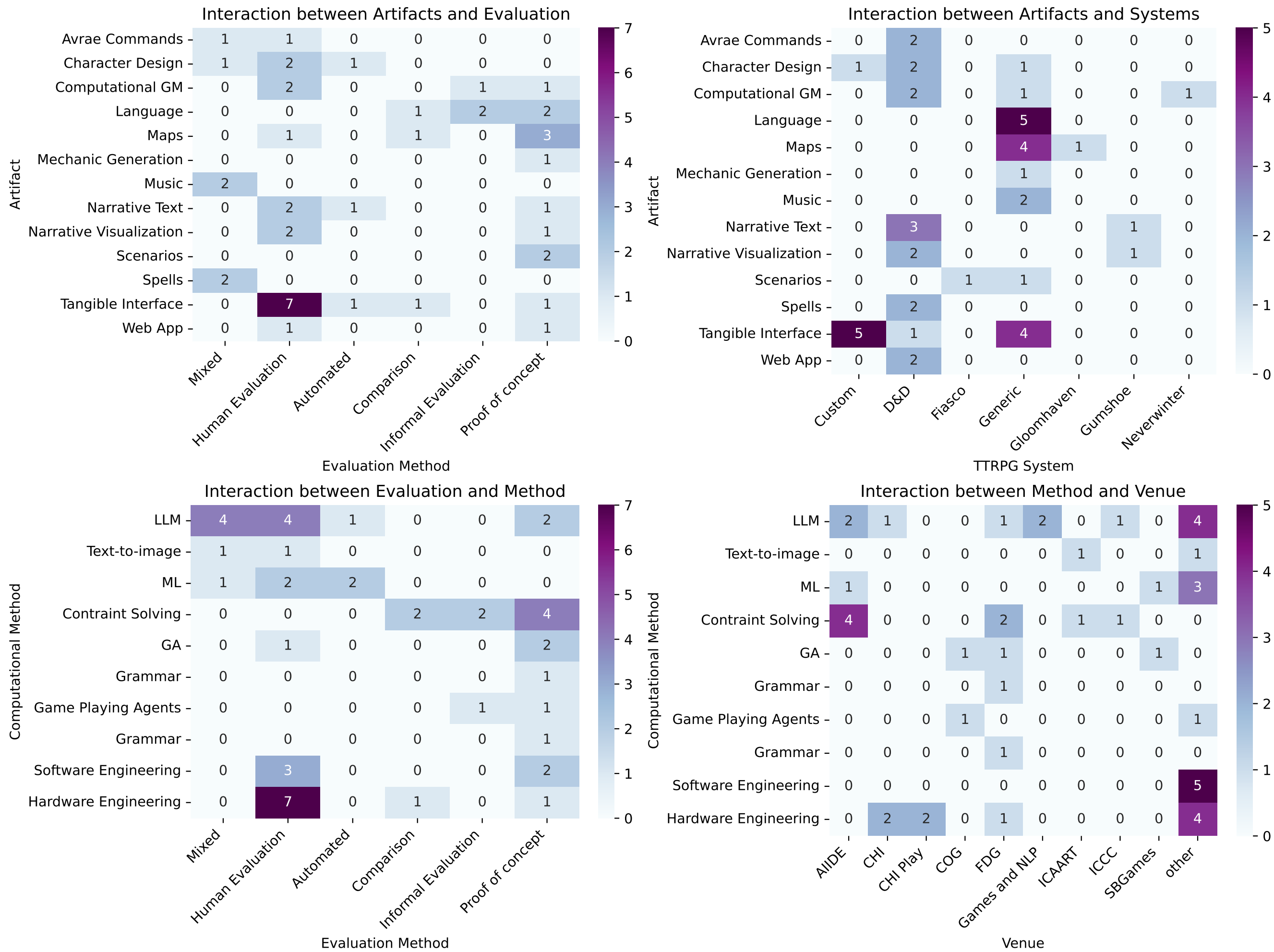

Figure 7: Heatmap of the count of papers at the intersection between different data items.

in realistic environments for TTRPG gameplay. Studies were most often done with parties that were not previously established, over games that were not designed by the GM, and only lasted a single session. This is not representative of the way TTRPGs are actually played, often between friends with a close bond, with a personalized gameplay experience, and campaigns that last multiple months to multiple years. These are all aspects of TTRPGs that make them unique from other types of gameplay, and the failure to represent these aspects brings into question the usefulness of these tools. A study that looks at the use of a computational tool throughout the entire process of creating and playing a TTRPG, from brainstorming and planning through running a campaign over multiple sessions, would be of great value to the research community. This type of study would, however, involve a significant amount of time and resources to conduct.

In terms of AI-based techniques, the investigation of models that work with a smaller amount of training data can be of value. This would allow tools to apply to settings and systems that are less popular (and therefore have less data) than larger systems like D&D. Many of the machine learning techniques presented here produce artifacts representitive of only those most commonly thought of as TTRPG content: Dungeon and Dragons artifacts set in a medieval high fantasy setting with oil paint styled illustrations. Procedural generation is most useful, not for replicating what already exists, but for expanding the diversity of content that exists; allowing amueter designers to bring their individual aspirations to life. To accomplish this, machine learning approaches would need to be adapted to only work with a small amount of previously created artifacts. This can also help mitigate some of the ethical issues behind current generative AI techniques. Smaller amounts of training data allow researchers to better vet the data in terms of copyright and access rights, along with biases that may be present. Additionally, smaller models are less computationally expensive and therefore have a smaller carbon impact. While models with smaller training data have their set of limitations, investigation into these types of models could be of great value.

A consideration for future work is what systems the computational tools are geared towards. For example, all included papers related to narrative focused on a specific system. It seems like it would be possible to create tools that are system agnostic, which can increase the population that may use them. Additionally, tangible interfaces were often designed based on a system designed for the study. Future work can consider more system-agnostic tangible interfaces, or creating tangible interfaces geared toward existing systems.

Another limitation that can guide future work is the accessibility of tools presented here. Future work can consider how to incorporate digital and physical equipment in a cost-efficient way that can allow people with fewer resources to access these tools. It is also interesting that none of the studies, particularly with tangible interfaces, addressed a major issue within TTRPG games: the experience of players when they are not in person. Several commercial tools give access to virtual tabletops for remote play, but this does not fully replicate the in-person experience. Additionally, to our knowledge, there are not tools designed for hybrid experience: where some players are remote and others in person. Digital game boards seem like a great opportunity to optimize the experience for both in-person and remote player, yet none of the studies explored this possibility. We can also consider ways to allow users to access the expressiveness of languages without needing the high amount of technical skills needed to use programming languages. Potentially some sort of visual interface can be developed as a wrapper for these proposed languages.

Finally, there are several aspects of gameplay in TTRPGs that were not presented in these studies, in particular combat and puzzle-solving. While Arnold et al. [5] presented a digital tool to generate combat encounters, this used the very simplistic model of random look-up tables (which were already available in physical material such as the Dungeon Master's Guide). More advanced techniques for generating encounters or monsters could be considered in future work. Further, while previous work [20] has investigated the generation of narrative-based puzzles, this has not been applied or tested in the environment of TTRPGs. Future work can investigate ways to help GMs develop puzzles that fit into the narrative of the campaign while being interesting and fun to solve.

6 Limitations

Our study has several limitations to its validity. The major limitation is that this work only investigates academic endeavors. There are many computational tools that GMs use outside of academia that would be of value to investigate. Further, due to the scope of this study, we excluded any theses from our evaluation. Our queries identified many undergraduate, master's, and doctoral theses that could be of value to the academic community. It is also possible that limiting this review to papers with implemented systems excluded papers that proposed novel ideas or evaluated previous work (like with Crain et al. [13]). We also limited our search terms to those of TTRPGs generally and for the most popular system "D&D." It is possible we could have identified more work if we had included other systems such as Pathfinder or Call of Cthulhu. The review process was also done by only one person, which means that there is some subjectivity in the inclusion and evaluation of papers.

7 Conclusion

We examined 46 papers related to computational tools for table-top role-playing games. These papers contained a wide variety of produced artifacts, computational methods, and evaluations. The most popular artifacts were tangible devices, the most popular method was large language model, and the most popular evaluation technique was human evaluation. We also identified major gaps in current research including mixed-initiative methods, long-term human evaluation, smaller AI models, and tools for combat or puzzles. These gaps provide opportunities for future research endeavors.

Acknowledgments

We would like to thank Regina Pagani and Philip Espinola Coombs for their help in organizing this review. Our current work on this review was not supported by external sources.

A All Included Papers

Title	Year	Artifact	Method	Evaluation	System
The TViews table role-playing game [40]	2008	Tangible Interface	Hardware	Human	Custom
Troll, a language for specifying dice-rolls [42]	2009	Language	Software	Prototype	Generic
Adventures in level design: generating missions and spaces for action adventure games [17]	2010	Maps	Grammar	Prototype	Generic
Tisch digital tools supporting board games [29]	2012	Tangible Interface	Hardware	Human	Generic
ToyVision: a toolkit to support the creation of innovative board-games with tangible interaction [39]	2013	Tangible Interface	Hardware	Comparison	Generic
Automatic generation of fantasy role-playing modules [6]	2014	Maps	GA	Prototype	Generic
Generating and adapting game mechanics [66]	2014	Mechanics	Constraint	Prototype	Generic
Fiascomatic: A framework for automated Fiasco playsets [30]	2015	Scenarios	Constraint	Prototype	Fiasco
WEARPG: game design implications for movement-based play in table-top role-playing games with arm-worn devices [9]	2016	Tangible Interface	Hardware	Human	Custom
WEARPG: Movement-Based Tabletop Role-Playing Game with Arm-Worn Devices and an Augmented Die [8]	2017	Tangible Interface	Hardware	Human	Custom
Bardo: Emotion-based music recommendation for tabletop role-playing games [45]	2017	Music	ML	Mixed	Generic
A BDI Game Master Agent for Computer Role-Playing Games [38]	2017	Computer GM	Agent	Informal	Neverwinter
Extracting Design Guidelines for Wearables and Movement in Tabletop Role-Playing Games via a Research Through Design Process [10]	2018	Tangible Interface	Hardware	Human	Custom
Imaginarium: A tool for casual constraint-based pcg [31]	2019	Language	Constraint	Prototype	Generic
Mysterious murder-mcts-driven murder mystery generation [35]	2019	Scenarios	Agent	Prototype	Generic
Adaptive Generation of Characters for Tabletop Role Playing Games [19]	2019	Character Design	GA	Human	D&D
The calculation of player's and non-player character's gameplay attribute growth in role-playing game with K-NN and Naive Bayes [58]	2020	Character Design	ML	Automated	Custom
Truesight Battle Grid - Enhancing the Game Experience of Tabletop Role-Playing through Tangible Data Visualization [48]	2020	Tangible Interface	Hardware	Human	D&D
Computer-generated music for tabletop role-playing games [21]	2020	Music	LLM	Mixed	Generic
A declarative PCG tool for casual users [32]	2020	Language	Constraint	Prototype	Generic
Kismet: a small social simulation language [53]	2020	Language	Constraint	Informal	Generic
Interviews towards designing support tools for TTRPG game masters [2]	2021	Visualization	Software	Human	D&D
Casual Creation of Tile Maps via Authorable Constraint-Based Generators [11]	2021	Maps	Constraint	Comparison	Generic
Chaotic Creations [5]	2021	Web App	Software	Prototype	D&D
Party Mascot: Experimental Prop Design for Streaming Actual Plays [52]	2022	Tangible Interface	Hardware	Prototype	Generic
Step: a highly expressive text generation language [33]	2022	Language	Constraint	Comparison	Generic
Generating descriptive and rules-adhering spells for dungeons &dragons fifth edition [44]	2022	Spells	LLM	Mixed	D&D
Using language models to convert between natural language and game commands [46]	2022	Avrae Commands	LLM	Human	D&D
Requirements analysis and speculative design of support tools for TTRPG game masters [3]	2022	Visualization	Software	Human	D&D
Conductive Fiducial Tangibles for Everyone: A Data Simulation-Based Toolkit using Deep Learning [51]	2022	Tangible Interface	ML	Automated	Generic
A Primer on Synthesis and Evaluation of a Domain-specific Large Data Set for Dungeons &Dragons [47]	2022	Narrative Text	LLM	Automated	D&D
Towards Computational Support with Language Models for TTRPG Game Masters [36]	2023	Narrative Text	LLM	Prototype	Gumshoe
WorldSmith: Iterative and Expressive Prompting for World Building with a Generative AI [15]	2023	Maps	Text-to-image	Human	Generic
Shoelace: A Storytelling Assistant for GUMSHOE One-2-One [1]	2023	Visualization	Constraint	Prototype	Gumshoe
Fireball: A dataset of dungeons and dragons actual-play with structured game state information [64]	2023	Avrae Commands	LLM	Mixed	D&D
CALYPSO: LLMs as Dungeon Master's Assistants [65]	2023	Narrative Text	LLM	Human	D&D
Exploring the Potential of ChatGPT as a Dungeon Master in Dungeons & Dragons tabletop game [55]	2023	Computer GM	LLM	Human	D&D

Continued on next page

Title	Year	Artifact	Method	Evaluation	System
Fable Reborn: Investigating Gameplay Experience between a Human Player and a Virtual Dungeon Master [4]	2023	Computer GM	LLM	Human	D&D
Enhancing Interactive Storytelling: A Computational Approach to Autonomously Generating Role-Playing Game Scripts with Natural Language Processing [63]	2023	Narrative Text	ML	Human	D&D
Computer Aided Content Generation – A Gloomhaven Case Study [24]	2023	Maps	GA	Prototype	Gloomhaven
RPG Creature Design: Cross-System Analysis and Conversion [22]	2023	Character Design	ML	Human	D&D
PAYADOR: A Minimalist Approach to Grounding Language Models on Structured Data for Interactive Storytelling and Role-playing Games [27]	2024	Computer GM	LLM	Prototype	Generic
A Multi-Stage Approach to Image Consistency in Zero-Shot Character Art Generation for the D&D Domain [57]	2024	Character Design	Text-to-image	Mixed	Generic
Leveraging Large Language Models for Spell-Generation in Dungeons &Dragons [43]	2024	Spells	LLM	Mixed	D&D
DungeonMaker: Embedding Tangible Creation and Destruction in Hybrid Board Games through Personal Fabrication Technology [50]	2024	Tangible Interface	Hardware	Human	Custom
New product design and implementation of aboleth: a mobile D&D character creator for enterprise mobile applications and metaverse [12]	2024	Web App	Software	Human	D&D

B Query Text

The full query text for the search is given below. This string was used for all databases.

```
(procedural OR automatic OR computational) AND (tool OR
     generation OR assistant) AND ("Tabletop role-play"
     OR "table top role play" OR "table-top role-play"
     OR "Tabletop roleplay" OR "TTRPG" OR "TTRPGS" OR "
     TRPG" OR "TRPGS" OR "Dungeons and Dragons" OR "
     Dungeons & Dragons") AND game
```

References

[1] Devi Acharya, Jack Kelly, William Tate, Maxwell Joslyn, Michael Mateas, and Noah Wardrip-Fruin. 2023. Shoelace: A Storytelling Assistant for GUMSHOE One-2-One. In *Proceedings of the 18th International Conference on the Foundations of Digital Games (FDG)* (Lisbon Portugal, 2023-04-12). ACM, 1–9. https://doi.org/10.1145/3582437.3582475

[2] Devi Acharya, Michael Mateas, and Noah Wardrip-Fruin. 2021. Interviews Towards Designing Support Tools for TTRPG Game Masters. *Interactive Storytelling* (2021), 283–287. https://doi.org/10.1007/978-3-030-92300-6_26

[3] Devi Acharya, Michael Mateas, and Noah Wardrip-Fruin. 2022. Requirements analysis and speculative design of support tools for TTRPG game masters. In *Proceedings of DiGRA 2022 Conference: Bringing Worlds Together* (2022-01-01). https://dl.digra.org/index.php/dl/article/view/1310 ISSN: 2342-9666.

[4] Charlene Ang, Lorraine Renee Cortel, Carlo Luis Santos, and Ethel Ong. 2023. Fable Reborn: Investigating Gameplay Experience between a Human Player and a Virtual Dungeon Master. In *Extended Abstracts of the 2023 Conference on Human Factors in Computing Systems (CHI)* (Hamburg Germany, 2023-04-19). ACM, 1–7. https://doi.org/10.1145/3544549.3585793

[5] Natalie Arnold, Leah Kramer, Christopher Lewis, Terra Williams, Sergiu Dascalu, and Frederick Harris. 2021. Chaotic Creations. In *29th International Conference on Software Engineering and Data Engineering (SEDE)*. 76–65. https://doi.org/10.29007/bz75

[6] Daniel Ashlock and Cameron McGuinness. 2014. Automatic generation of fantasy role-playing modules. In *2014 IEEE Conference on Computational Intelligence and Games (CIG)* (2014-08). 1–8. https://doi.org/10.1109/CIG.2014.6932867 ISSN: 2325-4289.

[7] Claus Bossen, Christian Dindler, and Ole Sejer Iversen. 2016. Evaluation in participatory design: a literature survey. In *Proceedings of the 14th Participatory Design Conference: Full papers-Volume 1*. 151–160.

[8] Oğuz Turan Buruk, Ismet Melih Özbeyli, and Oğuzhan Özcan. 2017. WEARPG: Movement-Based Tabletop Role-Playing Game with Arm-Worn Devices and an Augmented Die. In *Extended Abstracts Publication of the Annual Symposium on Computer-Human Interaction in Play (CHI-Play)* (Amsterdam The Netherlands, 2017-10-15). ACM, 639–646. https://doi.org/10.1145/3130859.3130868

[9] Oğuz Turan Buruk and Oğuzhan Özcan. 2016. WEARPG: game design implications for movement-based play in table-top role-playing games with arm-worn devices. In *Proceedings of the 20th International Academic Mindtrek Conference* (Tampere Finland, 2016-10-17). ACM, 403–412. https://doi.org/10.1145/2994310.2994315

[10] Oğuz Turan Buruk and Oğuzhan Özcan. 2018. Extracting Design Guidelines for Wearables and Movement in Tabletop Role-Playing Games via a Research Through Design Process. In *Proceedings of the 2018 Conference on Human Factors in Computing Systems (CHI)* (Montreal QC Canada, 2018-04-21). ACM, 1–13. https://doi.org/10.1145/3173574.3174087

[11] Dan Carpenter, John Thomas Bacher, Henry Crain, and Chris Martens. 2021. Casual Creation of Tile Maps via Authorable Constraint-Based Generators. In *Artifical Intelligence in Digital Entertainment (AIIDE)*.

[12] Victor Chang, Dan Lawrence, Le Minh Thao Doan, Ariel Qianwen Xu, and Ben S.C. Liu. 2024. New product design and implementation of aboleth: a mobile D&D character creator for enterprise mobile applications and metaverse. *Enterprise Information Systems* 18, 1 (2024), 2193557. https://doi.org/10.1080/17517575.2023.2193557

[13] Henry Crain, Dan Carpenter, and Chris Martens. 2022. Evaluating a Casual Procedural Generation Tool for Tabletop Role-Playing Game Maps. In *2022 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*. IEEE, 1–6.

[14] Jeremy Crawford, James Wyatt, Robert J Schwalb, and Bruce R Cordell. 2014. *Player's handbook*. Wizards of the Coast LLC.

[15] Hai Dang, Frederik Brudy, George Fitzmaurice, and Fraser Anderson. 2023. World-Smith: Iterative and Expressive Prompting for World Building with a Generative AI. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology (UIST)* (San Francisco CA USA, 2023-10-29). ACM, 1–17. https://doi.org/10.1145/3586183.3606772

[16] Barbara De Kegel and Mads Haahr. 2020. Procedural Puzzle Generation: A Survey. *IEEE Transactions on Games* 12, 1 (2020), 21–40. https://doi.org/10.1109/TG.2019.2917792 Conference Name: IEEE Transactions on Games.

[17] Joris Dormans. 2010. Adventures in level design: generating missions and spaces for action adventure games. In *Proceedings of the 2010 Workshop on Procedural Content Generation in Games (at FDG)* (Monterey California, 2010-06-18). ACM. https://doi.org/10.1145/1814256.1814257

[18] Brad Everman, Trevor Villwock, Dayuan Chen, Noe Soto, Oliver Zhang, and Ziliang Zong. 2023. Evaluating the carbon impact of large language models at the inference stage. In *2023 IEEE international performance, computing, and communications conference (IPCCC)*. IEEE, 150–157.

[19] Fabrício Guedes Faria, Leonardo Tórtoro Pereira, and Cláudio Fabiano Motta Toledo. 2019. Adaptive Generation of Characters for Tabletop Role Playing Games. In *2019 18th Brazilian Symposium on Computer Games and Digital Entertainment (SBGames)* (2019-10). 39–46. https://doi.org/10.1109/SBGames.2019.00016 ISSN: 2159-6662.

[20] Clara Fernández-Vara and Alec Thomson. 2012. Procedural generation of narrative puzzles in adventure games: The puzzle-dice system. In *Proceedings of the The third workshop on Procedural Content Generation in Games*. 1–6.

[21] Lucas Ferreira, Levi Lelis, and Jim Whitehead. 2020. Computer-Generated Music for Tabletop Role-Playing Games. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment* (2020-10-01), Vol. 16. 59–65. https://doi.org/10.1609/aiide.v16i1.7408 Number: 1.

[22] Artur De Oliveira Da Rocha Franco, Thiago Iachiley AraúJo De Souza, Paulo De Tarso Passos Da Costa Filho, José Wellington Franco Da Silva, José Gilvan Rodrigues Maia, and Miguel Franklin De Castro. 2023. RPG Creature Design: Cross-System Analysis and Conversion. In *Proceedings of the 22nd Brazilian Symposium on Games and Digital Entertainment (SBGames)* (Rio Grande (RS) Brazil, 2023-11-06). ACM, 96–105. https://doi.org/10.1145/3631085.3631332

[23] Sebastian Gajewski, Nour El Mawas, and Jean Heutte. 2022. A Systematic Literature Review of Game Design Tools:. In *Proceedings of the 14th International Conference on Computer Supported Education* (Online Streaming, — Select a Country —, 2022). SCITEPRESS - Science and Technology Publications, 404–414. https://doi.org/10.5220/0011137800003182

[24] Marcus Gerhold and Kristian Tijben. 2023. Computer Aided Content Generation – A Gloomhaven Case Study. In *Proceedings of the 18th International Conference on the Foundations of Digital Games* (Lisbon Portugal, 2023-04-12). ACM, 1–10. https://doi.org/10.1145/3582437.3587196

[25] Matthew Guzdial, Devi Acharya, Max Kreminski, Michael Cook, Mirjam Eladhari, Antonios Liapis, and Anne Sullivan. 2020. Tabletop roleplaying games as procedural content generators. In *Proceedings of the 15th International Conference on the Foundations of Digital Games*. 1–9.

[26] Gary Gygax and David Cook. 1989. *The Dungeon Master Guide, No. 2100, 2nd Edition (Advanced Dungeons and Dragons)*. TSR, Inc.

[27] Santiago Góngora, Luis Chiruzzo, Gonzalo Méndez, and Pablo Gervás. 2024. *PAYADOR: A Minimalist Approach to Grounding Language Models on Structured Data for Interactive Storytelling and Role-playing Games*.

[28] Andreas Halman and Alicia Oshlack. 2024. Catchii: Empowering literature review screening in healthcare. *Research Synthesis Methods* 15, 1 (2024), 157–165. https://doi.org/10.1002/jrsm.1675 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/jrsm.1675

[29] Ulf Hartelius, Johan Fröhlander, and Staffan Björk. 2012. Tisch digital tools supporting board games. In *Proceedings of the International Conference on the Foundations of Digital Games (FDG)* (Raleigh North Carolina, 2012-05-29). ACM, 196–203. https://doi.org/10.1145/2282338.2282376

[30] Ian Horswill. 2015. Fiascomatic: A Framework for Automated Fiasco Playsets. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE)*, Vol. 11. 22–29. https://doi.org/10.1609/aiide.v11i3.12823 Number: 3.

[31] Ian Horswill. 2019. Imaginarium: A Tool for Casual Constraint-Based PCG. In *International Conference on Agents and Artificial Intelligence (ICCART)*.

[32] Ian Horswill. 2020. A Declarative PCG Tool for Casual Users. 16, 1 (2020), 81–87. https://doi.org/10.1609/aiide.v16i1.7411 Number: 1.

[33] Ian Horswill. 2022. Step: A Highly Expressive Text Generation Language. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE)* (2022-10-11), Vol. 18. 240–249. https://doi.org/10.1609/aiide.v18i1.21969 Number: 1.

[34] Lei Huang, Weijiang Yu, Weitao Ma, Weihong Zhong, Zhangyin Feng, Haotian Wang, Qianglong Chen, Weihua Peng, Xiaocheng Feng, Bing Qin, et al. 2024. A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions. *ACM Transactions on Information Systems* (2024).

[35] Corinna Jaschek, Tom Beckmann, Jaime A. Garcia, and William L. Raffe. 2019. Mysterious Murder - MCTS-driven Murder Mystery Generation. In *2019 IEEE Conference on Games (CoG)* (2019-08). 1–8. https://doi.org/10.1109/CIG.2019.8848000 ISSN: 2325-4289.

[36] Jack Kelly, Michael Mateas, and Noah Wardrip-Fruin. 2023. Towards Computational Support with Language Models for TTRPG Game Masters. In *Proceedings of the 18th International Conference on the Foundations of Digital Games (FDG)* (New York, NY, USA, 2023-04-12) *(FDG '23)*. Association for Computing Machinery, 1–4. https://doi.org/10.1145/3582437.3587202

[37] C Kulkarni. 2022. Ethical implications of large language models in content generation. *Journal of Artificial Intelligence, Machine Learning & Data Science* 1, 1 (2022), 62–67.

[38] Bao Luong, John Thangarajah, and Fabio Zambetta. 2017. A BDI Game Master Agent for Computer Role-Playing Games. 15, 1 (2017), 1–16. https://doi.org/10.1145/2629700

[39] Javier Marco, Sandra Baldassarri, and Eva Cerezo. 2013. ToyVision: a toolkit to support the creation of innovative board-games with tangible interaction. In *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction* (Barcelona Spain, 2013-02-10). ACM, 291–298. https://doi.org/10.1145/2460625.2460674

[40] Ali Mazalek, Basil Mironer, Elijah O'Rear, and Dana Van Devender. 2008. The TViews Table Role-Playing Game. *JVRB - Journal of Virtual Reality and Broadcasting* 5(2008), 8 (2008). https://doi.org/10.20385/1860-2037/5.2008.8

[41] Mike Mearls and Jeremy Crawford. 2014. *Monster Manual: Dungeons and Dragons*. Wizards of the Coast.

[42] Torben Ægidius Mogensen. 2009. Troll, a language for specifying dice-rolls. In *Proceedings of the 2009 ACM symposium on Applied Computing (SAC)* (Honolulu Hawaii, 2009-03-08). ACM, 1910–1915. https://doi.org/10.1145/1529282.1529708

[43] Elio Musacchio, Lucia Siciliani, Pierpaolo Basile, and Giovanni Semeraro. 2024. Leveraging Large Language Models for Spell-Generation in Dungeons & Dragons. In *Proceedings of the 10th Workshop on Games and Natural Language Processing @ LREC-COLING 2024* (Torino, Italia, 2024-05), Chris Madge, Jon Chamberlain, Karen Fort, Udo Kruschwitz, and Stephanie Lukin (Eds.). ELRA and ICCL, 61–69. https://aclanthology.org/2024.games-1.7

[44] Pax Newman and Yudong Liu. 2022. Generating Descriptive and Rules-Adhering Spells for Dungeons & Dragons Fifth Edition. In *Proceedings of the 9th Workshop on Games and Natural Language Processing within the 13th Language Resources and Evaluation Conference* (Marseille, France, 2022-06), Chris Madge (Ed.). European Language Resources Association, 54–60. https://aclanthology.org/2022.games-1.7

[45] Rafael Padovani, Lucas Ferreira, and Levi Lelis. 2017. Bardo: Emotion-Based Music Recommendation for Tabletop Role-Playing Games. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment(AIIDE)*, Vol. 13. 214–220. https://doi.org/10.1609/aiide.v13i1.12958 Number: 1.

[46] Stefan Papazov, Wesley Gill, Marta Garcia Ferreiro, Andrew Zhu, Lara J Martin, and Chris Callison-Burch. 2022. Using Language Models to Convert Between Natural Language and Game Commands. In *The Third Wordplay: When Language Meets Games Workshop*.

[47] Akila Peiris and Nisansa de Silva. 2022. A Primer on Synthesis and Evaluation of a Domain-specific Large Data Set for Dungeons & Dragons. In *6th Widening NLP Workshop*.

[48] Bas A Plijnaer, Daisy O'Neill, Eloisa Kompier, Günter Wallner, and Regina Bernhaupt. 2020. Truesight Battle Grid - Enhancing the Game Experience of Tabletop Role-Playing through Tangible Data Visualization. In *Extended Abstracts of the 2020 Annual Symposium on Computer-Human Interaction in Play (CHI-Play)*. https://dl.acm.org/doi/abs/10.1145/3383668.3419931

[49] Clay Spinuzzi. 2005. The methodology of participatory design. *Technical communication* 52, 2 (2005), 163–174.

[50] Evgeny Stemasov, Tobias Wagner, Ali Askari, Jessica Janek, Omid Rajabi, Anja Schikorr, Julian Frommel, Jan Gugenheimer, and Enrico Rukzio. 2024. DungeonMaker: Embedding Tangible Creation and Destruction in Hybrid Board Games through Personal Fabrication Technology. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI)* (Honolulu HI USA, 2024-05-11). ACM, 1–20. https://doi.org/10.1145/3613904.3642243

[51] Benedict Steuerlein and Sven Mayer. 2022. Conductive Fiducial Tangibles for Everyone: A Data Simulation-Based Toolkit using Deep Learning. *Proceedings of the ACM on Human-Computer Interaction* 6 (2022), 1–22. Issue MHCI. https://doi.org/10.1145/3546718

[52] Colin Stricklin, Xingyu Li, and Michael Nitsche. 2022. Party Mascot: Experimental Prop Design for Streaming Actual Plays. In *ACM International Conference on Interactive Media Experiences (IMX)* (Aveiro JB Portugal, 2022-06-21). ACM, 369–374. https://doi.org/10.1145/3505284.3532986

[53] Adam Summerville and Ben Samuel. 2020. Kismet: A Small Social Simulation Language. In *International Conference on Computational Creativity (ICCC)*.

[54] Kevin Tang, Terra Mae Gasque, Rachel Donley, and Anne Sullivan. 2023. "It Has to Ignite Their Creativity": Opportunities for Generative Tools for Game Masters. In *Proceedings of the 18th International Conference on the Foundations of Digital Games*. 1–6.

[55] Tuul Triyason. 2023. Exploring the Potential of ChatGPT as a Dungeon Master in Dungeons & Dragons tabletop game. In *Proceedings of the 13th International Conference on Advances in Information Technology (IAIT)* (Bangkok Thailand, 2023-12-06). ACM, 1–6. https://doi.org/10.1145/3628454.3628457

[56] Breno M. F. Viana and Selan R. dos Santos. 2021. Procedural Dungeon Generation: A Survey. *Journal on Interactive Systems* 12, 1 (2021), 83–101. https://doi.org/10.5753/jis.2021.999 Number: 1.

[57] Gayashan Weerasundara and Nisansa De Silva. 2024. A Multi-Stage Approach to Image Consistency in Zero-Shot Character Art Generation for the D&D Domain. In *Proceedings of the 16th International Conference on Agents and Artificial Intelligence (ICCART)* (Rome, Italy). SCITEPRESS - Science and Technology Publications, 235–242. https://doi.org/10.5220/0012317600003636

[58] Nur Rohman Widiyanto, Supeno Mardi Susiki Nugroho, and Mauridhi Hery Purnomo. 2020. The Calculation of Player's and Non-Player Character's Gameplay Attribute Growth in Role-Playing Game with K-NN and Naive Bayes. In *2020 International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM)* (2020-11). 103–110. https://doi.org/10.1109/CENIM51130.2020.9297945

[59] Carole-Jean Wu, Ramya Raghavendra, Udit Gupta, Bilge Acun, Newsha Ardalani, Kiwan Maeng, Gloria Chang, Fiona Aga, Jinshi Huang, Charles Bai, et al. 2022. Sustainable ai: Environmental implications, challenges and opportunities. *Proceedings of Machine Learning and Systems* 4 (2022), 795–813.

[60] Daijin Yang, Erica Kleinman, and Casper Harteveld. 2024. GPT for Games: A Scoping Review (2020-2023). In *IEEE Conference on Games 2024* (2024-04-27). arXiv. https://doi.org/10.48550/arXiv.2404.17794 arXiv:2404.17794 [cs]

[61] Livia Yuliawati, Putri Ayu Puspieta Wardhani, and Joo Hou Ng. 2024. A Scoping Review of Tabletop Role-Playing Game (TTRPG) as Psychological Intervention: Potential Benefits and Future Directions. *Psychology Research and Behavior Management* 17 (2024), 2885–2903. https://doi.org/10.2147/PRBM.S466664 Publisher: Dove Medical Press _eprint: https://www.tandfonline.com/doi/pdf/10.2147/PRBM.S466664.

[62] Yuzhong Zhang, Guixuan Zhang, and Xinyuan Huang. 2022. A Survey of Procedural Content Generation for Games. In *2022 International Conference on Culture-Oriented Science and Technology (CoST)* (2022-08). 186–190. https://doi.org/10.1109/CoST57098.2022.00046

[63] Qijun Zhong. 2024. Enhancing Interactive Storytelling: A Computational Approach to Autonomously Generating Role-Playing Game Scripts with Natural Language Processing. In *2023 3rd International Signal Processing, Communications and Engineering Management Conference (ISPCEM)*. 714–719. https://doi.org/10.1109/ISPCEM60569.2023.00134

[64] Andrew Zhu, Karmanya Aggarwal, Alexander Feng, Lara J. Martin, and Chris Callison-Burch. 2023. FIREBALL: A Dataset of Dungeons and Dragons Actual-Play with Structured Game State Information. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 4171–4193. https://doi.org/10.18653/v1/2023.acl-long.229 arXiv:2305.01528 [cs]

[65] Andrew Zhu, Lara Martin, Andrew Head, and Chris Callison-Burch. 2023. CALYPSO: LLMs as Dungeon Master's Assistants. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE)* (2023-10-06), Vol. 19. 380–390. https://doi.org/10.1609/aiide.v19i1.27534 Number: 1.

[66] Alexander Zook and Mark O Riedl. 2014. Generating and Adapting Game Mechanics. In *Foundation of Digital Games (FDG)*.