Data modelling and/or multimedia in digital theatre research

An **IFTR Digital Humanities Working Group** Satellite Meeting, to be held in conjunction with the conference *Performing Arts in the Digital Age: Interactions and New Directions* at the University of Maryland (10-12 March 2022).

Convenors: Nora Probst (University of Paderborn) and Miguel Escobar Varela (National University of Singapore).

Saturday 12 March, 2022

Presenter(s)	Title	Affiliation	UMD time	Presenter Location	Presenter time
Julie Holledge[1] and Joanne Tompkins[2]	Virtual Theatre History: Exploring and Displaying Data and Design	[1]Flinders University, [2]The University of Queensland	09:05 AM	Australia	00:05 AM (+1 day)
Florian Siemund	FAIR Theatre Research	Johannes Gutenberg University of Mainz	09:25 AM	Germany	3:25 PM
Rafaella Gasparian and Gustavo Sol	Marash	Centro Universitário Belas Artes, São Paulo	09:45 AM	Brazil	11:45 AM
Discussion			10:05 - 10:30 AM		
Break			10:30 - 10:50 AM		
Mylène Maignant	Mapping British Theatre Criticism Thanks to Numbers	Ecole Normale Supérieure, Paris	10:50 AM	France	4:50 PM
Kirsten Dickhaut and Gabriel Viehhauser	Theater texts as sources for historical performances – a corpus-based digital approach	University of Stuttgart	11:10 AM	Germany	5:10 PM
Marjan Moosavi[1], Mehdy Sedaghat Payam[1] and Baharak Sahami[2]	Iranian Drama by Numbers: A Computational Analysis of Iranian Plays	[1]University of Maryland, [2]Islamic Azad University (Parand Branch)	11:30 AM	US / Iran	8 PM (Iran)
Discussion			11:50 AM - 12:15 AM		

Abstracts

Virtual Theatre History: Exploring and Displaying Data and Design
Julie Holledge and Joanne Tompkins

This joint presentation demonstrates two ways we have found to conduct performance research in and with virtual reality models of theatres that no longer exist. We have created five such virtual models from cultures and performance traditions around the world as part of a research project called Visualising Lost Theatres. The two models we focus on here, both nineteenth-century venues, are the Bergen Komediehus, Norway, where Henrik Ibsen learned how to be a playwright, and the Queen's Theatre in Adelaide, Australia, the country's oldest mainland theatre. In addition to developing historically accurate models that users can navigate, we are developing two ways in which such models can contribute to a broader context for theatre history. The first involves the display and interrogation of nineteenth century techniques of scenic art: four set designs have been reconstructed on the virtual stages of these two theatres from plays staged in 1841, 1857, 1886, and 1897. The second is the potential to use the venues as a portal to link to additional data as a means of collecting and layering relevant material in innovative ways. We have worked on two different ways of curating data related to these venues: a webpage connected to the five models, and a virtual exhibition that links our research with data stored in AusStage, the digital research resource of Australian performing arts. This research illustrates how different examples of digital research can work together to enhance our understanding of theatre history.

FAIR Theatre Research
Florian Siemund

As an academic discipline, theatre history depends on a large array of different types of source material – I am calling it "theatre data" – in order to approximately (re-)create past theatre events. Thus, machine-understandable and sustainable digital resources are helpful tools and key elements in keeping track of a vast and continuously growing amount of relevant theatre data. To avoid redundant work, or even data loss in this field, theater researchers of the 21st century should act thoughtful and efficiently, by using interdisciplinary methods of (digital) data management. Some of these methods for realizing sustainable data management are described in the so called "FAIR" concepts.

"FAIR" is an acronym which stands for ("Findable, Accessible, Interoperable and Reusable"[1]) resources, that can exist both on- and offline. Under each of the acronym's letters, FAIR describes a few specific technical measurements which could be applied to a set of data, in order to improve its scientific quality. In an ideal design, all FAIR principles can be applied to each phase of the research data lifecycle, which in itself is a model of sustainable and efficient research data management.[2]

Other aspects of (digital) data management, like security, ethics or rights are not addressed by the FAIR principles, so the principles should not be considered a "standard" which, once implemented, guarantees a better data quality. But they are designed as an orientation, in order to improve the scientific worth (in terms of reproducability, transparency, reusability etc.) of research works regardless of the underlying subject. I would like to demonstrate a few use cases, on how to implement FAIR-principles in theatre research, how theatre history might profit from FAIRs interdisciplinary approach and at which point, even theatre data that is 100% FAIR by definition, still can be further improved by applying additional concepts of (digital-) data management.

---

[1] Wilkinson et al., 2016. https://doi.org/10.1038/sdata.2016.18
[2] Sarretta, 2018. doi:10.1890/120375

Marash
Rafaella Gasparian and Gustavo Sol

In the performance *Marash*, the actress Rafaella Gasparian tells the story of her family that moved to Brazil because of the Armenian Genocide, caused by the Ottoman Turkish Empire in 1915, being a historical fact which to these days is denied by Turkey. From 2020, the situation in Eastern Europe worsens, starting a war between Armenia and Azerbaijan, with the support of Turkey, over the Nagorno-Karabakh region. With the risk of a possible genocide happen again, the Armenian people around the world are mobilizing in street demonstrations, but Rafaella decides to discuss the Armenian diaspora differently.

With the direction of Gustavo Sol, Gasparian started *Marash*, an artistic process based on a fictional autobiography. As poetic operations, she uses visualizations, narratives, and spatialization of memory with silent moments, like their maternal great-grandfather and grandmother, when they tell about their scape through the Syrian desert. Identity, political and cultural issues are discussed in the dialogues, besides to girl´s curiosity to know what was not and what is not said.

Gustavo Sol, in addition to directing, also contributed using an interface developed in his doctorate that classifies neurophysiological states associated with different states of poetic presence through artificial intelligence. The interface is used during the rehearsals to identify and help understand the states of performativity, measuring the actress´s heart rate during the scene that helps to get into the best states results that the scene asks for. This provide us with neurophysiological, cardiac and autobiographical data that are used to contribute to dramaturgy and scene development, as well as historical documents. The system that Sol has been developing for over 10 years, allows fluid and live interaction with audiovisual materials, in case videos and sounds collected during the Gasparian´s recent trip to Armenia.

Mapping British Theatre Criticism Thanks to Numbers
Mylène Maignant

Focusing on a corpus of more than 40 000 London theatre reviews going from 2010 to 2020, my research consists in analyzing the reception of London theatre through the eyes of two distinct communities: journalistic criticism on the one hand, and digital criticism (bloggers) on the other. Since 2010 in England, a wave of blogs written by authors coming from various horizons has arisen on the Internet. Students, theatre professionals but also mere amateurs began publishing their own theatre reviews. These new independent voices in the digital space progressively redefine the shape of classic journalistic criticism. Although discreet, they offer a new vision of the history of London theatres. By doing so, it sets itself apart from the canon of mainstream journalism.

By using various digital tools and techniques (computational stylistics, sentiment analysis, Geographic Information System, etc.), this research aims at modelling the differences and similarities between these two communities to better understand the cultural discourse they construct. An online platform displaying all the interactive graphs as well as a search engine based on the dataset have been recently created.

Theater texts as sources for historical performances – a corpus-based digital approach
Kirsten Dickhaut and Gabriel Viehhauser

In digital literary studies, drama texts have been popular areas for the application of computer-aided methods. Especially, network analysis has been used for a 'distant reading' of character constellations. However, although in principle concerned with the interaction of characters, such approaches often remain directed at the text and do not consider the performance aspect of theater and especially historical performances. Therefore, we want to present a novel approach to computer-aided analysis of drama texts that aims at exploiting those texts as sources for historical performance practice. We start from the assumption that the way in which drama texts thematise performance themselves by e.g. referring to terms that are related to scenographic practice can be informative for historical theatre practice and that a corpus-based distant reading of these terms can trace the discourses conducted about these practices. To employ such an approach, we use two methods from corpus linguistics resp. distributional semantic: 1) Collocation analysis and 2) the detection of semantic word fields with the help of word embeddings (Word2vec, Mikolov et. al. 2013).

1600-1699		1700-1799		1800-1899	
token	S_c	token	S_c	token	S_c
prodige	0.7308	tableau	0.7052	but	0.7088
présage	0.6909	changement	0.7033	tribunal	0.6940
sacrifice	0.6700	présage	0.7014	havre	0.6929
séjour	0.6654	tumulte	0.6731	tableau	0.6922
appareil	0.6407	délire	0.6616	piège	0.6810
accueil	0.6364	séjour	0.6603	parterre	0.6711
tableau	0.6361	gouffre	0.6560	chien	0.6672
climat	0.6352	désordre	0.6492	canada	0.6669
changement	0.6201	voile	0.6475	brillant	0.6668
mélange	0.6158	nuage	0.6456	siècle	0.6641
cri	0.6104	précipice	0.6449	réseau	0.6569
miracle	0.6083	péril	0.6400	genre	0.6521
trophée	0.6049	sacrifice	0.6398	port	0.6505
projet	0.5954	chaos	0.6382	parfum	0.6490
revers	0.5907	sommeil	0.6357	signal	0.6456
accident	0.5867	piège	0.6347	bourdonnement	0.6434
fantôme	0.5865	antre	0.6342	combat	0.6410
monstre	0.5865	prestige	0.6331	début	0.6369
vacarme	0.5828	monument	0.6293	centre	0.6366
gouffre	0.5822	récit	0.6292	supplice	0.6365

Figure 1: The 20 most similar words to the term 'spectacle', calculated with cosine similarity measured between the vectors of a word2vec-model build on the corpus separated into three timespans (17th, 18th and 19th century). Cells coloured in blue indicate words that point towards a metaphysic dimension.

We apply our approach to a corpus of 1560 French plays as collected by Paul Fièvre for his website Théâtre Classique and provided in standardized TEI format by the DraCor project (FreDraCor 2021). Figure 1 shows the 20 most similar words to the term 'spectacle' calculated with cosine similarity on the vectors of a word2vec-model of the corpus divided by century. As can be seen, the most similar

words in the plays written before 1700 show a strong metaphysic component which decreases in the following centuries.

1600-1699		1700-1799		1800-1899	
bigram	likelihood	bigram	likelihood	bigram	likelihood
quel spectacle	249.28	quel spectacle	453.36	quel spectacle	64.75
triste spectacle	172.17	spectacle enchanteur	103.82	spectacle effrayant	38.04
horrible spectacle	148.46	spectacle odieux	85.188	beau spectacle	37.55
tragique spectacle	113.55	spectacle funeste	74.67	spectacle douloureux	24.10
spectacle sanglant	112.05	spectacle cruel	60.49	spectacle réexcité	19.82
spectacle si	110.16	spectacle effroyable	43.08	pareil spectacle	17.35
spectacle odieux	97.62	spectacle terrible	43.06	spectacle naissait	16.00
spectacle affreux	88.52	spectacle affreux	41.63	spectacle risible	13.32
spectacle nouveau	86.60	spectacle effroyableâ	35.86	spectacle pitoyable	12.77
funeste spectacle	76.51	spectacle pompeux	35.42	spectacle ravissant	12.77
spectacle horrible	59.04	spectacle touchant	32.18	spectacle réel	11.78
spectacle pompeux	56.24	spectacle horrible	23.19	spectacle apparaît	11.59
spectacle assez	55.87	spectacle effrayant	22.89	spectacle applaudir	11.34
nouveau spectacle	45.79	spectacle intéressant	21.58	spectacle hideux	11.05
barbare spectacle	40.01	spectacle charmant	21.43	spectacle incroyable	10.98
spectacle inhumain	36.77	spectacle fatal	21.17	spectacle effroyable	10.63
spectacle effroyable	36.40	spectacle si	21.16	spectacle restera	10.47
spectacle charmant	31.55	spectacle épouvantable	20.46	spectacle amer	10.42
spectacle funeste	31.33	nouveau spectacle	19.33	spectacle merveilleux	9.98

Figure 2: The 20 most distinct collocations in the corpus calculated for each of the three timespans (17th, 18th and 19th century). Cells coloured in blue indicate words that point towards the dimension of *movere*.

Figure 2 shows the 20 most distinctive collocations (i.e. common appearances) of 'spectacle'. An increasingly high number of collocations can be linked to the dimension of movere that has been intensively discussed in poetics since the beginning of the 17 th century. As the term 'spectacle' is in the first case linked to something exposed to be seen, and in the second it is specified for spectacle in theatre, this observation might confirm the historical semantics, as metaphysics are shown and only in the second case drama is reflected.

References
Mikolov, Tomas & Chen, Kai & Corrado, G.s & Dean, Jeffrey. (2013). Efficient Estimation of Word Representations in Vector Space. Proceedings of Workshop at ICLR. 2013.
https://www.theatre-classique.fr/index.html

French Drama Corpus (FreDraCor): A TEI P5 Version of Paul Fièvre's "Théâtre Classique" Corpus. Edited by Carsten Milling, Frank Fischer and Mathias Göbel. Hosted on GitHub, 2021–.
https://github.com/dracor-org/fredracor

Iranian Drama by Numbers: A Computational Analysis of Iranian Plays
Marjan Moosavi, Mehdy Sedaghat Payam and Baharak Sahami

Many anthropologists of Islam and the Middle East believe that as scholars and educators we must be fully alert to the need to attend to the everyday lives of micro-level actors in their shifting multi-layered contexts. Dramatic literature is a superb example of a literary genre that is reflective of and responsive to the everyday socio-political reality and aesthetic sensibility of people. This study aims to explore the trends that a computational analysis of 100 Iranian plays reveals to us about the complexity of everyday life grounded in Iranians' aspirations, relations, patterns of pleasure-seeking, and consumption, as well as their ideological repositionings and religio-moral ambivalence. Such analysis reveals trends and topics that may not be visible from a close study of this corpus.

As a subfield of Natural Language Processing, topic modeling and its powerful text mining technique allows us to discover latent data in large corpora and the relationships between the data. In the current research, we apply the most common statistical technique Latent Dirichlet Allocation (LDA) to a corpus of 100 Iranian plays written between 1880 and 2010. This generative statistical model helps us identify collections of words that co-occur together in each playscript but also co-occur in other playscripts throughout the corpus. Since any single playscript is likely to contain a wide range of topics, i.e. be "about" more than one subject, the LDA model assumes that every playscript is a collection of topics in different proportions. Equally important is to assume that the entire corpus is a collection of topics in different proportions. Therefore, the analysis unfolds in two phases. In the first phase, LDA discovers topics in a group of playscripts. This requires removing the stop words, which are mainly the structure words in playscripts. In the second phase, it demonstrates the relevancy of topics to each other. The package that we have selected for our analysis is Mallet that has the stop word removal function built into it.

Presenter bios
In order of presentation

Joanne Tompkins, recently retired from the University of Queensland, has published widely on spatiality and theatre, among other topics. Most recently she has led a project that reconstructs lost theatres through virtual reality: the book from this project, *Visualising Lost Theatres*, co-authored with Julie Holledge, Jonathan Bollen and Liyang Xia, is forthcoming from Cambridge. She has co-edited/edited *Modern Drama* and *Theatre Journal*, and special issues of *Contemporary Theatre Review* and *Australasian Drama Studies*. The Australasian Drama Studies Association named an editing prize after her. She is a Fellow of Australia's Academy of Humanities; the recipient of an honorary doctorate from Queen Mary, University of London; and an executive committee member for the International Federation for Theatre Research. Following a three-year secondment to the Australian Research Council where she was Executive Director for Humanities and Creative Arts, she currently consults on strategic research grant writing and research management.

Professor Emeritus Julie Holledge FAHA, has published extensively in the field of women's performance and is author of *Innocent Flowers: Women in Edwardian Theatre* (Virago 1981); and is co-author with Joanne Tompkins of *Women's Intercultural Performance* (Routledge 2000). When she retired from Flinders University, she was appointed as a research professor to the Centre of Ibsen Studies at the University of Oslo (2011-2020), she published three books on Ibsen: *A Global Doll's House: Ibsen and Distant Visions* (Palgrave 2016); *Ibsen Between Cultures* (Novus forlag 2016) and *Ibsen on Theatre* (Nick Hern 2018)*.* In her role as a founder of AusStage, she has pioneered e-research methodologies in the field of theatre studies and is currently working on reconstructing stage settings and movement in nineteenth century plays using VR technology.

Florian Siemund. As a student of Theatre Studies (Theaterwissenschaft) at the Johannes Gutenberg University Mainz, he received his Master of Arts degree in 2021. In parallel, he studied Digital Humanities at the very same university, in connection with Hochschule Mainz. He expects to receive his degree in March, 2022 and is planning to start his PhD project in Digital Humanities, in the following months, at the University of Cologne.

Rafaella Gasparian had her first contact with theater at the age of 6, performing "The Wizard of Oz". Keeping her passion for theater and performing many other school plays, at the age of 15 she entered the vocational course of actors at the Atores Wolf Maya School of Acting, working with some of the prominent directors and choreographers of São Paulo. During the period she became more interested in her Armenian family history, learning many cultural aspects. In 2020, at the age of 17, she started her grading in Performing Arts at the Centro Universitário Belas Artes, of São Paulo.

Gustavo Sol (www.gustavosol.com.br) Actor and director specialized in neurocomputing of performing arts. Holds a PhD in Arts from ECA / USP (2017), with a Sandwich Doctorate at Paul Valery III University (France), a Master´s degree in Communication and Semiotics from PUC / SP (2008) and a Bachelor´s Degree in Performing Arts from UNICAMP (2000). He contributes to the Robotics and Rehabilitation Research Laboratory - LABORE / IF and is a professor at the Centro Universitário Belas Artes and is collaborator of Digital Dramaturgy Lab Squared, from University of Toronto, Canada. https://www.gustavosol.com.br/marash

Mylène Maignant is the founder of Drama Critiques (http://www.dramacritiques.com/), the first database which explores 10 years (2010 - 2020) of London contemporary theatre by using data science. Mylène holds two Master's degrees: the first one in English Language and Literature from Sorbonne Université, and the second one in Digital Humanities from Ecole des Chartes. She is

now a fully funded PhD candidate at Ecole Normale Supérieure (Ulm) in Paris where she is finishing her research in IT and literature.

Kirsten Dickhaut, Professor of French and Italian literature at the university of Stuttgart, Germany, is specialized in Magic and Theatre in Early Modern times. Her recent works concentrate on magical steering and magical figures on the one hand, and on theatre as an intelligent system on the other hand.

Marjan Moosavi is a lecturer, researcher, digital curator, and dramaturg. She is the Roshan Lecturer in Persian Studies and Performing Arts and Associate Director of the Roshan Initiative in Persian Digital Humanities. She holds a Ph.D. in Theatre and Performance Studies from the University of Toronto. Her work, whether academic or artistic, examines the dynamics of theatre-making in the Middle Eastern countries, particularly Iran and theatre's intersection with gender, religion, and politics. Her scholarly articles and interviews are published in The *Drama Review (TDR), New Theatre Quarterly, Modern Drama, Ecumenica, Asian Theatre Journal, Critical Stages*. Marjan's digital and curatorial collaborations include two pioneering and transnational digital projects: Digital Guide to Theater of the Middle East and Digital Photo Exhibit on the Middle Eastern Theatre. She is a longstanding Regional Managing Editor for TheTheatreTimes.com. Currently, she is working on a book manuscript titled: *Theatre Making for World Marking: Dissidence Played Decently on Today's Iranian Stage*.
Twitter: @marjan_moosavi

Mehdy Sedaghat Payam received his Ph.D. in experimental print fiction and digital literature from Victoria University of Wellington in New Zealand in 2014. Currently, he is doing his second Ph.D. in Comparative Literature at the University of Maryland. He has worked on electronic literature, but he uses the text mining method to analyze 600 Persian novels in his current doctoral studies. Since the last two years, he has also been working on a Mellon-awarded project to develop an OCR engine for Persian and Arabic scripts. He is also a published novelist and translator. His first novel was published in Persian in 2009, and his Persian translation of *Digital Humanities* by Anne Burdick et al. was published in 2016.
Twitter: @mspayam

Baharak Sahami is a senior lecturer at English Department in Islamic Azad University, Parand Branch, Iran. She is also a translator, theater critic and the regional Managing editor of the Iran section for TheTheatreTimes.com. She received her MA in English Language and Literature from Shahid Beheshti University, Tehran. Her main areas of interest are literary criticism and theater studies. Her Persian translation of Collin McGinn's *Shakespeare's Philosophy* has been recently published in Iran.