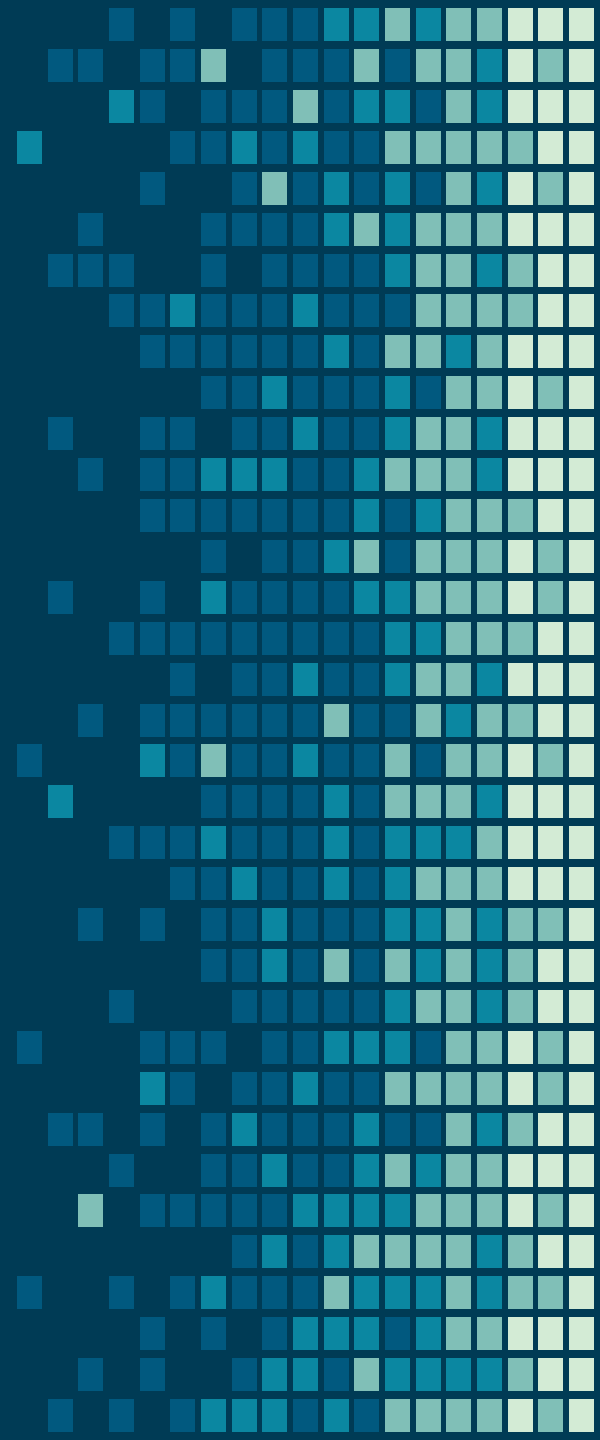


# Who is the author?

## Modeling creative relations in television writing

Joanna Byszuk  
@jbyszuk

# Project background



Distant ~~reading~~ watching audiovisual data

# Distant ~~reading~~ watching audiovisual data

## **Text-focused:**

Corpus linguistic – Monika Bednarek,

stylometry – number of works by Agata Hołobut, Monika Woźniak and Jan Rybicki, previous papers by me,

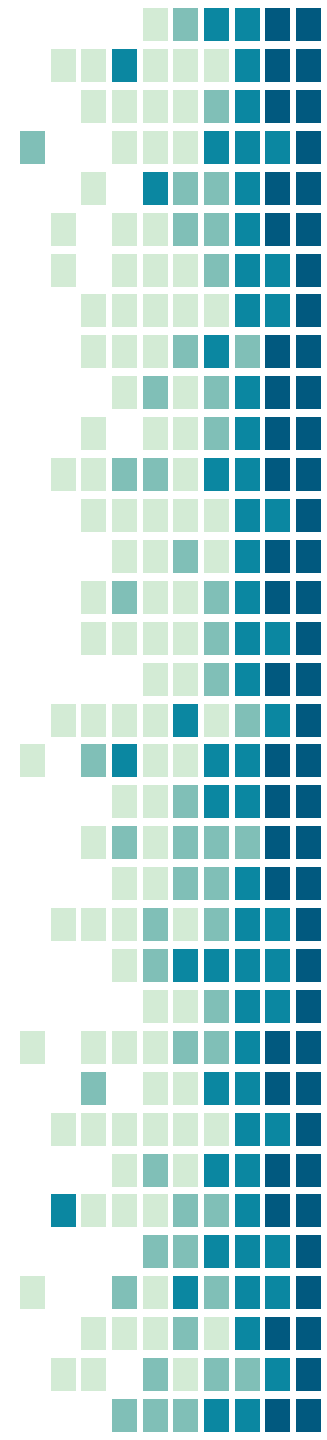
Topic modeling – various scholars, e.g. Ben Schmidt  
'Typical TV episodes: visualizing topics in screen time'

## **Visual and beyond:**

Taylor Arnold and Lauren Tilton:

<https://distantviewing.org/>

Also various visual and audio projects related, among others, to DHd AG Film und Video.

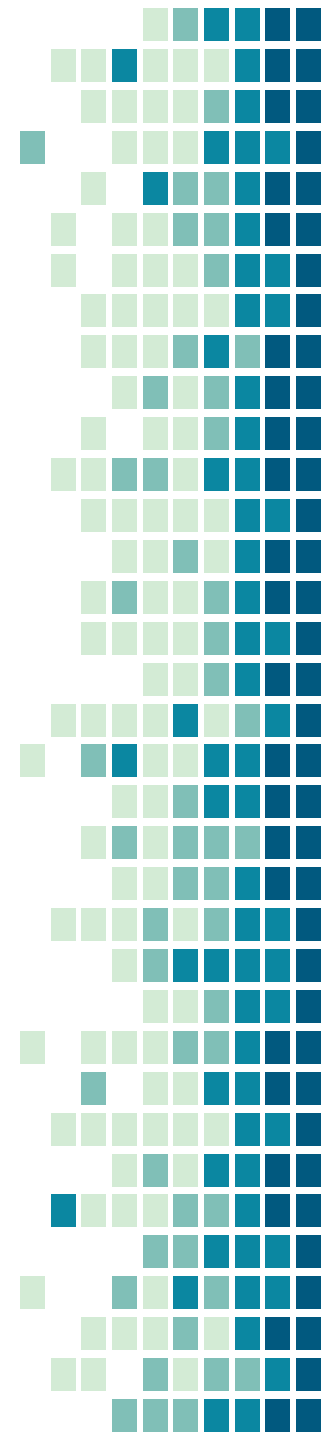


Analysis of textual layer of TV shows

# Analysis of textual layer of TV shows

Focus on:

- Topic *good*
- Character idiolect *good*
- Language features *good*
  
- Genre *okay*
- Author *meh*



Authorship in television

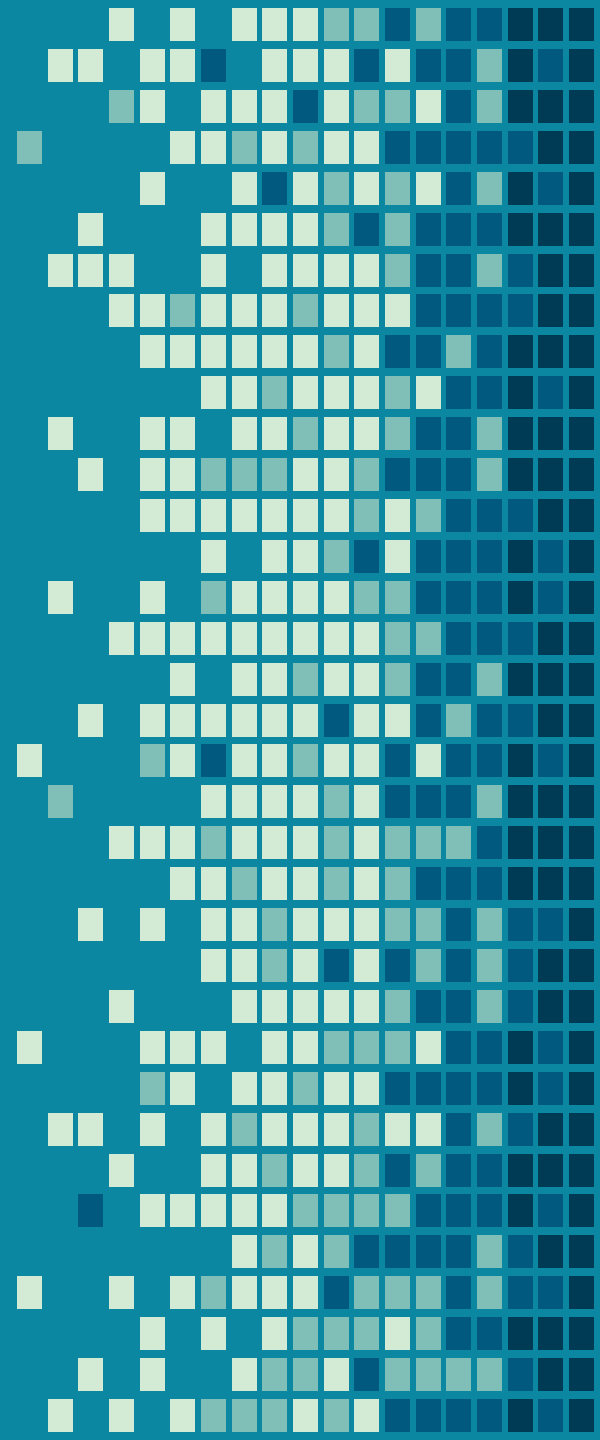
—

film studies perspective

“

*the issue of authorship 'has been rarely associated with television [as the authorship] was diffused among many individuals and may more properly have been claimed by a studio, company, or channel than by a person'.*

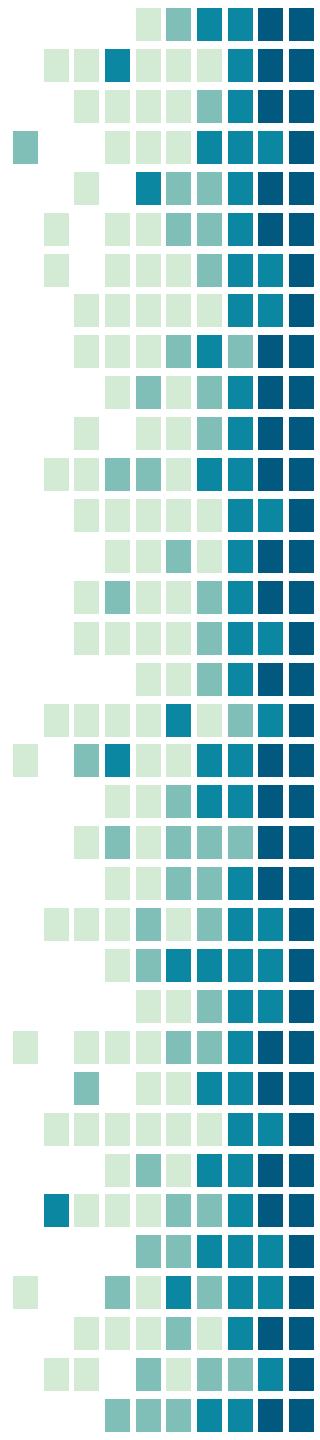
Hartley, (2004) From Republic of Letters to Television Republic? Citizen readers in the era of broadcast television, p. 401





# Foucauldian perspective on authorship

author-function – concept applicable to television studies?



“

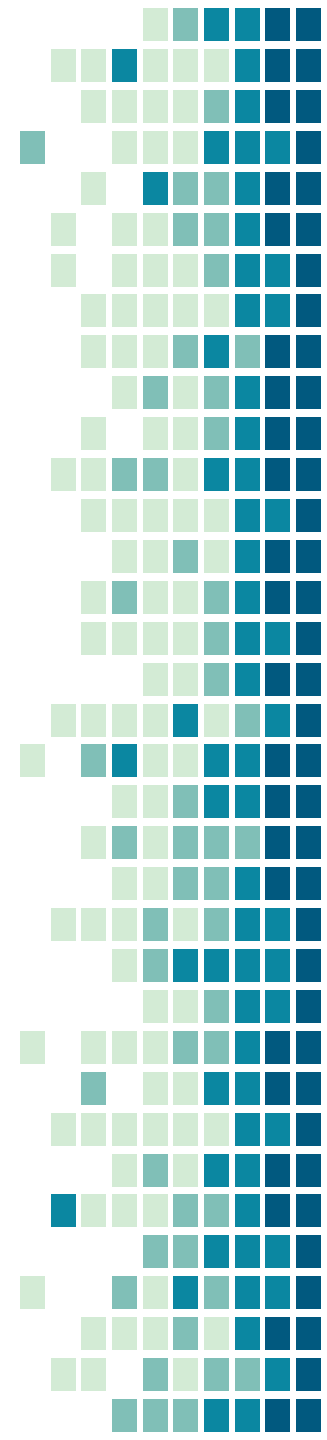
*Television fiction provides an interesting case study for Foucault's model of authorial function—the material creative process is far more collaborative and decentered than in most other media, and television authorship has been mostly hidden from the public eye for much of its history, buried in confusing credits far more than the prominent role of literary writers or film directors.*

Jason Mittell, (2015) *Complex TV: :  
The Poetics of Contemporary  
Television Storytelling*

Who writes TV?

# Who writes TV?

Role of showrunner vs film director vs producer



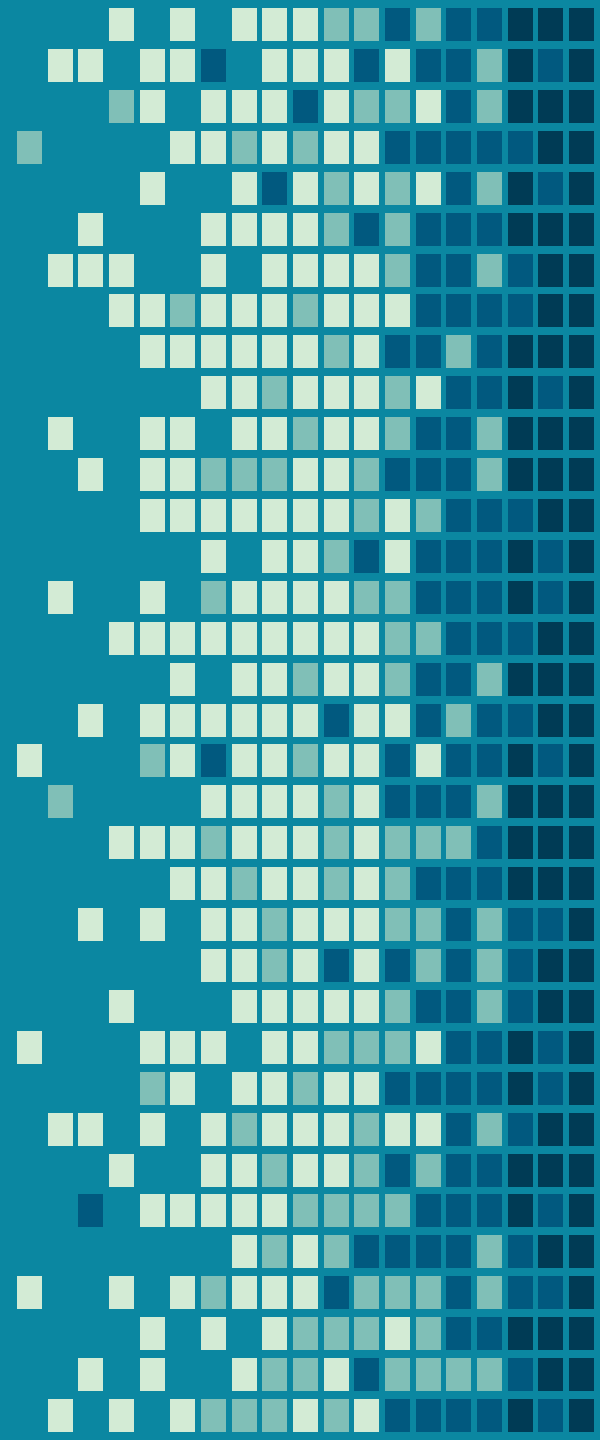
💬💬 *I might change at least 30 per cent of the material, often 60 per cent, sometimes almost 100 per cent. I go over every line of dialogue, either adding new stuff or refining what's there; sometimes that means enhancing a line that the original writer hasn't realised is good. I'll bring out themes, punch up moments, add signature dialogue, clarify stage directions and make cuts. To every single scene, if need be. Usually, the basic shape remains intact, but sometimes I'll invent brand new characters and subplots... while at the same time remaining faithful to the original writer. I'll even impersonate them.*

Russell T Davies

“

*what you're there to do  
is to say in different  
ways every single day, to  
send out a consistent  
message: this is the show  
we're trying to make, it's  
like this.*

Steven Moffat



“



**Lauren S. Hissrich** ✓ @LHissrich · 19 paź

I haven't talked about writing in a long time, but it's important to know what happens to all those words when they fly off the metaphorical page and collide with a very real production.

Most of the time, a lot changes. It has to. There are logistics to consider:



**Lauren S. Hissrich** ✓ @LHissrich · 19 paź

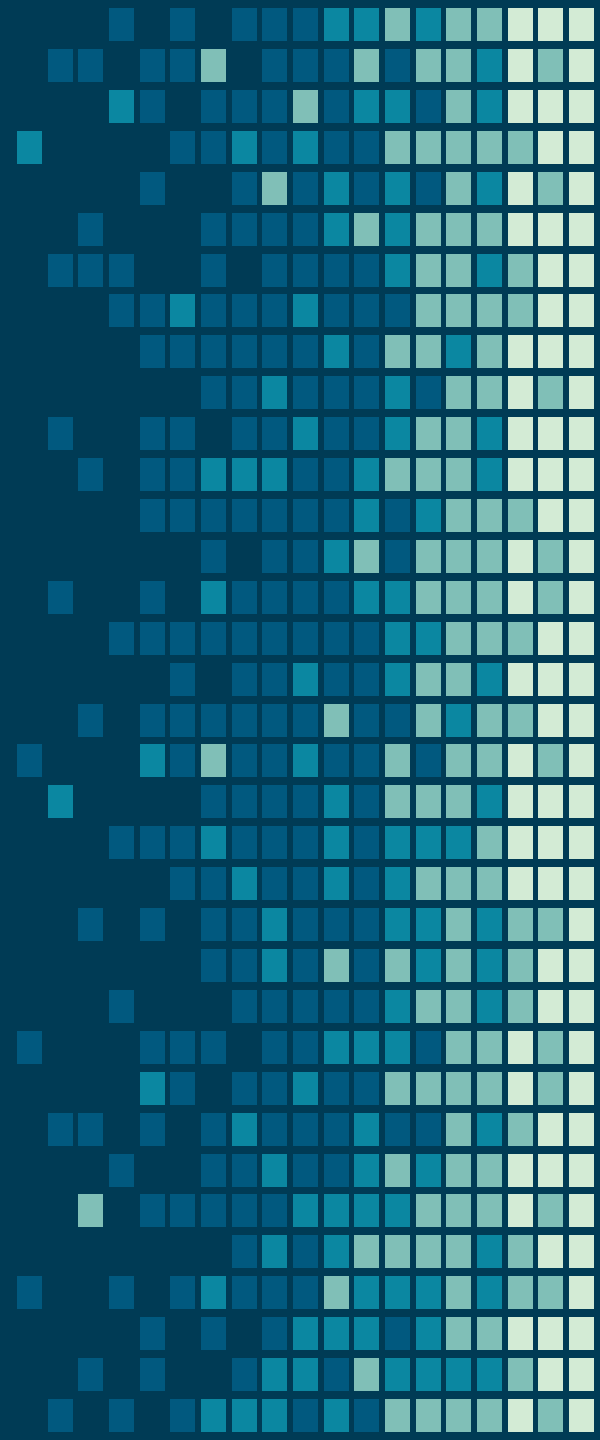
Creative vision alignment. Director notes. Cast notes. Horse notes. Self-flagellating notes. Budget. Run time of episodes. Rain. Snow. Stage availability. Crew availability. Actor availability.



**Lauren S. Hissrich** ✓ @LHissrich · 19 paź

Anyway, I take all those factors back to my computer, and start what we call the revision process. There's a lot of cool coding in revisions, stars and pretty colors, but let's be honest, often it's where the fun dies. Writers -- as we colloquially say -  
- must kill our babies.

# Hypothesis





# Tracing Showrunners' Impact

– presented at AIUCD 2018

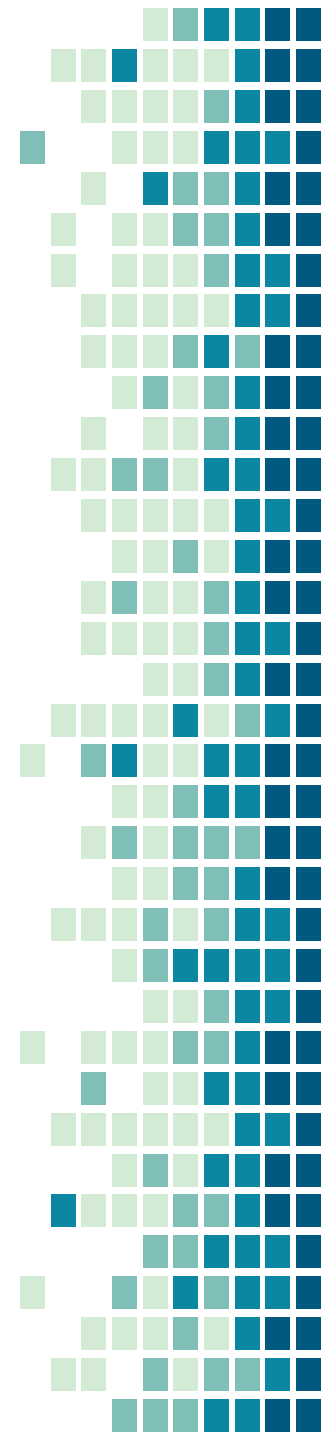
## **Varying showrunning styles:**

Solo showrunner and the rest – Moffat

Collaborative showrunner – Fey, Whedon

Distinct showrunner with big impact over other writers – Davies, Rhimes

Hidden showrunner – Whedon



# Build-up on the findings from Tracing Showrunners' Impact

## **Varying showrunning styles:**

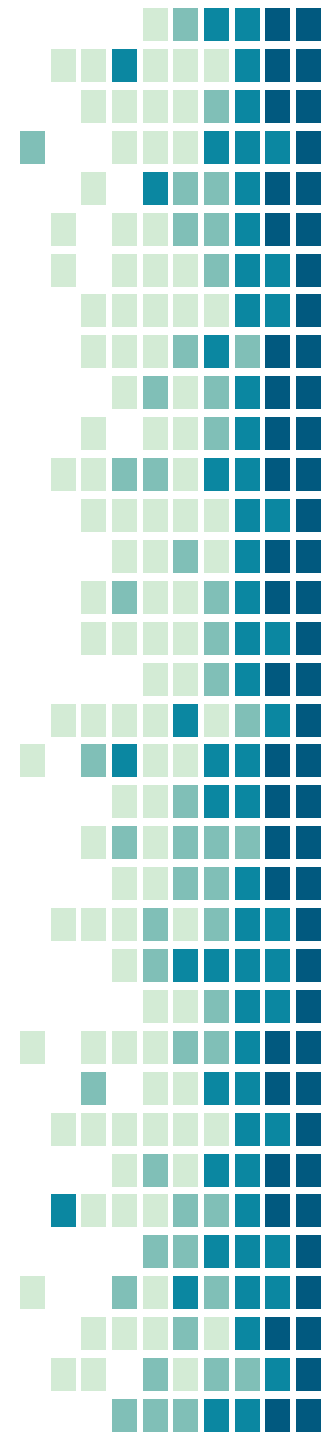
Solo showrunner and the rest – Moffat

Collaborative showrunner – Fey, Whedon

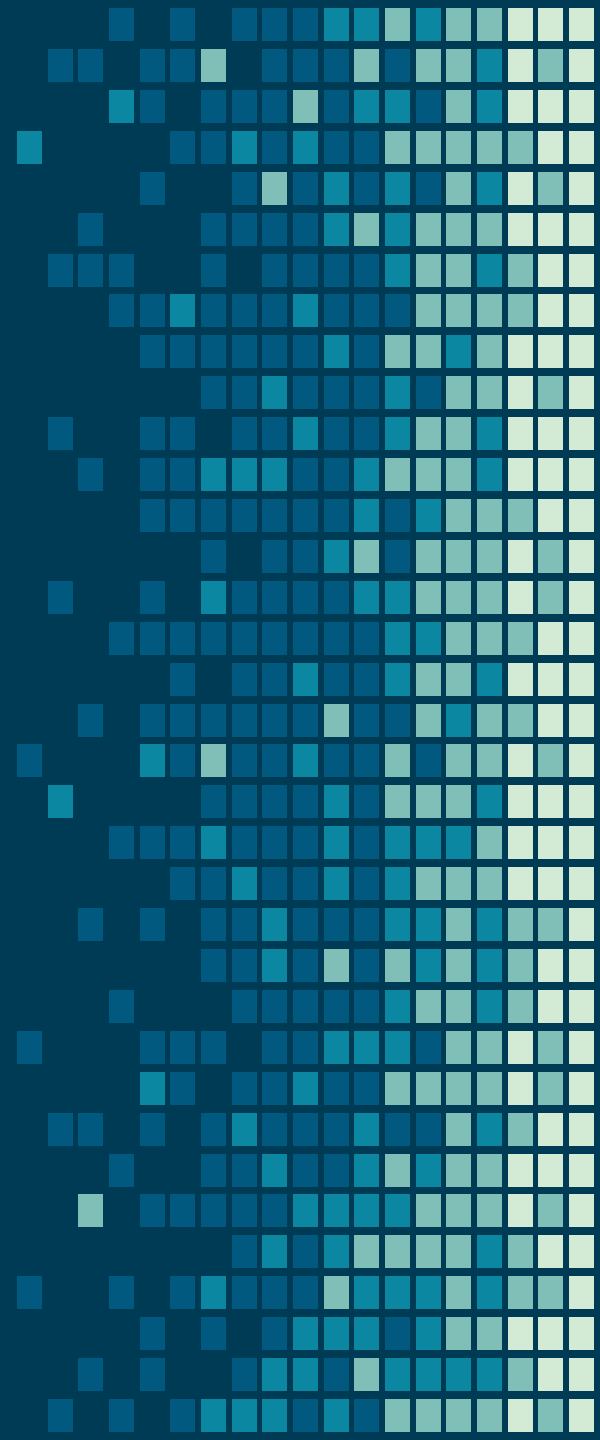
Distinct showrunner with big impact over other  
writers – Davies, Rhimes

Hidden showrunner – Whedon

+ Production-focused showrunner



# Study



Data

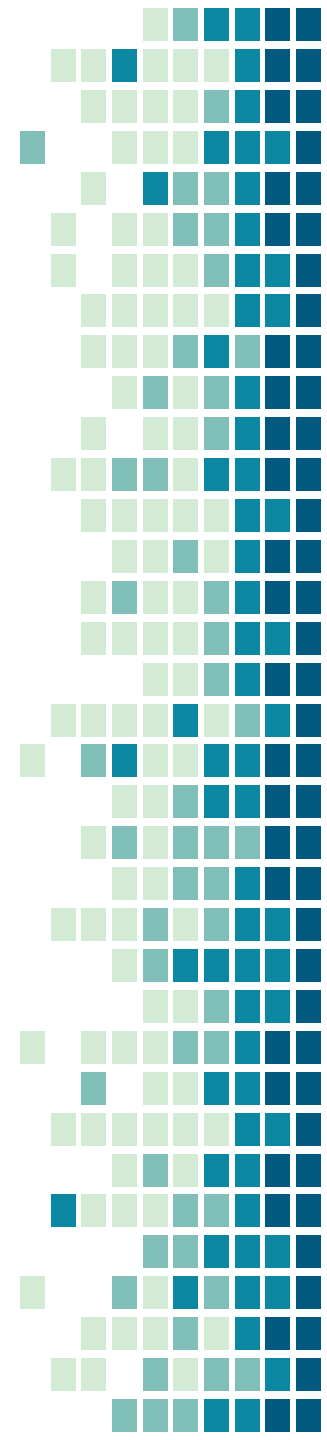
# Obtaining data: subtitles vs transcripts

## Transcripts:

- More accurate
- Additional information: who speaks?
- Very rare

## Subtitles:

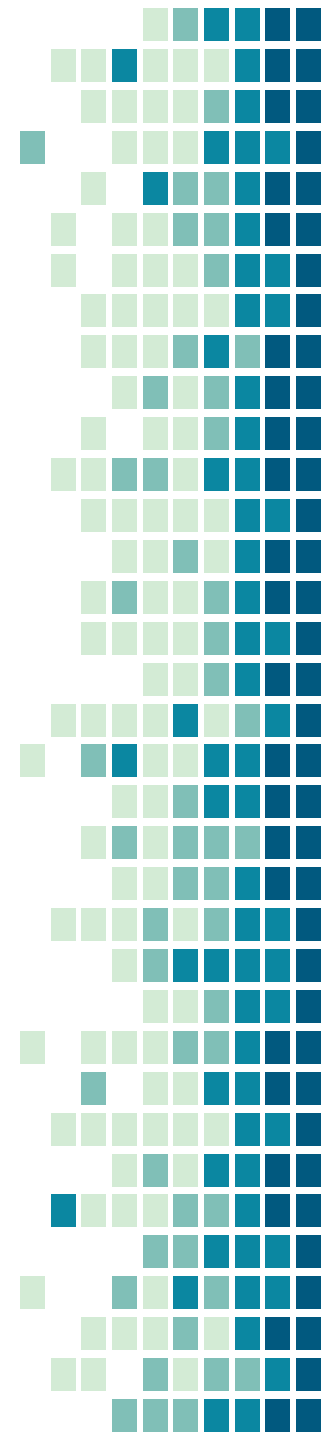
- Audiodescriptions fairly accurate
- Require more cleaning
- Easy to scrap and process



# Obtaining data: creator metadata

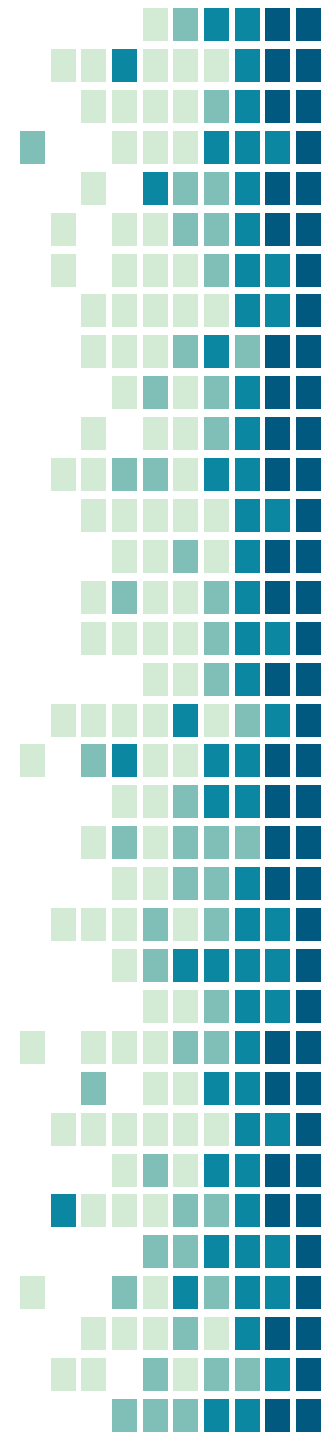
## Selection criteria:

- At least two shows per candidate author
- Two basic genres:
  - (medical / crime / social) drama
  - science-fiction
- Episodes at least 40 minutes long



# Obtaining data: creator metadata

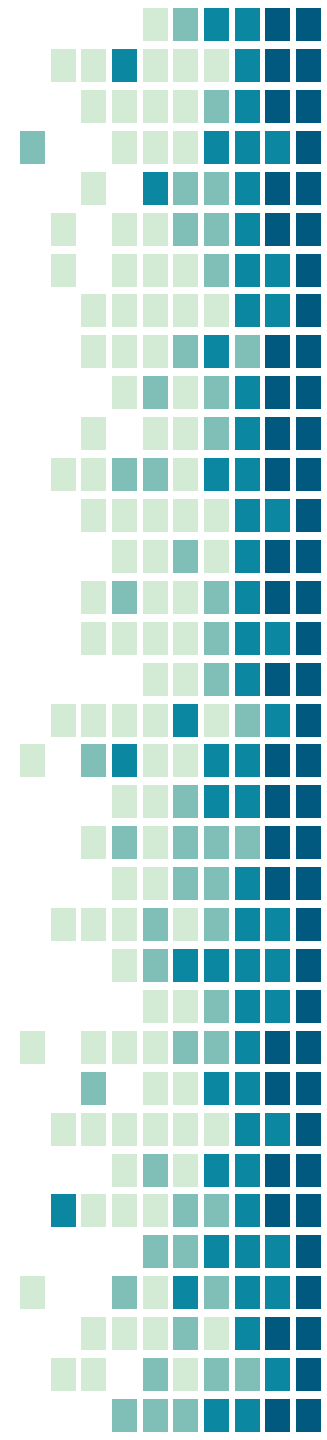
Imdb (and a number of other) API – problem:  
mapping the information across various types of  
description



# Method

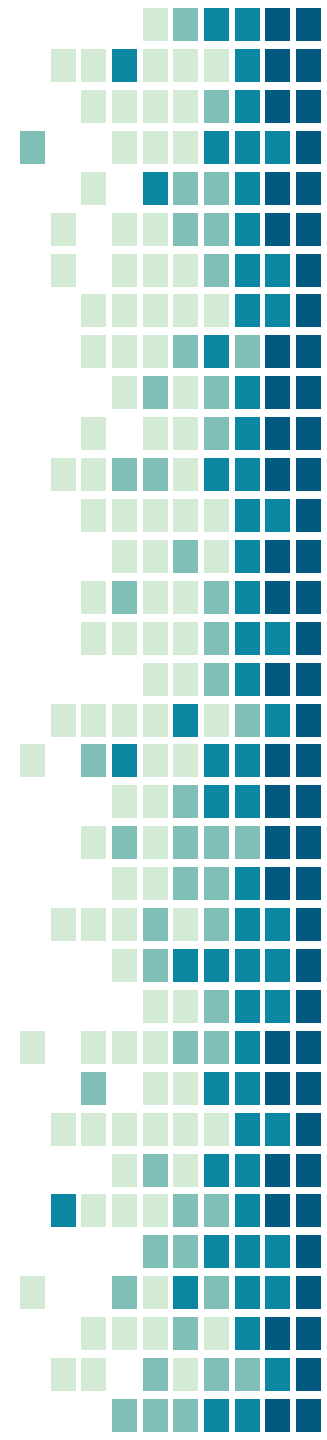


100-1000 Most Frequent Words, no culling  
bootstrap consensus networks (a variant of  
hierarchical clustering) – Stylo implementation\* +  
Gephi visualisation

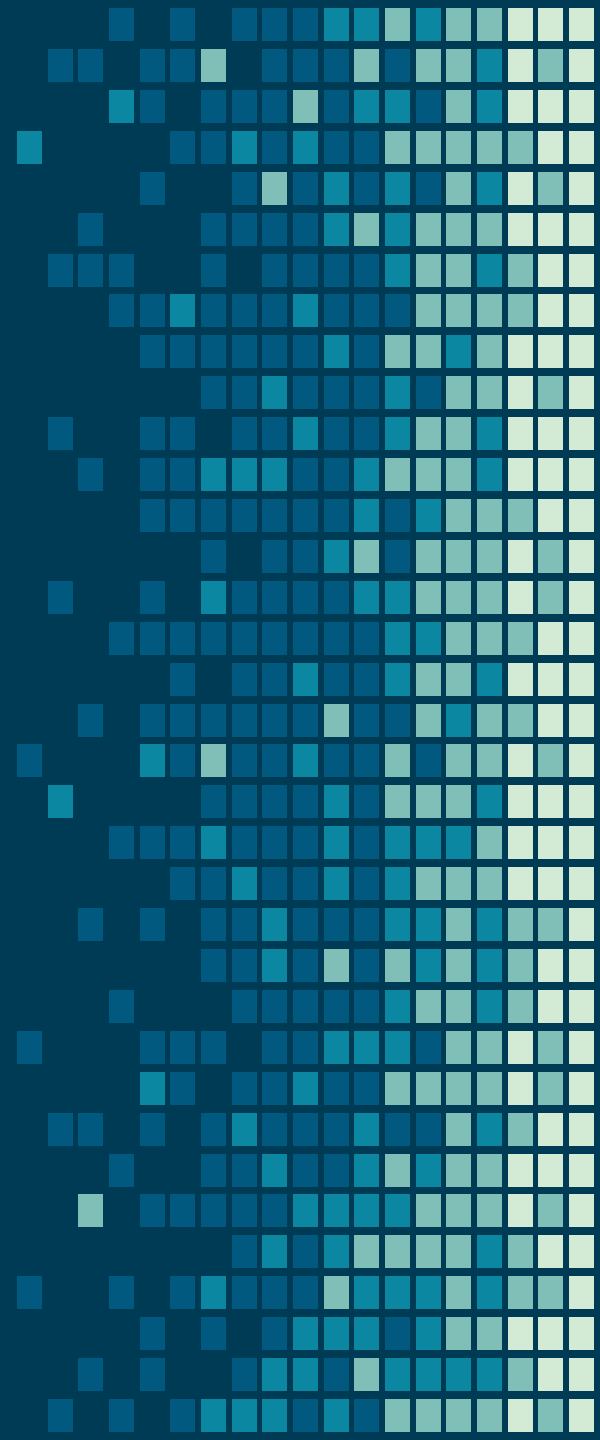


# Final corpus

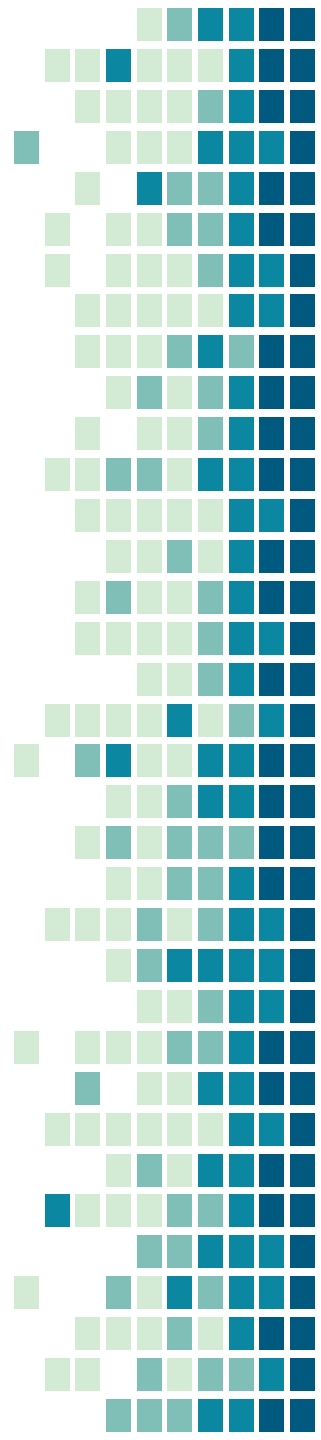
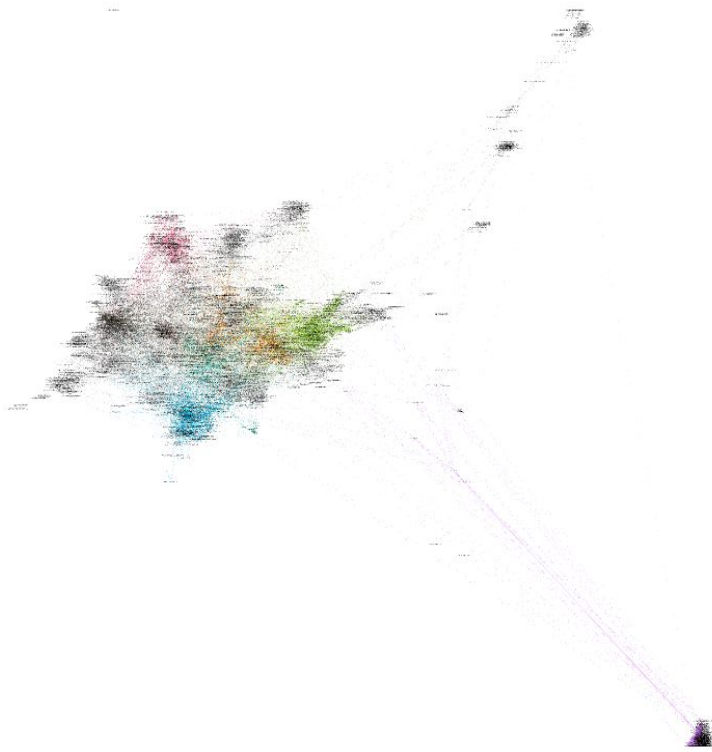
- 86 shows (+6)
- ~30 million tokens
- Representative for genre
- Unbalanced in terms of amount of texts per show (typical American format 10/13/24, typical British format  $\leq 13$ )
- Some spelling normalization required



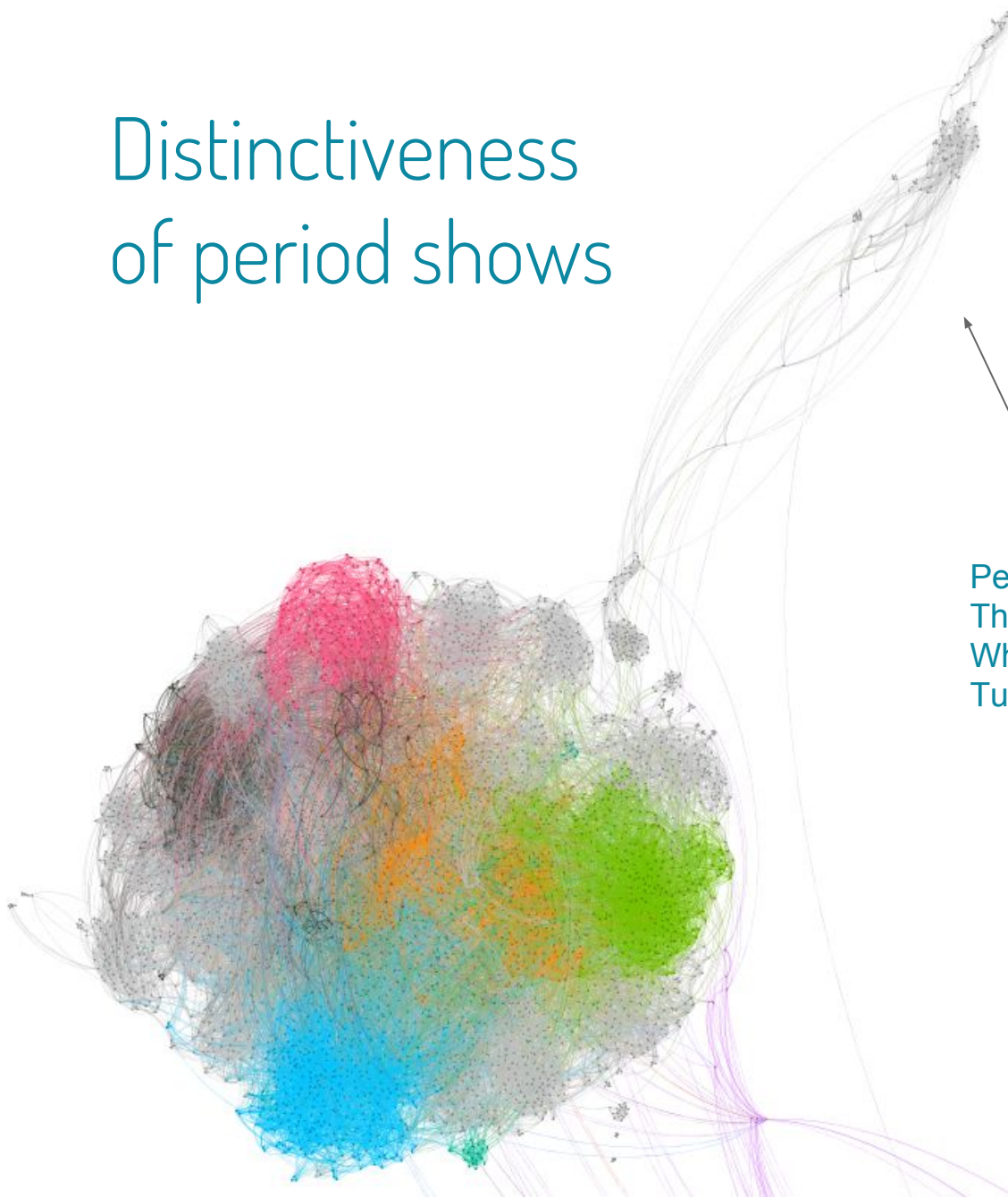
# Findings



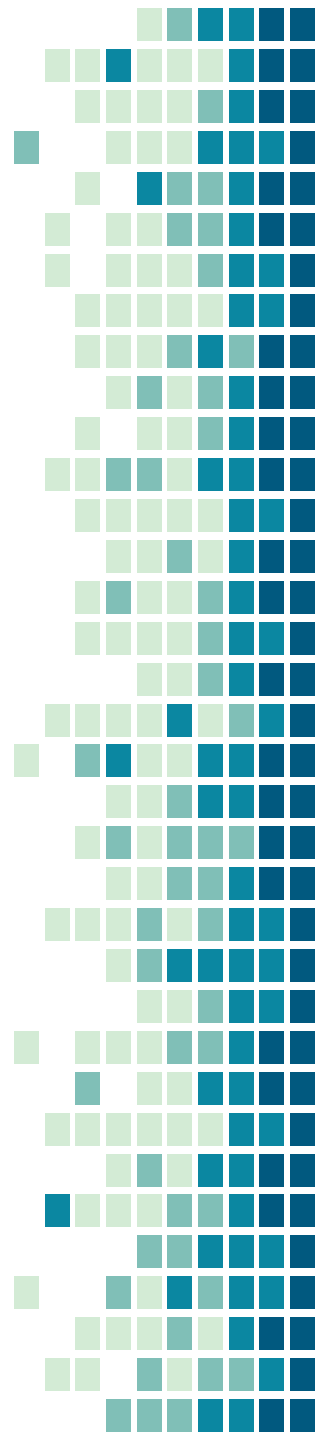
Another idea how to crush gephi!



# Distinctiveness of period shows



Period fiction:  
The White Queen, The  
White Princess, The  
Tudors, Victoria, etc.

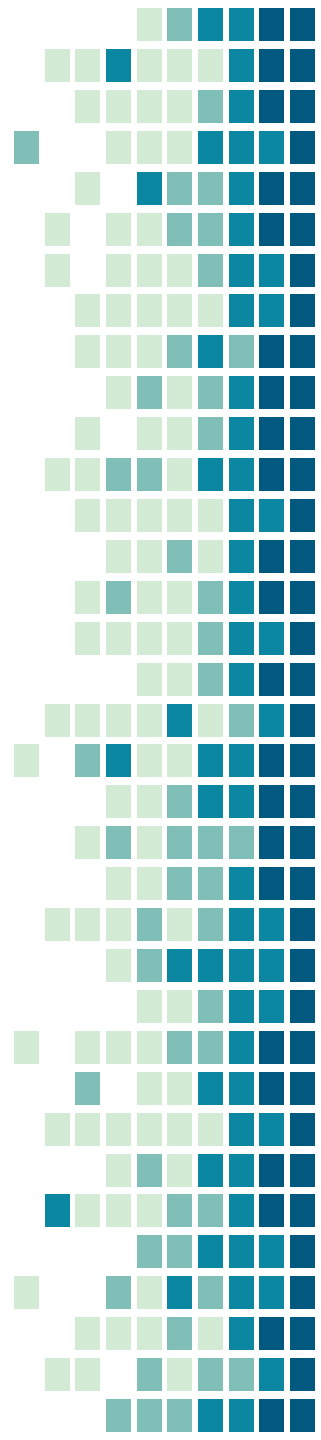




ABC + a little Showtime  
+ BBC's Spooks, no  
common genre or topic



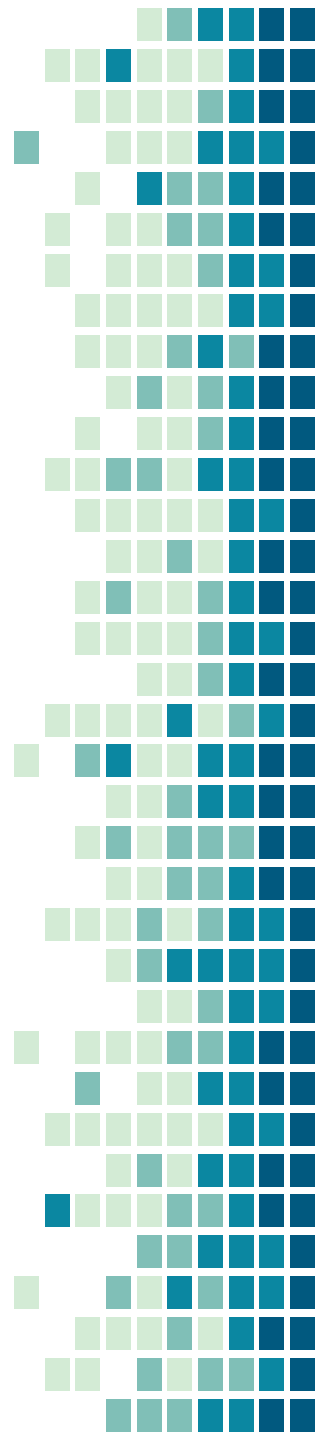
Production company?



# New patterns of collaboration

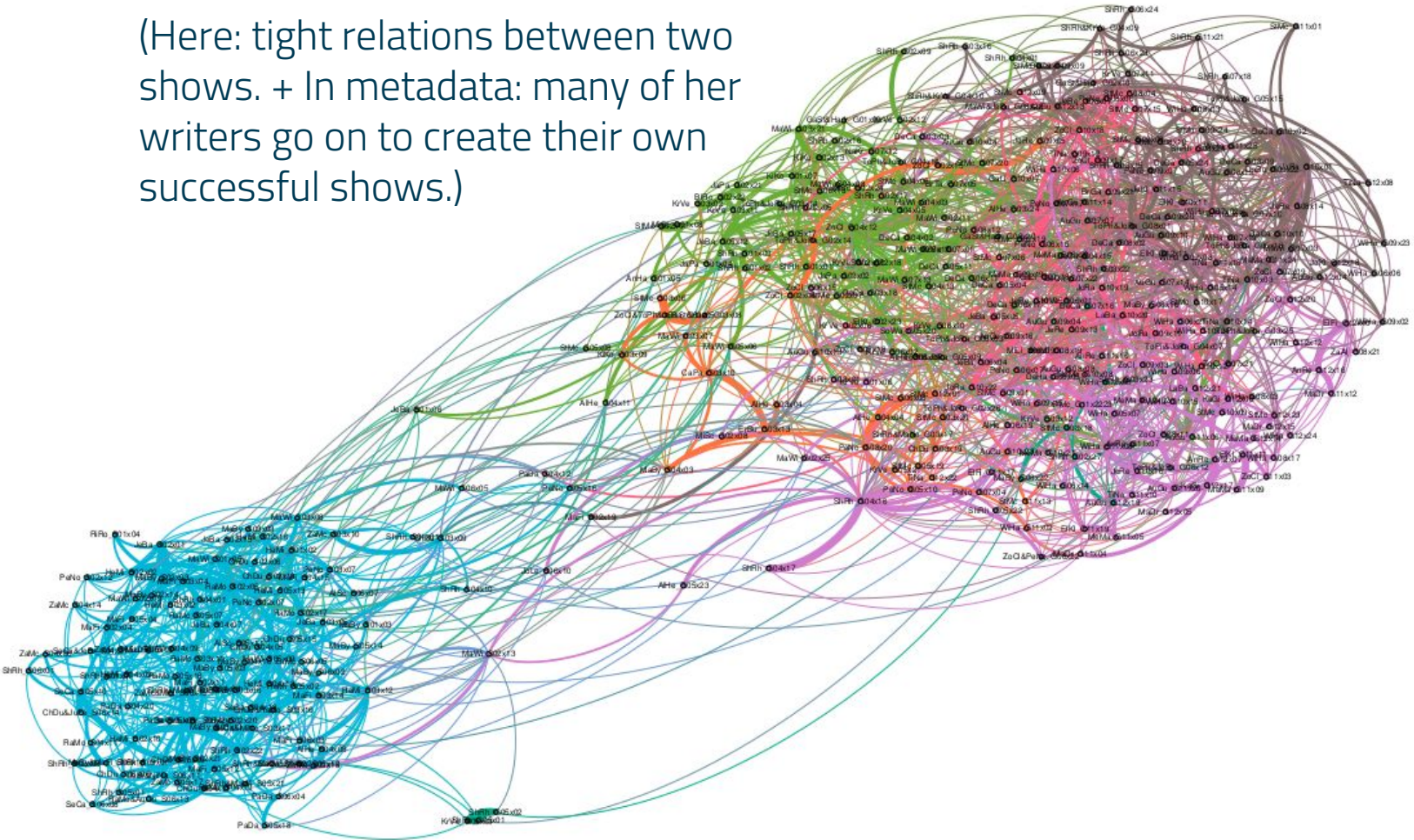
Authorial power couples

Favourite team



This heatmap visualizes the distribution of the number of children per woman across various countries and regions. The color scale indicates the fertility rate, with light green representing lower rates and dark blue representing higher rates. The data is organized into rows and columns, with some rows and columns highlighted in a darker shade of green.

(Here: tight relations between two shows. + In metadata: many of her writers go on to create their own successful shows.)



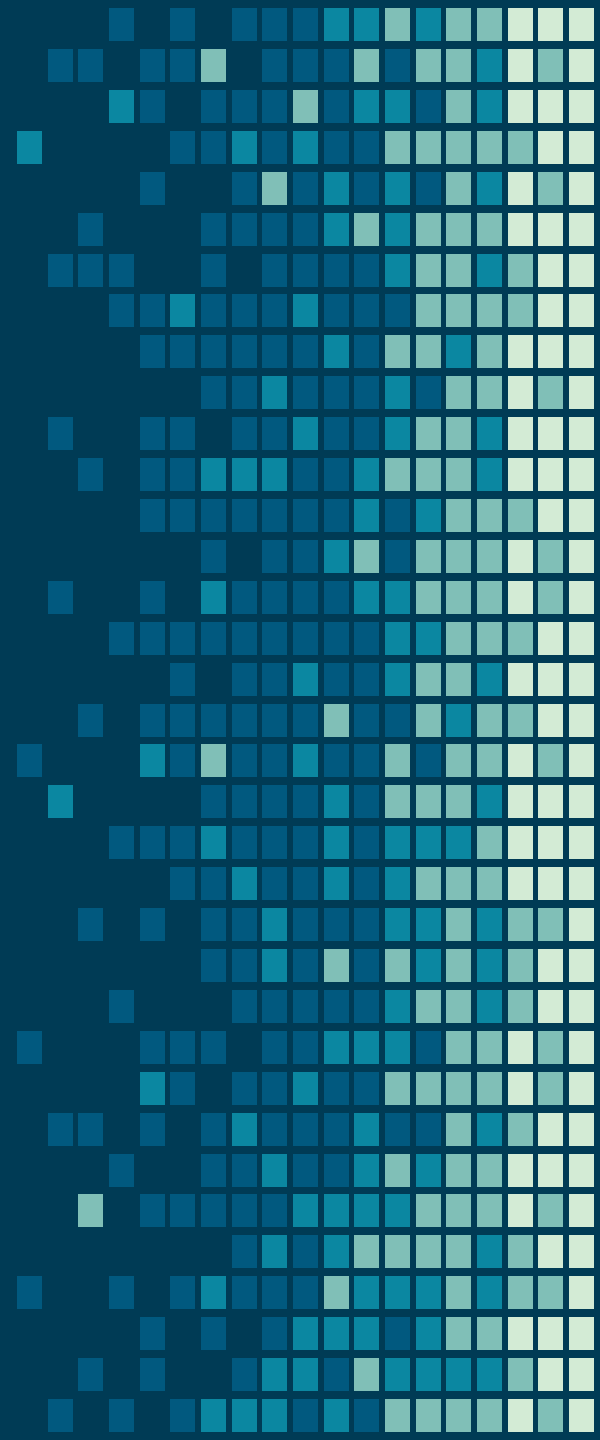


# Tina Fey – favourite collaborators

(Here: two shows, closer relations between authorial tandems of the moment rather than show or topic)

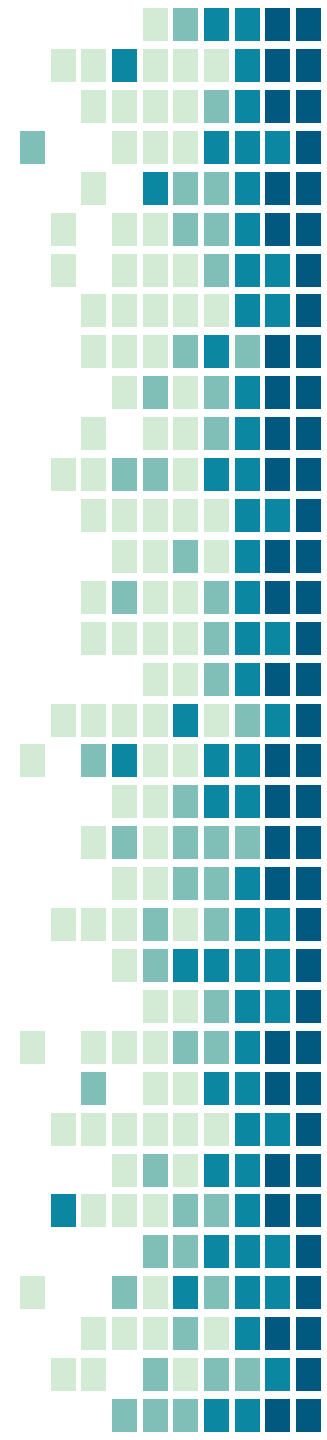


# Future work



# Future work

- More detailed analysis of particular influences
- Evaluating the strength of influences with authorship attribution and verification methods
- Comparison with results of visual and audio analysis



Thank you!

