Digital Studies / Le champ numérique

Leichman, Jeffrey M., and Ben Samuel. 2023. "Part I — Introduction: Social Physics and the VESPACE Project." *Digital Studies/Le champ numérique* 13(2): 1–33. DOI: https://doi.org/10.16995/dscn.8085.

Open Library of Humanities

Part I — Introduction: Social Physics and the VESPACE Project

Jeffrey M. Leichman, Louisiana State University, United States, jleichman@lsu.edu
Ben Samuel, University of New Orleans, United States, bsamuel@cs.uno.edu

This article introduces the VESPACE project, an international, multi-disciplinary digital humanities initiative to build a computer-mediated playable simulation—a video game—of the eighteenth-century Paris Fair theatre. As part of this project, a weeklong postgraduate workshop was convened by the authors in May 2020 to develop protocols and procedures for coding literary and historical data for the Ensemble social physics engine that will govern behaviour of non-player characters (NPCs) in the interactive model. This article lays out the history and theory of social physics, the potential impacts of this project on historiographic practice, and the methodology and outcomes of the workshop week. We conclude with a discussion of lessons learned and promising leads with respect to the future of applying social physics to humanities research. [This article is part of the collection *Computer Modelling and Simulation for Literary-Historical Research: VESPACE and Social Physics*.]

Cet article présente le projet VESPACE, une initiative pluridisciplinaire internationale dans le domaine des humanités numériques qui vise à construire une simulation ludique — un jeu vidéo — basée sur le Théâtre de Foire parisien au XVIIIe siècle. En mai 2020, les auteurs de cet article organisèrent un atelier postdoctoral d'une semaine dans le but de développer des protocoles et des procédures de codage de données littéraires et historiques pour le moteur de physique sociale Ensemble, qui règle le comportement des personnages non joueurs (PNJ) dans le modèle interactif. Cet article présente l'histoire et la théorie de la physique sociale, les impacts potentiels de ce projet sur la pratique historiographique, ainsi que la méthodologie et les résultats de la semaine d'atelier. Nous concluons par une discussion des leçons apprises et des perspectives prometteuses pour l'avenir de l'application de la physique sociale dans la recherche en sciences humaines. [Cet article fait partie de la collection *Modélisation et simulation informatiques pour la recherche littéraire-historique : VESPACE et physique sociale*.]

Editor's Note / Note de l'éditeur

The following article is part of Computer Modelling and Simulation for Literary–Historical Research: VESPACE and Social Physics, *a special issue that details the rationales, objectives, and strategies for authoring a playable experience of eighteenth-century public sociability in the context of the VESPACE project's goals to construct an immersive and interactive virtual reality model of an eighteenth-century Paris Fair theatre. The co-authored introductory article (**Jeffrey M. Leichman and Ben Samuel**) includes an in-depth discussion of the history, ambitions, and methodologies of the project. The second part of the special issue features separate contributions from two participants (**Paul François and Daniel DeKerlegand**) in the May 2020 social physics authoring workshop who provide technical insight into the affordances and challenges of this experience. The third part features four additional short essays that offer informed perspectives from emerging literary scholars who participated in the authoring workshop (**Louise Moulin, Charlee M. Bezilla, Julien Le Goff, and Chiara Azzaretti**), drawing on individual research pursuits to outline the disciplinary stakes of authoring computer simulation rules in the framework of the VESPACE project, while also looking ahead to potential futures for this kind of experimental work in literary-historical digital humanities.*

L'article suivant fait partie de Modélisation et simulation informatiques pour la recherche en histoire littéraire : VESPACE et la physique sociale, *un numéro spécial qui détaille les justifications, les objectifs et les stratégies pour rédiger une expérience jouable de la sociabilité publique du XVIIIe siècle dans le contexte du projet VESPACE, qui vise à construire un modèle immersif et interactif en réalité virtuelle d'un théâtre de Foire parisien du XVIIIe siècle. L'article d'introduction (co-écrit par **Jeffrey M. Leichman et Ben Samuel**) comprend une discussion approfondie de l'histoire, des ambitions et des méthodologies du projet. La deuxième partie du numéro spécial présente les contributions respectives de deux participants à l'atelier expérimental sur le codage en physique sociale de mai 2020 (**Paul François et Daniel DeKerlegand**), qui donnent un aperçu technique des moyens et des défis de cette expérience. La troisième partie comporte quatre courts essais, publiés ensemble, qui offrent des points de vue informés de nouveaux spécialistes de la littérature qui ont participé à l'atelier de codage (**Louise Moulin, Charlee M. Bezilla, Julien Le Goff et Chiara Azzaretti**), et qui s'appuient sur des recherches individuelles pour définir les enjeux disciplinaires de la rédaction de règles de simulation informatique dans le cadre du projet VESPACE, tout en regardant également vers l'avenir potentiel pour ce type de travail expérimental en humanités numériques historico-littéraires.*

Introduction to this special issue of *Digital Studies / Le champ numérique*

This special issue is one of three novel research outcomes from a weeklong experimental workshop conducted in May 2020. Unlike a traditional special issue, in which scholars are invited to contribute research around a common theme, the works collected here are all reflections on a shared digital humanities *experience*: a week of intensive discussion and coding of literary texts to create executable instructions for a playable social simulation. The rules generated during this workshop comprise another novel outcome in the form of a JSON file that is primarily "readable" through the effects the rules exercise on characters in the simulation. One way of understanding this publication—the present article and the individually authored field reports from participants—is as a critical annotation of these rules by the scholars who authored them, lending intellectual depth and rhetorical eloquence to otherwise abstract logical statements. The last major outcome of this early-pandemic experience was the authoring tool that served as our common interface during our intensive week together, custom-built for authors with little or no previous digital humanities experience to enable coding literary interpretation as executable JSON files in the Ensemble application.

The structure of what follows reflects differential levels of engagement with the larger project that gave rise to this experiment in harnessing literary expertise for building a social simulation. The long initial essay explains the background of the "Interactive VR Simulation of an Eighteenth-Century Paris Fair Theatre: VESPACE" project (National Endowment for the Humanities Award HAA 266501–19, https://vespace.cs.uno.edu), under whose auspices the workshop was convened. Two VESPACE project leads, Jeffrey M. Leichman and Ben Samuel, who have been developing aspects of this work since 2017, detail the history and objectives of VESPACE, which seeks to create a multi-dimensional model of an eighteenth-century Paris Fair theatre, to study the architectural, performance, and social characteristics of this unique *ancien régime* institution. Focusing particularly on the use of the computing metaphor of social physics to build an interactive simulation of this boisterous public space, this essay sets out the rationales, challenges, and strategies for our approach to authoring a playable experience of eighteenth-century public sociability.

The remainder of the contributions in this multi-part publication, written by participants in the May 2020 workshop, are each about half as long as a traditional peer-reviewed article. At the time of our workshop, all participants were doctoral students in either France or the USA (several have since completed their studies). One of them, Dr. Paul François (currently Research Engineer at the CNRS in Aix-Marseille), has been integral to the VESPACE project since its inception. He is the author of the virtual reality restitution of a Foire Saint-Germain marionette theatre, created as part of his doctoral dissertation, which serves as the virtual environment for the entire VESPACE experience; his essay reflects on the similarities and distinctions in methodology that underpin the

visual-spatial model, and the social physics-driven character interactions that have since been integrated within it (François and DeKerlegand 2023). Another participant, Daniel DeKerlegand (a PhD candidate in Computer Science who also holds an MA in English), is responsible for the authoring tool that framed and supported the collaboration of the workshop week; developing this software, in addition to continuing at a furious pace during the actual workshop, had begun nearly a year earlier. Mr. DeKerlegand's essay offers a detailed perspective on the technical facilitation of a complex, distributed database-building process able to respond to shifting technical, literary, and historical needs that emerged over the course of the workshop week (François and DeKerlegand 2023). These two contextual essays are presented together in order to more completely situate the goals and techniques of the authoring workshop.

The remaining four essays, also presented together, comprise contributions from researchers who had no previous contact with VESPACE, social physics, or the ideas, techniques, or ambitions of the workshop week—Louise Moulin, Chiara Azzaretti, Julien Le Goff, and Charlee M. Bezilla (Moulin et al. 2023). One of the goals of the workshop was to see how accessible this kind of coding could be made to literary scholars with no previous experience with this system, and in some cases little to no experience with digital humanities projects. In turn, this special issue aims to highlight how emerging scholars pursuing training in literary fields can also embrace, and be included in, digital initiatives that are transforming academic practice across the humanities.

These four contributions offer insight into the specificities of the protocols developed for this project, as well as a revealing look at how digital humanities appears from the perspective of scholars primarily trained in non-quantitative literary-historical studies. The first essay, by Louise Moulin, situates the ideas encountered in this workshop within a larger consideration of "literary data" in the digital humanities, confronting the challenges of reconciling technological processes with the subjective traditions of literary study (Moulin et al. 2023). Chiara Azzaretti next analyzes the unexpected similarities between the emotional engagement that drives much literary scholarship, and the rich and unexpected ways that this is transformed in a project to devise a computer-mediated representation of emotion (Moulin et al. 2023). Following this, Julien Le Goff drills down on a particularly compelling source, the *Anecdotes Dramatiques*, reconciling these short, sociologically acute, yet highly literary narratives with the experimental protocols of social physics coding (Moulin et al. 2023). The concluding essay, by Dr. Charlee M. Bezilla, turns on the notion of translation as both process and metaphor for the activities of our workshop, reflecting on how the technical and hermeneutic skills honed in advanced literary study also prepare humanists to meaningfully contribute to digital humanities projects (Moulin et al. 2023).

The goal in presenting this special issue, beyond providing a forum for high-quality work by early-career scholars, is also to influence ongoing conversations about knowledge representation in the twenty-first-century humanities. We hope that this unique publication can illustrate one model for how to implicate emerging researchers in collaborative, interdisciplinary, project-based work, from the planning phase through the peer-reviewed publication of results, and thus to further integrate digital humanities research in the intellectual and professional formation of the next generation of literary historians. The insight into the ways in which these scholars situate their own work with respect to digital humanities techniques, which often present forbidding barriers to entry, illustrates the disciplinary value of these reflections for traditional humanities fields. At the same time, the rigorous evaluation of both the techniques of the workshop and the personal challenges and rewards of working with simulation-building software underscores the value of actively recruiting and engaging scholars whose work is not otherwise oriented towards digital humanities in the development of powerful, accessible digital tools for literary-historical study.

VESPACE project background: A new approach to the history of eighteenth-century Fair theatre

The VESPACE hypothesis is that a computer-mediated model, capable of representing research on Fair theatre as a virtually rendered architectural environment containing a performance and allowing for social interaction between spectators, can better illustrate the current state of knowledge around this complex historical object. Before delving into further explanation of the computer models behind this ambition, it will be useful to provide a quick overview of the historical object we are modelling, the eighteenth-century Fair theatre (see **Figure 1**).

Figure 1: Nicolas van Blarenbergh, *Snuffbox with Theatrical Scenes of a Rope Dancer and a Puppet Show*, 1778–79, New York, Metropolitan Museum of Art (Creative Commons).

It can be misleading to speak of Fair theatre in the singular; in eighteenth-century Paris, the two main Fairs that hosted theatrical entertainments were located at the Saint-Germain abbey (winter) and the Saint-Laurent abbey (late summer), and the theatres that commissioned and staged works there frequently changed hands according to who could assemble the resources to rent a space from the host while also (in those years where an agreement had been reached) paying for the right to present their entertainment without significant interference from royally protected theatres (see Martin 2002 and Isherwood 1989 for overviews of Paris Fair theatre history). When Fair theatre managers and their officially sanctioned counterparts could not agree on terms, the result was often legal proceedings and cancellations, sometimes of entire seasons. While the Académie Royale de Musique (known as l'Opéra) and the Comédie Française enjoyed royal privileges granting them exclusive rights to perform lyric or dramatic works on the public stages of Europe's largest city, the Fair theatre received no subventions and enjoyed no special protections from the state. Despite this, Fair entrepreneurs consistently drew large and socially diverse crowds to their shows. As in all public theatres in early eighteenth-century France, the *parterre* at the Fair was an open space in front of the stage that was a forum for (primarily masculine) sociability among spectators who remained standing throughout the spectacle. Contrary to prevailing theatrical norms today, this boisterous section of the audience was able to move about the *parterre*, interacting with each other, the seated portions of the audience, and even the actors onstage. Attending the theatre at this time was a far less passive experience than what modern audiences have come to accept and expect—spectators were illuminated, upright, and vocal, rather than mutely seated in the dark—and it is this dynamic, diverse, and occasionally violent atmosphere that our social physics-based simulation seeks to represent.

This quick sketch of the both the mores of eighteenth-century theatre attendance and the institutional marginalization of Fair theatre indicates some of the numerous challenges that attend to historical research around this little-known but highly influential theatrical form. Aesthetically, the "Fair theatre" included a variety of different spectacles at a single fair; for example, an original three-act play with music and singing (by both performers and audience), staged in a theatre capable of sophisticated scenic effects, or a simple marionette show with two musicians and a white-clad Pierrot narrating from in front of a "castelet" (a miniature proscenium insert that frames the puppet action and hides the marionettists), could simultaneously be on offer at the Foire Saint-Germain, which also hosted dozens of merchant stalls selling luxury goods and food. As a result of the Fair's creative variety and dubious legal standing, administrative and financial records available to help scholars understand the material and administrative life of these theatres, despite their widespread popularity,

are both disparate and incomplete. Whereas scrupulous registers have facilitated the modern understanding of the artistry and professional life of the Comédie Française (see in particular the Comédie Française Registers project [RCF 2022],), no such resource exists for the Fair regarding expenses, theatre locations, and even repertoire. Work conducted by the Centre d'Etudes des Théâtres de Foire et de la Comédie-Italienne (CETHEFI) lab in Nantes, France, directed by VESPACE lead Françoise Rubellin at the Université de Nantes, has begun to reveal the richness and sophistication of the comic theatre presented at these venues. (For online resources, see CETHEFI 2022 and Theaville 2022. Also see publications by Rubellin 2005, Beaucé 2016, and Rubellin 2022, and CETHEFI-affiliated doctoral dissertations by Prou 2019, Sakhnovskaia 2013, Chahine 2014, and François 2021.)

The VESPACE project seeks to build on this research and, in particular, to reconceive the forms of knowledge representation that can usefully emerge from the database-building capacities of computers. We contend that a multi-dimensional, sensory, interactive model can also point the way to new areas of research, as much for the Fair theatre as for other historical phenomena whose cultural meaning emerged from ephemeral public interactions in a specific time and place.

After completing a prototype of the architectural restitution, the VESPACE research team pivoted to populating this space with computer-controlled non-player characters (NPCs), a necessary element in modelling one of the *ancien régime*'s most socially diverse and artistically vibrant public spaces. The initial immersive architectural model was experienced by hundreds of users in Europe and dozens in the United States before Covid-19 abruptly halted our American rollout in March 2020. (A completed playable model is now available for download at our project website, https://vespace.cs.uno.edu/.) In designing the social interaction system, we used the second generation of an AI game engine initially developed for the independent video game *Prom Week*, a tool called Ensemble that allows for the design of an internally coherent method for deriving NPC desires and actions (as well as framing and channeling in-game choices for human players) based on designated characteristics and relationships. The initial interest in this system for the VESPACE project lies in its openness to different cultural and historical specificities, and its ability to transparently document the relationship between computer-generated interactions and the primary source materials used to lay the foundation of the social simulation. Ensemble provides a tool that goes beyond footnoting a computer-generated social interaction by allowing researchers to ground this action in scholarship, establishing an epistemological pendant to the careful documentation that underpins the VESPACE architectural reconstitution and its integrated research module, Prouvé. (See the essay in this collection by Paul François, which compares the methodologies developed for creating the architectural

model with the principles that guided the workshop week spent coding social physics predicates [François and DeKerlegand 2023].) By offering the ability to assess, critique, and eventually revise the interactive portion of our theatre model, Ensemble provides a means to fulfill our goal to prototype a digital, playable historical research resource, using Fair theatre as the initial case study.

For the purposes of this workshop, we asked participants to select period works, including fiction, manuals of comportment, and anecdotes related to the theatre, as sources for our Ensemble coding. One goal of this project is to lower barriers to participation in sophisticated digital projects. In this instance, we were interested in leveraging these students' existing expertise in order to help build up the knowledge base of schema and rules that determine gameplay in Ensemble; participants were encouraged to choose a work with which they were already familiar as the basis of their coding during the workshop week. (See, in particular, the contribution by Julien Le Goff, whose research on dramatic anecdotes was of particular relevance to our goal of capturing behavioural norms amongst theatre audiences [Moulin et al. 2023].) However, the choice to work from literary sources does not reflect a naive conviction that these works can provide contemporary scholars with a precise image of the past. Rather, we are extending methodological choices elaborated during the composition of the VR architectural restitution by Paul François, whose main visual source contains its own fair share of distortions. This painting, a tiny 1779 miniature by Louis-Nicolas van Blarenbergh adorning a jewel-encrusted snuffbox, tells us at least as much about the market for luxury goods in the latter eighteenth century as it does about the exact interior disposition of a Fair theatre (see **Figure 1**). The scarcity of documentation concerning these venues, as compared to other theatres, obligates us to explicitly foreground the biases and lacunae inherent in our sources. To that end, Dr. François authored a Unity plugin called Prouvé, which allows users to access research from within the immersive digital model to help understand or explain the visual representations that surround them (see François et al. 2021).

This approach illustrates a fundamental design principle of the VESPACE project called "image depth"—a depth, in this instance, that is epistemological rather than spatial or perceptual. Deep images, whether referring to a visual representation (as in the VR model) or a social representation (as in the Ensemble-based rules), allow users access to underlying research in order to better assess the relationship between the archival record and the technological processes capable of creating highly persuasive digitally mediated models. While a historical object like the Fair theatre calls out for precautions of this kind, we feel that image depth is a principle of ethical design with very broad applicability across digital humanities applications and projects—even

those that do not have mimetic visual representation as their primary output—and as such it also informs the work we are undertaking with Ensemble to build out a social model of the Fair theatre.

Digital projects often benefit from auras of authority and objectivity reflexively assigned to computer-based research outcomes, especially in the case of an immersive VR simulation whose sensory illusions can trick even the most self-aware users into perceiving the projected space—and thus the historical theory that underpins its restitution—as factually real. This overwhelming impression of presence is perhaps the most immediately obvious advantage to the VR medium but is also a potentially serious drawback for users who assign truth-value to its representations based solely on sensory inputs. A successful game illustrating a set of social conventions about a historical period runs the same risk of displacing the critical skepticism appropriate to a scholarly hypothesis with an impression of truth. Crucially, this is the outcome we seek to avoid. In its place, our project aims to instill practices of ethical computer model design that allow for and encourage dialogue about methods, materials, and results as a way to contribute to a twenty-first-century digital humanities deontology that acknowledges and embraces the necessary place of subjectivity in a field that often swathes itself in a misleading aura of computer-generated objectivity.

Explainable AI and social physics

Though the study of artificial intelligence is relatively nascent, its advances have shaped the world. Through both real-world technological developments and in the spectre of media depictions of "rogue AIs" like Hal 9000 and the Terminator, the boons and banes of AI advances are affecting everyday life at an impressive clip. Though a complete history of AI is beyond the scope of this writing, a brief introduction to some of the outstanding goals of AI research, as well as to the sometimes synergistic (and, at times, competing) notions of "statistical" and "symbolic" approaches to the field, can help the reader better understand what social physics—VESPACE's approach to autonomous agents—is, and what it is not.

Artificial intelligence has long pursued the goal of creating rational agents: autonomous beings—either virtual or physical—that have the capacity to sense their environment and act upon and within it in ways appropriate to their form and function. Such agents perform myriad tasks, from providing tutoring assistance on a specific subject, to assembling cars on a factory line, to curating social media feeds with content predicted to maximize engagement and the content generation provided by applications like ChatGPT and DALL-E. One area that has proven to be a valuable space

for breakthroughs in artificial intelligence techniques is in games. The constrained environments of games (or game-like experiences, such as the scenarios outlined in economic game theory) make them ideal for focusing on solving specific problems in well-defined spaces, which can then be generalized and relaxed to apply to problems in the real world. While highly mediatized examples of this work include IBM's chess-playing Deep Blue and Google's AlphaGo each defeating the human world-champions of these games, one of the most original uses of AI in gaming was demonstrated in the landmark interactive drama *Façade*, whose system enables the characters Grace and Trip, inspired by the toxic couples in Edward Albee's play *Who's Afraid of Virginia Woolf?*, to dynamically respond to player input in a believably lifelike manner (Mateas and Stern 2003).

The above examples contain a mix of both "statistical" and "symbolic" approaches to artificial intelligence research. Popular successes like AlphaGo have highlighted the ludic potential of statistically based "deep learning" AI techniques that are also increasingly integral to advanced industrial practices. Borrowing metaphors from human cognition such as "neural networks," these approaches are examples of machine learning, whose algorithms apply copious amounts of training data in order to build a statistical model that can then be applied to future data to predict outcomes. What the model "learns" from its previous experiences informs its future decision-making, producing outputs based on prior data that can yield spectacular results. At the same time, the insatiable appetite for training data makes this approach vulnerable to perpetuating and reifying biases in ways that have potentially damaging real-world consequences (Dastin 2018). In contrast to the statistical nature of machine learning, the symbolic approach to artificial intelligence is founded on using discrete logical calculus to represent the world and solve problems within it. (Many terms beyond "statistical" and "symbolic" have been used to describe these schools of thought; see Minsky 1991.) Symbolic AI avoids some of the issues encountered by statistical approaches, as it requires less training data, and inherent bias is far more easily traced back to specific authorial decisions and thus more readily correctible. However, symbolic AI "learns" in a different way from statistical systems, making its sphere of application distinct from that of statistical approaches. For machine learning, the power of the system comes from the ever-increasing amount of information fed to it, whereas the dynamism of symbolic AI systems derives from their ability to creatively refigure the logical relationships between agents based on a static data set. If machine learning offers the possibility of modelling *the* world, at the potential cost of reinforcing its preexisting inequities, symbolic AI holds out the promise of modelling *a* world, offering up perspectives on how to navigate complexities—including data bias—on a human scale.

The enormous complexity of machine learning and other statistical AI models underscores the fundamental problems inherent to an artificial intelligence system that is not *explainable*. These problems can be understood from the standpoint of ethics, as well as from the perspective of the system's expressive aims and abilities. Despite the primordial role of human contributions to programming, a common fallacy of computing holds that "machines have no bias," resulting in undue authority granted to the oracular outputs of complex systems. In reality, developers' conscious and unconscious partialities are still baked into programming procedures. In response to systems built on a lack of transparency, research communities have recently turned toward "explainable AI" to grapple with problems of unrecognized biases being represented as reasonable and eventually hardening into real-world "truths," advocating instead for systems that provide straightforward means to "look under the hood" and investigate and question the machine's decision-making process.

Explainable AI is also valuable for another reason: not only are results easier to read and understand by humans, the authoring process is also made more comprehensible and accessible to a wider range of people. Given the VESPACE project goal to facilitate meaningful contributions to digital humanities projects by non-technically trained domain experts, focusing on an explainable, intelligible approach to AI was a natural choice. However, this intelligibility can come at a cost to efficiency: instead of training a model to determine processes automatically, humans must take responsibility for delineating the rules and decision-making processes of these systems. Moreover, symbolic approaches to AI acquire information differently, requiring more pre-authored content. While more onerous from a labour standpoint, transparent human authorship makes it far easier to identify and track down the source of bias in rules and processes (a traceability that is especially important to the VESPACE project, which places significant value on representing historical objects according to scholarly consensus). This hand-authoring approach also makes AI systems broadly available for bespoke applications, not least in gaming and other creative fields, for which efficiency and rationality may be less important goals than believability and emotional engagement.

The authoring challenge for symbolic systems can be daunting, not least because there are far too many contexts to be able to account for every contingency, and a human author attempting to write everything would never be able to achieve completion, as new situations constantly arise. In response to the impossibility of completely capturing the world in data, social physics was developed to allow for both an explainable mechanism in which a human interactor can examine the underlying rules that govern the autonomous agent's behaviour, as well as providing a way for domain authors and experts to contribute relevant knowledge in a structured format

that could be understood and parsed by a computer system. Inspired by the theories of social psychologists Eric Bern and Erving Goffman, social physics furnishes a computational representation of what Goffman terms "dramaturgical analysis," which holds that a single individual might inhabit slightly different versions of themselves based on shifting contexts, just as a single actor might adopt multiple roles in a theatrical production (Berne 1964; Goffman 1978). The choice to work with a Goffman-inspired model diverges from the "Big Five" personality traits, a five-factor model that has been successfully used in other systems, including the natural language generation tool PERSONAGE and the social simulation system Talk of the Town (see Digman 1990; Mairesse and Walker 2007; Ryan, Mateas, and Wardrip-Fruin 2016). Social physics designers felt that the Big Five (which represents personality as a set of scalar values, each representing how significantly a particular key trait applies to them: extraversion, openness to experience, consciousness, agreeableness, and neuroticism) did not result in sufficiently nuanced characters. Social physics instead adapts Goffman's theory that individual behaviour is influenced by both social needs and the nuances of current context. A character's location, the other people in their company, and their current feelings all influence the consequent behaviours of an individual.

Social simulation design with Ensemble: Rules authoring

The desire to be able to represent and model a diverse array of characters using descriptive language, and for any given character to be able to behave consistently "like themselves" while allowing for appropriate variation according to differing circumstances, drove the development of the precursor of the Ensemble system, Comme il Faut (CiF), the first social physics engine. (For preliminary work on the system and its inspiration in Goffman's theories, see McCoy and Mateas 2009.) A common problem encountered in previous systems relates to character inability to respond appropriately to situations that were unanticipated by the programmers, as highlighted by Noah Wardrip-Fruin in his discussion of quest and dialogue logics, commonly used tools for telling interactive narratives (Wardrip-Fruin 2009, chap. 3, pp. 58–69). By contrast, characters in CiF could be described and discussed using a multitude of terms and notions, taking into account many different types of social affiliation with other virtual agents within the system.

The atomic logical building block of this system (and others like it) is the *predicate*; a single predicate represents a single fact about the world. These facts—and thus, these predicates—can speak to myriad aspects of any individual virtual agents and the relationships they might hold with one another. For example, some predicates might detail immutable character traits that are applied to an individual. Others speak

to how characters relate to one another within social cliques and "networks" across a variety of dimensions (such as friendly affinity, enmity, or romantic interest). Within any given predicate lies room for additional granular specificity; some predicates, such as the aforementioned character traits, are represented as Boolean values (true/false statements—either a character has a trait or does not), while other values are scalar, giving characters an ability to recognize just how strongly they feel about each using a numeric range (for example, a middling enmity score might signify simmering tension, while a high score bespeaks outright hostility). In addition to modelling characters and determining how they wish to engage with one another, social physics systems also elegantly handle the cascading aftermath of any transpired social exchange. For example, if two characters just completed a messy divorce, social physics allows the system to represent the animosity between the exes, as well as the impact of this new situation on their friends and acquaintances throughout the world. CiF was designed to model these repercussions in a manner analogous to the computer representation of real-world kinetic phenomena—balls bouncing, water falling, cars exploding—that are governed in video games by a "physics engine." In turn, the "social physics engine" determines how multiple social forces collide and interact to form the sometimes-contradictory social desires that inform human engagement with the world.

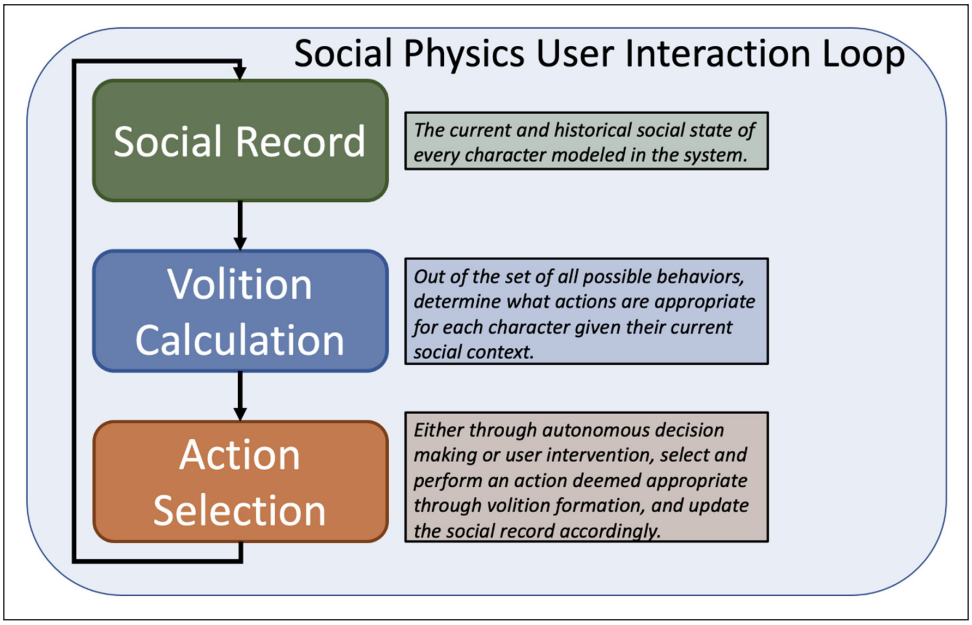

Figure 2: The social physics user interaction loop. The social record contains all predicates detailing every atomic "fact" about the characters of the simulation. These predicates are fed into the volition calculation process, which determines the types of actions each character wishes to take. From here, one such action is ultimately taken, which then adds new (and updates existing) predicates in the social record, which in turn will yield new character volitions.

Social physics systems, then, depend largely on a user interaction loop that includes three processes (see **Figure 2**). First, it maintains a social state (internally referred to as the "social record"), which is a collection of current truths about the world represented as logical predicate data. Each predicate in the social record represents individual characteristics of agents and their attitudes and feelings towards one another. Second, it uses "social rules" to compute character volition values, the type of behaviours that characters wish to engage in to fulfill their social needs. Third, it allows characters to select, and ultimately undertake, "actions" or "social exchanges" to then actually fulfill the needs determined in the previous step. Taking an action can affect the social state in any number of ways (e.g., former enemies are now friends, and their friends are consequently friendlier as well), which then sets the loop in motion again. Rules and actions are authored not with specific characters explicitly stated, but rather with noncommittal "roles" that, in theory, any character could potentially "bind" to. Thus, to author a rule in social physics is not to determine the behaviour of an individual character, but rather to project a possibility space containing the potential thoughts, feelings, and actions available to all characters in the system (see **Figure 3**).

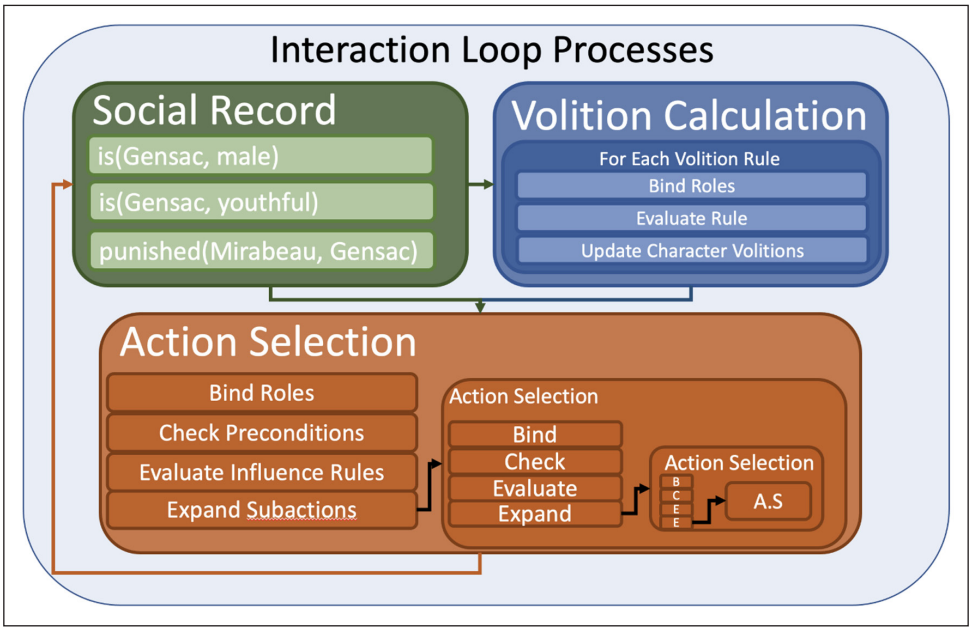

Figure 3: Interaction loop processes. This image further details the specific processes that occur within each of the three phases of the social physics user interaction loop. The social record is full of predicates that describe the current social state of the world (here, Gensac is a youthful male, punished by Mirabeau). Volition calculation loops through each volition rule, binds specific characters to the rules' abstract roles, and determines if the rule holds true for that set of characters. If so, it adjusts the volition scores of the bound characters accordingly. Lastly, action selection follows a similar structure, in which characters determine what acts they most want to perform based on their previously calculated volition scores. This is a recursive process, as actions in Ensemble can have arbitrarily deep hierarchies of sub-actions, with each layer capturing more specific or nuanced types of behaviour.

The CiF social physics engine has been the basis of a variety of projects beyond the flagship title *Prom Week*, including the mystery game *Mismanor*, the DARPA-funded training simulation IMMERSE, and as a mod for the Bethesda Softworks blockbuster *Skyrim* (Sullivan et al. 2012; Shapiro et al. 2013; Guimarães, Santos, and Jhala 2017). A second-generation social physics engine, dubbed Ensemble, was developed in 2015; while similar to its predecessor CiF, Ensemble offers significantly more authoring flexibility. Notably, whereas CiF enforced the use of certain schema elements (such as the "traits" and "networks" described above), Ensemble enables the author to specify their own schema elements. Rule and action authoring remains largely the same between the two systems. (For more on these elements of authoring in social physics, see the discussion of workshop methods and materials below.)

Social physics, then, does not aim to tell characters "what to do," but rather allows authors to determine character attributes and allow them to interact with each other according to "social rules" that take the form of predicate logic statements. In conjunction with a knowledge base that stores the current and historical social state of the world (social record), these social rules allow characters to form desires or "volitions" for the kinds of actions they would like to take with one another. Incorporating the previously defined traits and networks that describe individuals and aspects of their social relationships, any written rule can apply to any character if they meet the criteria of the predicate. What results is a classic rule-based AI system combined with utility scores. (Utility scores, sometimes known as "fitness scores" or "costs," are a frequent feature of AI systems in which an agent's choices are ranked according to expected value, allowing for evaluation of multiple choices that may be presented in a given situation.) Unlike work done in the branch of AI design known as "planning," in which agents form complex chains of action in order to affect the world to achieve particular goal states, agents in social physics systems consider a wealth of information to determine the various actions they can take in their present moment to satisfy their social needs. To be clear, the social physics approach and the work done in the planning community are by no means mutually exclusive. Though a project integrating these approaches remains future work, one could imagine a system where agents maintain and pursue long-term goals (as in planning), but in which the decision-making to achieve said goals isn't based purely on efficiency, relying instead on the wealth of social and cultural considerations that social physics systems such as Ensemble provide (see Riedl and Young 2010). Agents in social physics are beset by a multitude of social and cultural forces, at times pulling them in different directions. These conflicting forces can result in ambivalent feelings on the part of the agent, modelling situations akin to those found in fictional narratives where characters are forced to confront their values, and the values of the environment in which they find themselves, in order to arrive at decisions that drive the action. In this way, social physics

provides a compelling basis for interactive experiences, as character choices are not pre-selected by authors, but rather emerge through intersecting rules. This enables the player to make choices that may exercise an unpredictable influence on character fates, situating the user as co-author of an ergodic text crafted in collaboration with the computer system.

Playable experiences for humanities research

One of the major goals of VESPACE is to reconceive how humanities research is communicated to its various publics. Writing up results of research and analysis in the form of essays and reports—as in the case of this essay and the contributions that follow—is a standardized academic process whose rhetoric, vocabulary, and citational practices result from centuries of Western intellectual tradition, a set of practices applied to both textual and non-textual objects of study (see Leichman 2021). Theatre history, which sits at the confluence of literature, material and economic cultures, visual arts, performance, economics, and sociology, illustrates both the advantages and disadvantages of the exclusive reliance on discursive outputs to further research in the field, with textual description in constant pursuit of visual, aural, and experiential phenomena that lie outside the logosphere. The variety of archival material in theatre history makes it an ideal field in which to test our hypothesis that new technology can help contemporary scholars break free from the constraints of strictly textual research outcomes. Our challenge is to show that these new forms can both adhere to the common standards for transparency and documentation in advancing original research, while also broadening the conversation to include students and researchers whose extra-academic lives are saturated with interactive computer-based interfaces even as their education and research often remains exclusively tethered to print culture.

Creating historical scholarship for an interactive simulation involves reorienting expectations of both authors and readers. From a research perspective, fitting evidence into narrative form has always been the historian's great challenge, not least in finding the correct balance between detail and overview that allows readers to understand the particularities of the subject, as well as its place within larger historical developments. Seen this way, the reader occupies both the top and the bottom rungs of the intellectual ladder: access to an ordered, hierarchized, and (traditionally) linear story affords a sense of mastery over historical truth, at the same time that the singular path traced by the author from the beginning to end of the argument situates the reader as primarily a recipient, open to the authorial manipulations of narrative and perspective that comprise the novelist's—and the historian's—rhetorical arsenal. Style, in this formulation, becomes a crucial index of success, underscoring the tension between fidelity to the archive and the demands of a narrative charged with revealing the historian's theory of causality and of the repercussions of historical phenomena. Popular histories lean into the emotional

engagement permitted by story form, while academics often defer to the ambiguities of the archive, sacrificing the comforts and certainties of omniscient narration in favor of a style whose dryness can also be read as signalling seriousness of purpose.

By contrast, the work that is written and the work that is read are more clearly distinct, and interdependent, when writing history with the interactive simulation engine Ensemble. The outcome cannot be consumed in a linear fashion, as with a text that introduces an argument, develops it, and retrospectively inserts it into an existing body of knowledge that is also constructed of the same linguistic and grammatical materials. Rather, this kind of simulation is "read" through the performance of the gamer, its theory of sociability presented inductively through the choices faced when controlling a character during gameplay. While in a limited sense, no two readers of a traditional historical text will have the same experience, video games present a far starker illustration of the indeterminacy of "reading": access to the historical text, encoded in the rules, depends entirely on gamer choices, making identical experiences extremely rare. All readers will presumably read the top of page three after finishing the bottom of page two, receiving information according to the way in which the historian has ordered it. On the other hand, the choice to greet someone heartily or haughtily can lead to widely diverging experiences in a performance-based computer model, reflecting a degree of freedom that has helped to make "ergodic literature" a compelling and immensely popular way to engage with fictional narrative and an intriguing medium for constructing academic arguments. (Ergodic literature, a term coined by Espen Aarseth [Aarseth 1997], refers to narratives that require "non-trivial" effort to understand; in the case of video games, this refers to the decision-making and input requirements of gamers that allow for the story to advance.)

The historian, rather than writing a linear, progressive argument, sets out to define the conditions of the video game world such that the gamer's experience of a situation or exploration of a narrative furnishes a demonstration of the historical hypothesis through a series of individual choices in relation to the evidence-based attributes of the environment or other characters. Concretely, this involves authoring rules for the Ensemble system that reflect research findings. Rather than writing sentences contributing to a narrative, the output is mapped onto the rich intersection of mathematical and discursive reasoning emblematized in the conditional phrase ("if . . . then . . ."), which comprises the first-order predicate logic of the social physics engine. The system evaluates these conditions in order to determine the actions and reactions of the non-player characters (NPCs) who populate the simulated environment. The user controls one character, but the machine controls all of the others, imposing the social and communicative expectations of an interactive simulation on the historian's interpretation of the archive. The human interactor reads the text by making decisions

on behalf of a character, each one unlocking a set of potential responses grounded in historical research, allowing for first-person emotional investment while also ensuring that the feedback from the model adheres to the scholar's understanding of the evidence. While no two play-throughs are ever likely to be the same, it's also true that no single play-through is ever likely to be complete; player choices will determine which rules fire, meaning that a full understanding of the proposed theory would require multiple experiences giving rise to different combinatorics (see Leichman 2022).

The VESPACE social physics workshop: Materials and methods

The workshop, convened in May 2020, was the culmination of a six-month development process that resulted in a new authoring tool and a significantly revised understanding of the workflows and challenges related to integrating traditional literary-historical research into a digital simulation paradigm. (See DeKerlegand, Samuel, and Leichman 2020. The May 2020 workshop was the fourth, and most complete, iteration of the VESPACE effort to introduce social physics coding to students of literary history.) Starting with initial discussions around leveraging the computing metaphor of social physics for the VESPACE project in 2017, a clear part of our interest in this work was to test the validity of this methodology for literary research. The decision to work with graduate students in the prototyping phase reflects younger scholars' increasing familiarity with digital interfaces, as well as their being more likely to seek out and embrace new tools and methods for conducting their research and teaching. (See, in particular, the contributions to this special issue from Chiara Azzaretti and Charlee M. Bezilla, which detail their reservations around digital methods, and the ways in which the social physics workshop experience changed these perceptions [Moulin et al. 2023].) The rapid onset of COVID travel bans and lockdowns shifted the logistics from an in-person meeting in May at LSU in Baton Rouge—the loveliest season on our campus—to an online meeting, with a total of 53 hours of Zoom-based working sessions attended by seven graduate students and two faculty members representing seven institutions in the US and Europe. During this workshop, participants learned how data is organized in Ensemble, how to set up a social world, and how to code and document information from literary sources as the basis for historically and culturally plausible interactions between computer-controlled NPCs and a human player within an interactive model of the eighteenth-century Fair theatre in Paris.

Our meetings took place via Zoom, with three sessions per day; European participants (plus project leads) worked together in the first session, all participants collaborated in the second session, and American participants (plus project leads) logged on for the final session of the day. In the week prior to the workshop, organizers posted a series of six videos explaining the background of the VESPACE project, the principles

of AI and the development of social physics, and the basic look and functioning of the custom authoring tool; offloading this information for asynchronous consumption was also an important means to lessen the cognitive load of the workshop week and preserve our time together for interaction and discussion rather than lecturing. While video conferencing was an important aspect of our collaboration, most of the actual work took place through the authoring tool, adapted by Daniel DeKerlegand from a standalone Node.js javascript application to a web version hosted by Heroku, leveraging a central, collectively editable MongoDB database (on the influence of literary scholarship on authoring tool design, see Mr. DeKerlegand's contribution to this special issue [François and DeKerlegand 2023]). The ability to seamlessly and simultaneously update the database through a user-friendly authoring tool was a critical component of our success in maintaining attention and energy in an all-online workshop. In the middle of the workshop week, a significant build-out of the tool allowed us to work on social exchange authoring and to develop a character set.

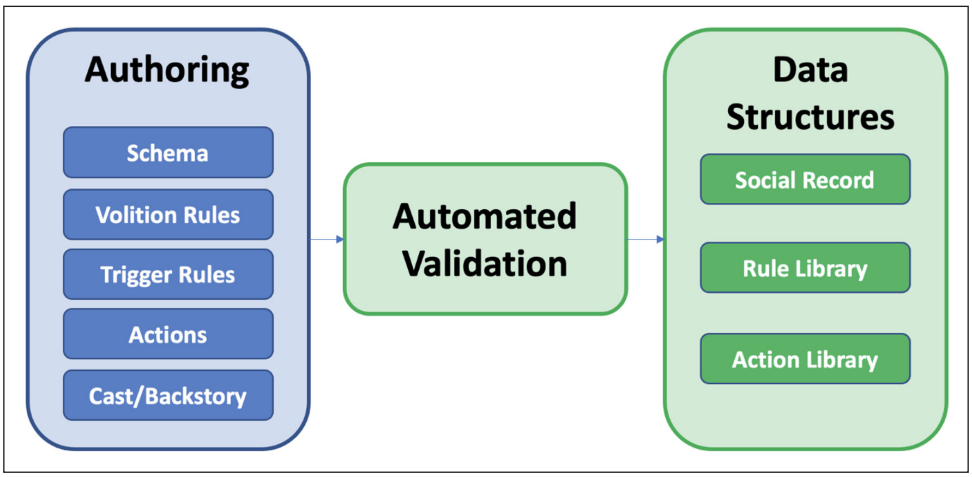

Figure 4: Authoring for a social physics system. The human author is responsible for specifying the components contained within the authoring module. The social physics engine is then responsible for validating the authored content for proper formatting, and then automatically populating its internal data structures with the validated authoring.

The workshop itself revealed the full complexity of writing history in this novel format (see **Figure 4**). The rules-authoring process can be broken down into three major phases, each dependent on the previous for its logic and coherence: schema authorship, volition and trigger rule authorship, and social exchange authorship (also referred to as action authorship). Additionally, social physics authors specify the characters that populate their virtual world and their relationships with one another, effectively establishing the backstory of these agents. The following paragraphs explain the principal challenges related to authoring schemas, rules, and actions in Ensemble.

The first phase, schema authorship, allows the historian to define the basic conditions of the social universe to be modelled. This involves determining what kinds of characteristics exist in this world; if there is to be friendship or jealousy, schema authorship is where these qualities must be outlined and categorized. Each category of characteristic is defined according to four factors that determine 1) its temporal aspect; 2) its capacity to affect other characters; 3) its mathematical representation within the system; and 4) its availability and effects during gameplay. Respectively, these qualities identify a characteristic's duration, directionality, value type, and actionability.

Some characteristics inhere within a person at all times, whereas others represent transitional states that might be true or false depending on when one asks the question. Being tall does not change over time, but jealousy over someone else's accomplishments might. Similarly, some characteristics are undirected or self-directed (being tall) while some are other-directed (jealousy) or reciprocal (friendship). Mathematically, a characteristic like being tall can be represented through a Boolean (true/false) value, whereas something like jealousy might be better captured through a scalar value, situating the intensity of the feeling along a numeric scale. Finally, actionability is an essential factor for Ensemble, defining those categories that one could plausibly aspire to affect or achieve in the course of gameplay; to return to previous examples, one can't do much to achieve being taller, but one could plausibly take actions to increase another's jealousy, situating the latter as a potentially actionable characteristic. (In our workshop, we defined six categories of characteristics, a schema architecture inherited from the most complete previous game to use social physics, *Prom Week*, developed in 2012. See Samuel et al. 2015.)

Schema authorship introduces many of the important ideas around computational thinking that represent a significant contribution of this kind of work to humanities research, as well as one of its signal challenges. A written output can rely on certain shared assumptions that the author is not required to explain; while the details may vary, it's safe to assume that jealousy and height don't need to be carefully defined for most readers. Computers, however, need to be told the exact contours of the universe every time. Moreover, the way in which the computer comes into this knowledge is through logical and mathematical values rather than grammar, with a corresponding rigidity (or, alternatively, clarity) that defines its role: jealousy is an object in a class that will function just like other objects in its class. The simulation acquires depth and nuance through the rule-based interactions of objects, rather than through variety within any individual object.

This workshop experience revealed the degree to which effective authorship at one phase of the rules-authoring process required an understanding of the downstream effects of the rules. Only *after* writing volition rules does the true importance of the schema emerge, just as the purpose of the volition rules, and thus the knowledge of how to orient

them in order to foster a satisfying playable experience (and one that reflects the theory of sociability the game seeks to advance), becomes clear upon beginning work authoring social exchanges. While the same could be said of writing a traditional text, humanists have long experience of ordering words into sentences, sentences into paragraphs, and paragraphs into arguments, whereas thinking in terms of computational stacks is often a new skill for literary and historical scholars. Learning curves are necessary and useful in all fields of endeavour, but streamlining the familiarity with the logical relationships between the three phases of rules-authorship in Ensemble remains an important goal for the broader application of this method to humanities research.

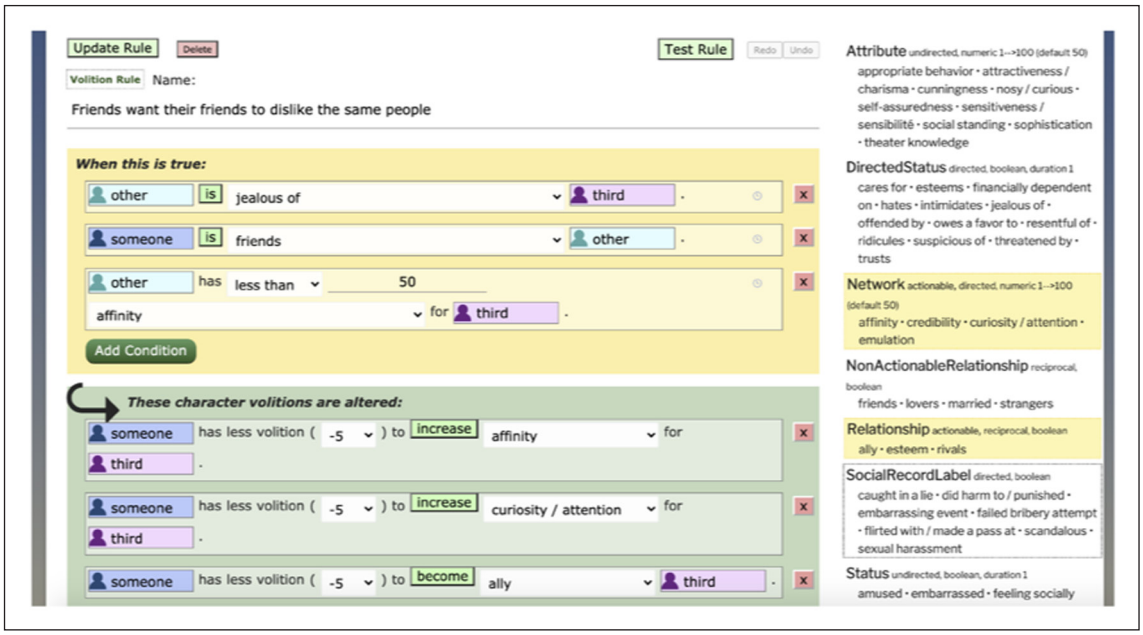

Figure 5: A snapshot of the Ensemble Rule Editor for the volition rule "Friends want their friends to dislike the same people." This rule has three precondition predicates: "someone" and "other" must be friends, "other" must be jealous of "third," and "other" must have less than 50 affinity for "third." If these conditions hold, then the volitions of "someone" to increase affinity and curiosity/attention and to become an ally drop. The noncommittal role names of "other," "someone," and "third" remind us that any set of three characters for whom these conditions hold will cause this volition rule to fire.

The next phase of rules-authoring in Ensemble relates to volition and trigger rules (see **Figure 5**). Of these two, volition rules are more complex, and constitute the overwhelming majority of the authorship during this phase. (Trigger rules operate a direct effect on the social state, bypassing volition formation and social exchanges; because of their very direct mechanism of action, they must be used sparingly in order to avoid unintended consequences.) Volition rules are a unique feature of social physics,

allowing for characters to form desires rather than simply respond to given set-points in pre-scripted ways. Volitions can be thought of as emotional vectors, endowed with both a directionality and a magnitude, and are susceptible to multiple influences (an individual's personality, current mood, relationship to interlocutors, etc.) that can exercise a different overall effect as they encounter other competing volitions at different moments in the character or game history. A volition rule takes the form of a conditional statement, "if x then y," in which the left-hand side of the equation (the "if" statement) consists of an element or a combination of elements from the schema that defines characters who might interact over the course of gameplay. Each volition rule is meant to capture a socially plausible desire for a given historical context, written in the form of predicate logic. If all of the predicates of the left-hand side evaluate to true, then a character's volition is influenced, altering the emotional vector in terms of magnitude, direction, or both.

Executed volition rules reveal a character's "intent," which is an index of their desire with respect to a schema category defined in the schema as "actionable." Each intent in turn contains social exchange (action) categories that are themselves populated with instantiations, the most granular level of social exchange authorship. (In our workshop, instantiations are also anchored to textual examples in order to ensure coherence within the universe modelled on these literary sources.) Thus, from the relatively small list of schema types, a huge number of volition rules can be created; the outcomes of these are funnelled through the small number of actionable types to become intents, for which a large number of social exchange instantiations can again be authored. At this point, player choice intervenes to determine the course of the simulation: based on the results of volition formation, the player selects a social exchange for her character, and can observe how this choice ricochets through the social universe of the game.

First-time rule authors often try to code the exact scene from the source they are using, but rules authorship is above all an interpretive gesture in which the literary instance serves as the basis for an abstraction that can be applied across different specific situations. For example, rather than writing a rule for the character Candide, one might write a rule for a character ("someone") who is innocent or unworldly; similarly, his interlocutor would not be Cunegonde, but rather a character ("other") who is beautiful. In using literary sources to code social influence rules in Ensemble, we have situated these works of fiction as an archive of plausible (*vraisemblable*) interactions that would have been comprehensible to readers at the time; indeed, novels and plays have long been a valuable resource for historians seeking to illustrate the sociability of bygone eras. The goal in using these sources in Ensemble is similar: not to make a game of Voltaire's *Candide*, but rather to create a game universe where the interactions in *Candide* are able to take place.

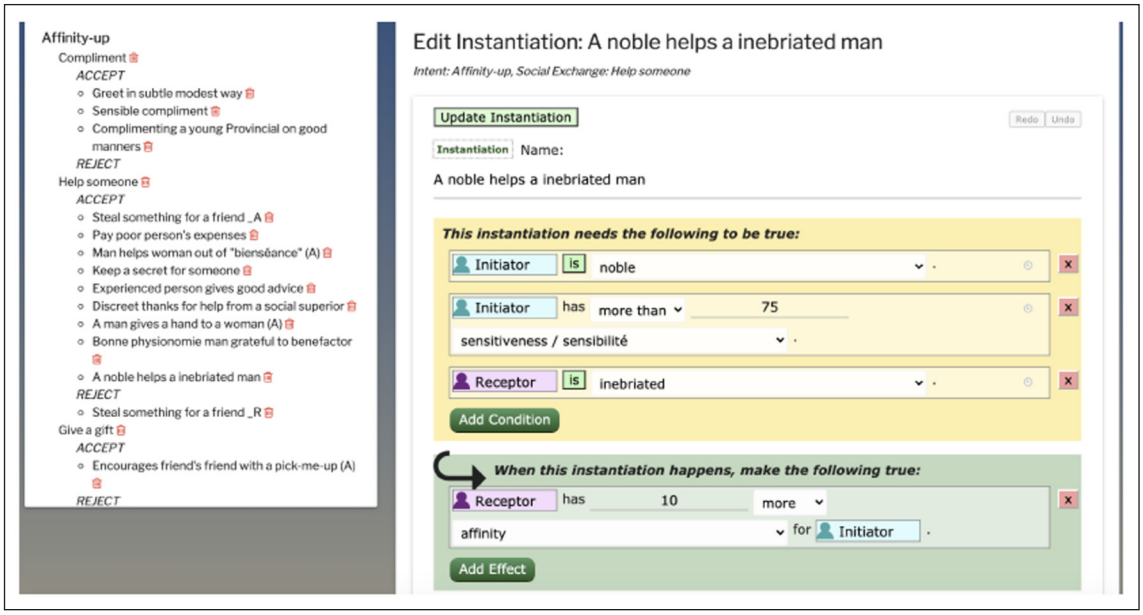

Figure 6: A snapshot of the Action/Social Exchange Editor. This depicts the preconditions and effects of an instantiation for the "Help Someone" action, in which a sensitive noble assists someone inebriated and receives their affinity in return.

The third and final mode of authoring in Ensemble is social exchange authoring, also called action authoring (see **Figure 6**). Social exchanges complete the circle, connecting the other two modes of authoring by providing an expressive output for character intent as it arises from the original building blocks of the schema and the volition-formation rules built from these elements. Social exchanges, in turn, give characters the capacity to act upon their volitions and in so doing, affect interlocutors and themselves. The resulting alterations in the entire social world in turn form the ground state for the next round of volition formation, triggering Ensemble's foundational AI loop, allowing the system to respond to, and help to create, emergent situations.

In turn, instantiations—actions selected by human interactors based on choices made available through system calculations—can have one of two outcomes: "accept" or "reject." Simply because one character wants to engage in a social action with a second does not mean that the action will go through as the initiator originally planned; aggressors can be met with pacifist responses, romantic propositions can be rejected. Mechanically, an action will be rejected if the volition of the responder (i.e., the target of the initiator's actions) towards the initiator is less than zero for the intent of the initiator's action. Regardless of whether the instantiation is labelled as an "accept" or as a "reject," the result of these rules (right-hand side) specifies how the carrying out of an instantiation affects the underlying simulation's representation of the social state

of these virtual characters—how they feel about one another and themselves, similar to the structure of volition rules. In a game situation, a completed instantiation affects the current social state of the world, adjusting the values of the elements defined in the schema for all characters, which in turn affects which social norms now apply to them (i.e., which volition rules hold true), establishing the ground state for the next set of actions they wish to take with one another, and initiating yet another iteration of the AI loop.

This last level of authorship is the most perceptible for naive users, who understand the Ensemble-driven experience as a string of instantiations triggered by player choices, just as a theatrical piece is composed of a string of finished scenes for an audience unaware of the dramatic structure that supports the sequencing of onstage action. The layer of social exchange authoring within Ensemble is intentionally "thin," inasmuch as it focuses on the prerequisite social state for an exchange to take place and its resulting effects. Social exchange authoring does not focus on the social exchange's performance realization. Potential realizations might include lines of dialogue, choreographed movement, musical excerpts, or other symbolic representation of emotion that would be perceived by the player; these presentations are the subject of a separate authorship process. (One potential benefit of the careful literary documentation is the possibility of procedurally deploying this period-appropriate text as the basis of instantiation realizations, potentially easing the authorship burden while also affording access to authentic eighteenth-century prose.) While allowing system designers to focus on authoring at the simulation level, this again demands that authors work towards certain abstractions, in which their focus is not dialogue, for example, but the underlying effects of potential dialogue (or dance, or music, etc.) on the characters involved, with the understanding that the actual text itself will be produced at a later date. While beat-by-beat and scene-by-scene moments of a play might explain the fluctuating fortunes of characters, the story would likely feel overly mechanical if dialogue expressing these reversals were simply replaced with a list of the reversals themselves, a log of the underlying state changes of the dramatis personae—which is nonetheless the task that social exchange authors face.

Discussion: In the shadow of *Prom Week*

The first game created using CiF (Comme il Faut, the progenitor of the social physics engine Ensemble that shares many of its fundamental metaphors and processes) was *Prom Week*. *Prom Week* was a labour of love designed and developed by a large team of graduate students and undergraduate researchers with the goal of making a personally and emotionally engaging AI-driven game experience. Players of *Prom Week* frequently

cited an empathetic connection with the characters powered by this system, an attachment that made them feel bad about manipulating these characters into taking certain actions, a gameplay experience that elicited many moments of self-reflection and introspection on the part of players. At the time of its release in 2012, *Prom Week* was recognized as a technical achievement in artificial intelligence in research and independent game-developer circles, and has remained the primary case study for use of the underlying social physics technology.

The clear advantage of this kind of precedent is that it furnishes proof that an interactive AI-driven game can be created using this technology, and that it can provide meaningful experiences to players. However, VESPACE developers also need to assess the relevance of this example, which draws on similar technologies but to very different ends. Not only are the historical and geographical settings of these two experiences widely divergent (in lieu of recreating the eighteenth-century Paris Fair theatre, *Prom Week* instead transpires in a fictional early twenty-first-century American high school), but the contextual foreknowledge that players bring with them into these experiences varies greatly as well. Most American players of *Prom Week* could leverage personal anecdotal experiences of various high school rites of passage, as well as the mediation of these experiences in film and TV, to inform their play. The game's characters and social rules specifically made use of many of the classic tropes of fictionalized—and highly dramatic—high schools in order to enable players to immediately ground their understanding of the characters and their relationships to one another. Save for scholars of the eighteenth century, most players of the VESPACE project will have no such pre-existing knowledge. VESPACE then, inverts the player formula; instead of leveraging one's own anecdotal understanding of a time, space, and culture to inform play (which then, through simultaneously reinforcing and subverting these understandings, can lead to self-reflection), it instead expects the player to enter the world with limited understanding, and to fill in knowledge inductively through engaging with the characters and watching them engage with one another. The same contemporary sensibilities that assisted players of *Prom Week* are fully expected to clash with the social expectations coded into VESPACE rules; whether this clash leads to greater insight and appreciation of the system, or frustration with it, remains to be seen.

Beyond the player experience, the authoring lessons of *Prom Week* still echo within Ensemble itself, influencing the organization of rules in the VESPACE workshop as well. One concrete example is the decision to retain only three types in the "relationship" schema category (one of only two actionable categories, along with "network," for the workshop schema), which reflects a choice initially made for *Prom Week.* In *Prom Week*, this decision was based in part on ethnographic analysis done on the media sources

that inspired the world of the game, but the decision to create only a small number of actionable types in VESPACE was much more motivated by game design considerations: three relationships seemed like a small enough number of concepts that players could keep track of them in their mind's eye. Relationships appeared the most consistently across *Prom Week*'s volitions rules, and thus keeping the number of relationships low was in part inspired by the desire to reduce authorial burden and the combinatorial explosion that can result when introducing even just a single new element. Likewise, the Boolean aspect of "relationship" lends itself to a notion of quantifiable, cathartic achievement in which making changes to the behaviours that the characters within the newly forged relationship wish to take with each other produces easily perceptible effects. This emphasis can be felt across the volition rules, social exchanges, and instantiations, in which relationships figure prominently in character volition formation and the rules that result from this. In the general flow of gameplay in *Prom Week*, characters mostly engage in actions that affect the scalar, directed, permanent networks of "buddy," "romance," and "cool." These network values capture small fluctuations in characters' attitudes towards one another, eventually building up enough shared history between characters to allow the player to attempt to cement one of the three official "relationships" between a pair of characters (Friends, Dating, or Enemies), which in turn reflect the quantifiably cut-and-dried and alarmingly public "relationship statuses" characteristic of modern social media platforms. In this way, *Prom Week* attempted to capture a satisfying progression of a slowly evolving story involving multiple characters over time, allowing the player to guide them into a new relationship dynamic and consequently be rewarded with the opportunity to explore a wealth of new behaviours between them.

This example illustrates how further development of social physics in the VESPACE project (or similar DH initiatives) will have to abandon some of the structures that underpin the success of *Prom Week*. Ensemble, unlike CiF, allows for a nearly limitless degree of flexibility in schema definition, but for the workshop we imported the category setup from *Prom Week* almost unchanged, in part because these elements had already served as the basis of a satisfying game experience. But what game experience? Here, the vast discrepancy in the objectives of *Prom Week* and VESPACE would seem to argue for a more thorough re-conceptualization of, for example, actionable categories. In *Prom Week*—and in the VESPACE workshop—these were limited to "Relationship" and "Network," but given the more compressed time scale of the VESPACE simulation (which unfurls over the course of a single performance, rather than an entire week), entering into a new relationship might be both inappropriate and unachievable. Moreover, the world of *Prom Week* is one in which emotion is very nearly

the sole content of in-game decisions, as different campaign goals all revolve around leveraging teenagers' (notoriously labile) feelings in order to achieve social goals in a school setting, while abstracting away many of the unsavory aspects of this experience (neither race, gender, wealth, nor sexual orientation figure into volition formation in *Prom Week*, allowing players to focus on emotional manipulation in an idealized high school setting). By contrast, relationships in the theatrical space represented in the VESPACE simulation will often be less overtly expressive, and sometimes explicitly deceptive; capturing and representing the eighteenth-century interplay of social inequality, desire, and publicity requires a new schema architecture. In addition to rearranging or re-labelling existing categories, the spatial specificity of the VESPACE game also requires designers to push further into Ensemble's entirely customizable feature-set to account for factors including proximity, eye contact, and the social topography of the theatre, which exert a strong influence on the quality (or even the possibility) of social exchanges in this environment. In looking to *Prom Week* as a model for how to build a game with Ensemble, we must acknowledge that its success as proof of a new technological approach to AI-driven social interaction games also relies on the appealing simplicity of its spatial and narrative ambit. This lesson is perhaps the most important for future projects that seek to leverage social physics and may also turn out to be the most difficult to achieve.

The future of the past with Ensemble

In many respects, Willard McCarty's injunction to think of modelling as the primary activity of digital humanities has been fulfilled (McCarty 2004). This is certainly the case in primarily quantitative DH approaches, for which Franco Moretti's famous triad of graphs, maps, and trees continues to define the expected transformation of literary and historical texts into computer-generated abstractions whose forms can help perceive previously hidden meaning within corpora (Moretti 2007). In this sense, the idea of a digital humanities outcome taking an extra-literary form is not remotely controversial and indeed has become a kind of cliché, with scatter-graphs and heat maps offered as proof of a certain kind of seriousness that legitimizes the recourse to computers, required to assimilate categories of knowledge that can coexist, however uneasily, with the narrative outcomes of "traditional" humanities research.

Building an interactive computer model using social physics, by contrast, seems to sit uncomfortably in the no-man's-land between the database and the monograph, with the resulting work available to be read as both a "model of" a world and a "model for" how to behave in it (Geertz 1973, as cited in McCarty 2004, 255). The schema and rules operate a reduction of historical source texts to a set of quantifiable values

and relationships, which can be traced back to their origins and re-evaluated for accuracy and interpretation (a model *of* literature). But when set in motion, these same elements become the basis of an emergent narrative, written in collaboration with the digital blueprint, but whose final form the initial set of instructions (now a model *for* the interactive experience) is unable to predict. In addition to anxieties around instrumentalizing literature, the performance role of the human interactor, which determines the outcome of the simulation, also contributes to the discomfort—and fascination—elicited by this approach. Ceding control to an unknown interpretive agent, situating ludic engagement as an attribute of historical research, and prioritizing emotional decision-making as the narrative spine of scholarship are all propositions raised by this work that have potentially far-reaching consequences for the conduct of humanities research in the twenty-first century.

The learning curve over the course of our week working with Ensemble was very steep. Workshop participants went from having a very basic understanding of the principles of social physics, to authoring schemas and rules that were supposed to give rise to a plausible vision of eighteenth-century sociability. Nearly all participants are literature scholars, trained to avoid treating literary artifacts as mere "content" with a transparently indexical relationship to reality, yet they were tasked with finding individual instances within these works that could serve as the basis for abstract laws that not only reflect the reality of the simulation, but define it. At the same time, authors were instructed to not attempt to re-create the scenes from which their rules are derived, but rather to think in terms of a "possibility space." We spent a lot of time trying to understand the meaning and directionality of basic interactions—at the end of an exchange, am I more motivated to try to make you like me, or are you more motivated to make me like you, and which one of these situations is captured by "affinity-up"? Halfway through the week, participants came to understand that volition rules are not the end result, but merely the necessary preparation for social exchanges, revealing a whole new level of rules-authoring required to complete the system. And then when it came time to play, the system did what systems do at the end of a long week of building and programming: it bugged and crashed. (A complete, playable version of the VESPACE VR simulation is now available to download at https://vespace.cs.uno.edu/.)

Our student coders included three French citizens, a Swiss citizen, and two Americans; our three working languages were French, English, and JSON. One exceptionally clear lesson of this workshop is to confirm the importance of working collaboratively in planning and developing procedural tools; the work must be undertaken with contributions from scholars representing a diversity of intellectual positions and life experiences to dilute the propagating effect of each individual's

latent or overt biases. Along the way, it became clear that the benefit of the workshop may not reside principally in the outcome, but that the experience of confronting literary training with a very specialized interpretive and classificatory scheme was itself an intellectually rewarding exercise that required extending towards unfamiliar concepts and protocols, while also taking a more careful inventory of the technical arsenal already available in the literary scholar's toolkit. The VESPACE model is intended to be a valuable pedagogical and research tool for those who use it, but it also turns out that *making the simulation* accomplishes some of these same goals (albeit for a smaller population), situating Ensemble game design as a point of entry into both specific content-area studies and procedural literacy, independent of the resulting playable experience.

Further development of our literary-historical work with social physics will proceed in two directions, based on the insights afforded by this workshop. The first movement expands on the goals of the workshop, delving deeper into how this work can be used to advance research in historical fields. This will involve a thorough reconsideration of the schema categories used to build the model in Ensemble, and a further systemization, and even partial automation, of the process for selecting, preparing, and coding literary exempla that underpin the volition and social exchange rules. We see this work as oriented primarily toward scholarly communities, both in the humanities and in computer science, as we attempt to push the technical capabilities of Ensemble to take into account an ever-greater number of inputs (e.g., space, gesture, intonation) while keeping calculation overload at bay. The second movement is towards a greater simplification, to develop a modular form of this tool in which certain schema categories and principles of volition and social exchange rules are already established, and instructions are provided for relatively quick customization in order to sketch out social simulations in primarily pedagogical or recreational settings that prioritize playability over theorization. Looking back on the extraordinarily rich experience of this experimental workshop, and forward toward new and expanded use-case scenarios for this technology, we are excited about the potential for social physics to open up new opportunities for thinking, perceiving, and experiencing the humanities.

Competing interests

The authors have no competing interests to declare.

Contributions

Authorial

Authorship is alphabetical after the drafting author and principal technical lead. Author contributions, described using the CASRAI CredIT typology, are as follows:

Author name and initials:

Jeffrey M. Leichman (JL)

Ben Samuel (BS)

Authors are listed in descending order by significance of contribution. The corresponding author is JL

Conceptualization: JL, BS

Methodology: JL, BS

Software: BS

Writing – Original Draft Preparation: JL, BS

Writing – Review & Editing: JL, BS

Editorial

Special Guest Issue Editor

Jeffrey M. Leichman, Louisiana State University, USA

Copy Editor

Christa Avram, The Journal Incubator, University of Lethbridge, Canada

Layout/Production Editor

Virgil Grandfield, The Journal Incubator, University of Lethbridge, Canada

References

Aarseth, Espen. 1997. *Cybertext: Perspectives on Ergodic Literature*. Baltimore: Johns Hopkins University Press. DOI: https://doi.org/10.56021/9780801855788.

Beaucé, Pauline. 2016. *Parodies d'opéra au siècle des Lumières: évolution d'une forme dramatique*. Rennes: Presses Universitaires de Rennes.

Berne, Eric. 1964. *Games People Play: The Basic Handbook of Transactional Analysis*. New York: Ballantine Books.

CETHEFI (Centre d'Etudes des Théâtres de Foire et de la Comédie-Italienne). 2022. "Accuiel." Accessed December 13. http://cethefi.org/.

Chahine, Loic. 2014. "Le théâtre et la pratique du vaudeville: établissement et jalons d'analyse d'un corpus." PhD diss. Université de Nantes/École centrale de Nantes.

Dastin, Jeffrey. 2018. “Amazon Scraps Secret AI Recruiting Tool That Showed Bias against Women.” *Reuters*. October 10. Accessed September 27, 2022. https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G. DOI: https://doi.org/10.1201/9781003278290-44.

DeKerlegand, Daniel, Ben Samuel, and Jeffrey Leichman. 2020. “Encoding Socio-Historical Exegesis as Social Physics Predicates.” In *FDG ‘20: Proceedings of the 15th International Conference on the Foundations of Digital Games*, edited by Georgios N. Yannakakis, Antonios Liapis, Penny Kyburz, Vanessa Volz, Foaad Khosmood, and Phil Lopes, 1–9. New York: Association for Computing Machinery. DOI: https://doi.org/10.1145/3402942.3409608.

Digman, John M. 1990. “Personality Structure: Emergence of the Five-Factor Model.” *Annual Review of Psychology* 41(1): 417–440. DOI: https://doi.org/10.1146/annurev.ps.41.020190.002221.

François, Paul. 2021. “Outils de réalité virtuelle pour l’histoire et l’archéologie. Recherche, diffusion et médiation: le cas des théâtres de la Foire Saint-Germain.” PhD diss. Université de Nantes/Ecole centrale de Nantes.

François, Paul, and Daniel DeKerlegand. 2023. “ Part II – Technical Contexts: Building Digital Tools for Humanities Research.” *Digital Studies* 13(2). DOI: https://doi.org/10.16995/dscn.10371.

François, Paul, Jeffrey Leichman, Florent Laroche, and Françoise Rubellin. 2021. “Virtual Reality as a Versatile Tool for Research, Dissemination, and Mediation in the Humanities.” *Virtual Archaeology Review* 12(25): 1–15. DOI: https://doi.org/10.4995/var.2021.14880.

Geertz, Clifford. 1973. *The Interpretation of Cultures*. New York: Basic Books.

Goffman, Erving. 1978. *The Presentation of Self in Everyday Life*. London: Harmondsworth.

Guimarães, Manuel, Pedro Santos, and Arnav Jhala. 2017. “*Prom Week* Meets *Skyrim*.” In *AAMAS ‘17: Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems*, edited by Kate Larson, Michael Winikoff, Sanmay Das, and Edmund Durfee, 1790–1792. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems. Accessed December 13, 2022. https://dl.acm.org/doi/10.5555/3091125.3091441#.

Isherwood, Robert. 1989. *Farce and Fantasy: Popular Entertainment in Eighteenth-Century Paris*. New York: Oxford University Press.

Leichman, Jeffrey M. 2021. “Video Games as Cultural History: Procedural Narrative and the Eighteenth-Century Fair Theatre.” In *Modes of Play in Eighteenth-Century France*, edited by Fayçal Falaky and Reginald McGinnis, 168–182. Lewisburg, PA: Bucknell University Press. DOI: https://doi.org/10.2307/j.ctv2v55bw0.13.

———. 2022. “Game Never Over: Replaying History in Interactive Digital Simulation.” *L’Esprit Créateur* 62(2): 62-74. DOI: https://doi.org/10.1353/esp.2022.0014.

Mairesse, François, and Marilyn Walker. 2007. “PERSONAGE: Personality Generation for Dialogue.” In *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, edited by Annie Zaenen and Antal van den Bosch, 496–503. Stroudsburg, PA: Association for Computational Linguistics. Accessed December 13, 2022. https://aclanthology.org/P07-1063/.

Martin, Isabelle. 2002. *Le théâtre de la foire, des tréteaux aux boulevards*. Oxford: Voltaire Foundation.

Mateas, Michael, and Andrew Stern. 2003. "Façade: An Experiment in Building a Fully-Realized Interactive Drama." In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, edited by Nathan Sturtevant and Brian Magerko, 12(1): 204–210. DOI: https://doi.org/10.1609/aiide.v12i1.12877.

McCarty, Willard. 2004. "Modelling: A Study in Words and Meanings." In *A Companion to the Digital Humanities*, edited by Susan Schreibman, Raymond George Siemens, and John M. Unsworth, 254–270. Malden, MA: Wiley-Blackwell. DOI: https://doi.org/10.1111/b.9781405103213.2004.00022.x.

McCoy, Joshua, and Michael Mateas. 2009. "The Computation of Self in Everyday Life: A Dramaturgical Approach for Socially Competent Agents." In *AAAI Spring Symposium: Intelligent Narrative Technologies II*, edited by Sandy Louchart, Manish Mehta, and David L. Roberts, 75–82. Menlo Park, CA: AAAI Press. Accessed December 13, 2022. https://www.aaai.org/Papers/Symposia/Spring/2009/SS-09-06/SS09-06-013.pdf.

Minsky, Marvin L. 1991. "Logical versus Analogical or Symbolic versus Connectionist or Neat versus Scruffy." *AI Magazine* 12(2): 34–51. DOI: https://doi.org/10.1609/aimag.v12i2.894.

Moretti, Franco. 2007. *Graphs, Maps, Trees: Abstract Models for Literary History*. London: Verso.

Moulin, Louise, Chiara Azzaretti, Julien Le Goff, and Charlee M. Bezilla. 2023. "Part III – Literary Scholarship and Social Physics: Reading and Writing with Social Physics." *Digital Studies* 13(2). DOI: https://doi.org/10.16995/dscn.9978.

Prou, Fanny. 2019. "Pour une nouvelle historiographie foraine: constitution, analyse et édition d'un répertoire (1720–1727)." PhD diss. Université de Nantes/École centrale de Nantes.

RCF (Registres de la Comédie-Française). 2022. "Le programme des Registres de la Comédie-Française." Accessed December 13. https://www.cfregisters.org/.

Riedl, Mark O., and Robert Michael Young. 2010. "Narrative Planning: Balancing Plot and Character." *Journal of Artificial Intelligence Research* 39: 217–268. DOI: https://doi.org/10.1613/jair.2989.

Rubellin, Françoise, ed. 2005. *Théâtre de la foire. Anthologie de pièces inédites (1712–1736)*. Montpellier: Editions Espaces 34.

———, ed. 2022. *Marionnettes du XVIIIe siècle: Anthologie de textes rares*. Les Matelles: Editions Espaces 34.

Ryan, James, Michael Mateas, and Noah Wardrip-Fruin. 2016. "Characters Who Speak Their Minds: Dialogue Generation in *Talk of the Town*." In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, edited by Nathan Sturtevant and Brian Magerko, 12(1), 204–210. DOI: https://doi.org/10.1609/aiide.v12i1.12877.

Sakhnovskaia, Anastasia. 2013. "La naissance des théâtres de la foire: influence des Italiens et constitution d'un répertoire." PhD diss. Université de Nantes/Ecole centrale de Nantes.

Samuel, Ben, Dylan Lederle-Ensign, Mike Treanor, Noah Wardrip-Fruin, Josh McCoy, Aaron Reed, and Michael Mateas. 2015. "Playing the Worlds of *Prom Week*." In *Narrative Theory, Literature, and New Media: Narrative Minds and Virtual Worlds*, edited by Mari Hatavara, Matti Hyvärinen, Maria Mäkelä, and Frans Mäyrä, 87–105. New York: Routledge. DOI: https://doi.org/10.4324/9781315722313-6.

Shapiro, Daniel G., Josh McCoy, April Grow, and Ben Samuel. 2013. "Creating Playable Social Experiences through Whole-Body Interaction with Virtual Characters." In *Ninth Artificial Intelligence*

and Interactive Digital Entertainment Conference, 79–85. Accessed December 13, 2022. https://www.aaai.org/ocs/index.php/AIIDE/AIIDE13/paper/view/7408/7592. DOI: https://doi.org/10.1609/aiide.v9i1.12691.

Sullivan, Anne, April Grow, Michael Mateas, and Noah Wardrip-Fruin. 2012. "The Design of *Mismanor*: Creating a Playable Quest-Based Story Game." In *FDG '12: Proceedings of the International Conference on the Foundations of Digital Games*, edited by Magy Seif El-Nasr, Mia Consalvo, and Steven Feiner, 180–187. New York: Association for Computing Machinery. DOI: https://doi.org/10.1145/2282338.2282374.

Theaville. 2022. "The Project." CETHEFI (Centre d'Etudes des Théâtres de Foire et de la Comédie-Italienne). Accessed December 13. http://www.theaville.org/kitesite/index.php.

Wardrip-Fruin, Noah. 2009. *Expressive Processing*. Cambridge, MA: MIT Press.