

discourse network analysis

using



github.com/brunosj/

Bruno St-Jacques
15 June 2019

WHY DISCOURSE NETWORK ANALYSIS?

- Personal interest to explore the debate surrounding the most recent EU Copyright Directive (approved in March 2019)
- DNA allowed the comprehensive mapping of the discourse coalitions engaged in this process and the analysis of their characteristics
- Because networks and R = fun!



DISCOURSE NETWORK ANALYSIS

qualitative content analysis



DISCOURSE NETWORK ANALYSIS

qualitative content analysis



DISCOURSE NETWORK ANALYSIS



social network analysis

Article 13 Open letter Monitoring and Fighting Internet Content

October 16, 2017 · by LibertiesEU

Here is the open letter from 57 signatories to Article 13 of the new proposal on copyright.

Liberties and EDRI initiated an open letter to the European Parliament to withdraw Article 13 from the new copyright proposal.

Over 57 human rights and digital rights groups have signed the letter.

... enabled a process ...

EDRI, BONNIER, ETC, CRAC, SIRE, STIM, CEPIC, OVAP, WAKOM, RIAA, SACEM, stamp, holtzbrinck, Indie, FRA, NME, NME

CALL TO HONOURABLE MEMBERS OF THE EUROPEAN CREATIVE SECTORS ASK FOR

UPHOLD THE MANDATE ON COPYRIGHT IN PLENARY VOTE

We represent 4.5% of EU GDP and 12 million European jobs. We call on you to back the mandate.

On July 5 we ask for you to back the mandate.

There is a growing consensus among the creative sectors.

BA urges MEPs to support controversial copyright law

Published July 5, 2018 by Katherine Cowdrey

Share



Tim Godfrey, executive chairman of the Booksellers Association, has joined publishers in their fight against copyright infringement by voting for the new Copyright Directive.

The BA's efforts join those of the Publishers Association, which has voted in support of approval.

According to the BBC, the new law proposed would put greater responsibility on platforms to prevent copyright infringement, meaning they must have means to assess whether content uploaded to their site infringes copyright (per Article 13 of the Directive).

It also raises the possibility of a "link tax", which in the interests of copyright holders, would prevent online content-sharing platforms and news aggregators from sharing news stories without paying for them (per Article 11 of the Directive). According to Axel Voss, a German MEP, "for a private purpose" would be exempt.

Supporters have said the introduction of the new draft Copyright Directive and its Article 13 would "address the value gap and the ecosystem and its creators, fans and digital music services all

Your internet is under threat. Here's why you should care about European Copyright Reform

Wikimedia Policy | Follow
Sep 4, 2018 · 4 min read

By María Sefidari Huici, Chair, Wikimedia Foundation

You can also read this post in Spanish and in French and in Italian.

Back in 2001, the European Parliament came together to pass regulations and set up copyright laws for the internet, a technology that was just finding its footing after the dot com boom and bust. Wikipedia had just been launched and there were 29 million websites. No one could imagine the rapidly growing ecosystem—and today, the internet has over a billion websites, countless mobile users. We are more interconnected than ever.

But 17 years later, the internet has kept up with us.

The Copyright Directive: Misinformation and Independent Enquiry

The Copyright Directive: Misinformation and Independent Enquiry Statement from European Academics to Members of the European Parliament in advance of the Plenary Vote on the Copyright Directive on 5 July 2018

We have worked with legal, economic and social scientists to bring an independent academic perspective to the public debate surrounding the Proposed Directive on Copyright in the Digital Single Market.

The final version of the EU Copyright Directive is an improvement, but we remain concerned.

Article 17 (formerly Article 13) could still have unintended consequences that may harm Europe's creative and digital economy.

JOINT STATEMENT: THE EU URGENTLY NEEDS REMUNERATION RULES FOR AUTHORS AND PERFORMERS

European and international authors and performers' organisations from all sectors urge Member States to support the much-needed Article 14 providing for fair and proportionate remuneration for authors and performers in the Directive on Copyright in the Digital Single Market.

Proposed by the European Parliament and backed by a large majority of its members on 12 September, the opening provision in Chapter 3 sets a fundamental guiding principle for the remuneration of authors and performers, which should be fair and proportionate to the potential and actual value of the transferred or licensed rights.

Capital and business in the cultural and creative industries have long enjoyed a dominant position over authors and creators, who are grossly underpaid for their work and can barely sustain a decent living as freelancers or employees. The overall regulatory environment does not adequately protect them, as the weakest party, it mostly supports the industry.

Article 13 Open letter Monitoring and Fighting Internet Content

October 16, 2017 · by LibertiesEU

Here is the open letter from 57 signatories Article 13 of the new proposal on copyright.

Liberties and EPP initiated an open letter to Parliament to withdraw Article 15 from the proposal by 57 human rights and digital rights groups.



CALL TO HONOURABLE MEMBERS OF THE EUROPEAN CREATIVE SECTORS ASK FOR

UPHOLD THE MANDATE ON COPYRIGHT IN PLENARY VOTE

We represent 4.5% of EU GDP and 12 million European jobs. We are asking you to back the mandate.

On July 5 we ask for you to back the mandate.

There is a lot at stake.

Article 17 (formerly Article 13) could still have

unintended consequences that may harm Europe's creative and digital economy.



BA urges MEPs to support controversial copyright law

Published July 5, 2018 by Katherine Cowdrey

Share



Tim Godfrey, executive chairman of the Booksellers Association, has joined publishers in their fight against copyright infringement by voting for the new Copyright Directive.

The BA's efforts join those of the Publishers Association, which has also voted in support of approval.

According to the BBC, the new law proposed would put greater responsibility on platforms to tackle copyright infringement, meaning they must have means to assess whether content is infringing or not under Article 13 of the Directive.

It also raises the possibility of a "link tax", which in the interests of copyright holders would prevent online content-sharing platforms and news aggregators from linking to copyrighted material (per Article 11 of the Directive). According to Axel Voss, a German MEP, "for a private purpose" would be exempt.

Supporters have said the introduction of the new draft Copyright Directive and its Article 13 would "address the value gap and improve the ecosystem and its creators, fans and digital music services all over Europe".

JOINT STATEMENT: THE EU URGENTLY NEEDS REMUNERATION RULES FOR AUTHORS AND PERFORMERS

European and international authors and performers' organisations from all sectors urge Member States to support the much-needed Article 14 providing for fair and proportionate remuneration for authors and performers in the Directive on Copyright in the Digital Single Market.

Proposed by the European Parliament and backed by a large majority of its members on 12 September, the opening provision in Chapter 3 sets a fundamental guiding principle for the remuneration of authors and performers, which should be fair and proportionate to the potential and actual value of the transferred or licensed rights.

Capital and business in the cultural and creative industries have long enjoyed a dominant position over authors and creators, who are grossly underpaid for their work and can barely sustain a decent living as freelancers or employees. The overall regulatory environment does not adequately protect them, as the weakest party, it mostly supports the industry.

Your internet is under threat. Here's why you should care about European Copyright Reform

Wikimedia Policy | Follow
Sep 4, 2018 · 4 min read

By María Sefidari Huici, Chair, Wikimedia Foundation

You can also read this post in Spanish and in French and in Italian.

Back in 2001, the European Parliament came together to pass regulations and set up copyright laws for the internet, a technology that was just finding its footing after the dot com boom and bust. Wikipedia had just been launched, there were 29 million websites. No one could imagine the rapidly growing ecosystem—and today, the internet has over a billion websites, countless mobile users. We are more interconnected than ever. But 17 years later, the internet has kept up with us.

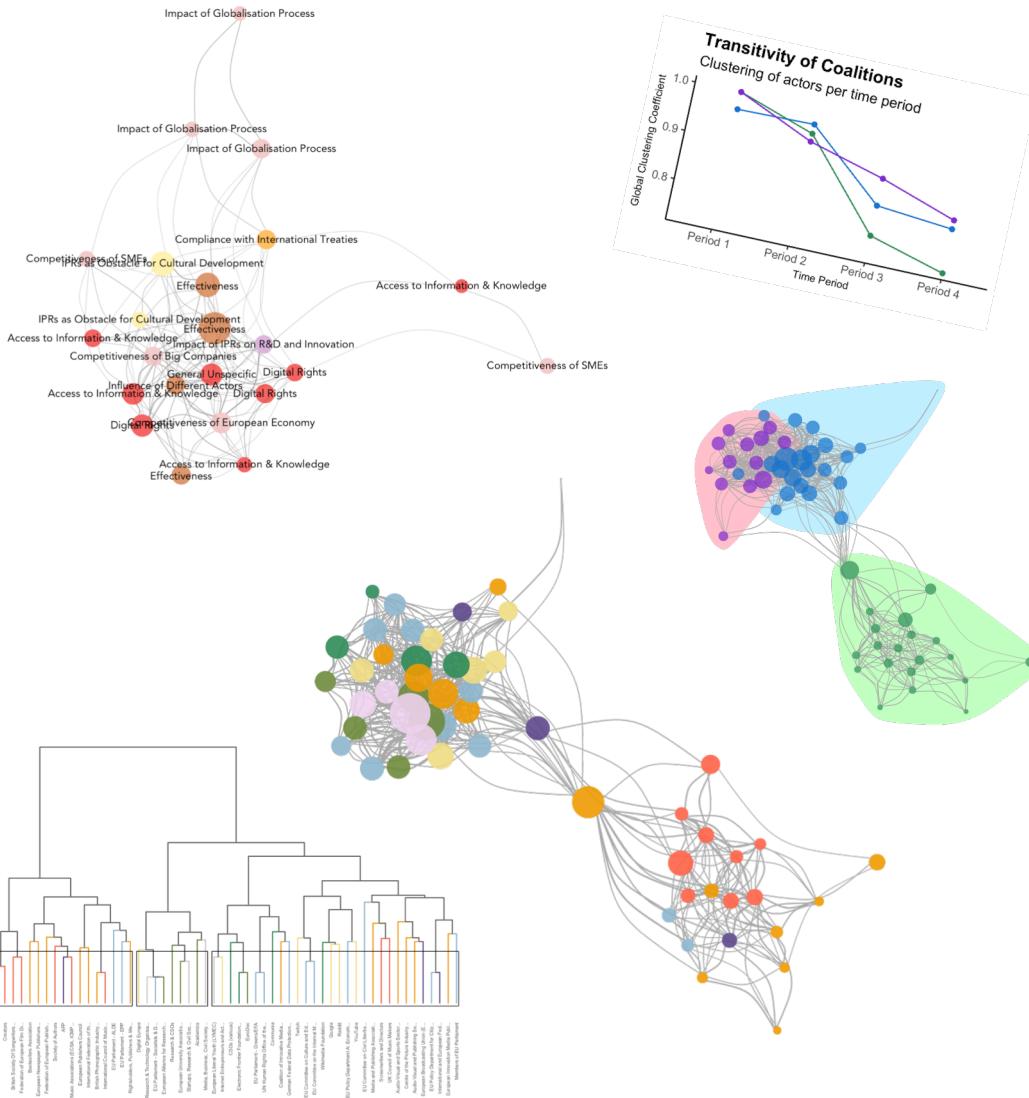
The Copyright Directive: Misinformation and Independent Enquiry

Statement from European Academics to Members of the European Parliament in advance of the Plenary Vote on the Copyright Directive on 5 July 2018

We have worked with legal, economic and social scientists to bring an independent academic perspective to the public debate surrounding the [Proposed Directive on Copyright in the Digital Single Market](#), supported by 25 leading research institutions.

The final version of the EU Copyright Directive is an improvement, but we remain concerned.

Article 17 (formerly Article 13) could still have unintended consequences that may harm Europe's creative and digital economy.



... and much more!

4 EASY STEPS



**Collect and select
text sources**



**Annotate raw text using
the *Discourse Network
Analyzer (DNA)* software**

4 EASY STEPS



**Collect and select
text sources**



**Import the structured data
in R using the package rDNA**



**Annotate raw text using
the *Discourse Network
Analyzer (DNA)* software**



**Explore, visualise and
interpret the data using the
toolbox of network analysis**

MEET YOUR NEW FRIEND, DNA

The screenshot shows the Discourse Network Analyzer (DNA) software interface. The main window has a menu bar with File, Document, Export, and Settings. On the left, there's a Coder panel showing a dropdown for Admin and buttons for edit, add, and delete. Below it is a Document properties panel with fields for Title (90-EFF's Letter to the EU's Cop), Date (2018-10-23 00:00:00), Coder (Admin), Author (Cory Doctorow, EFF), Source (Nonprofits & CSOs), Section, Type (Open Letter), and Notes. At the bottom are Save and Cancel buttons. The central area contains a large text block from Cory Doctorow. A tooltip highlights DNA Statement ID: 186, start: 1339, end: 1689, with options to edit, copy, or delete. Below the statement are four coded variables: person (Cory Doctorow), organization (Electronic Frontier Foundation), concept (Copyright Directive will be deti...), and agreement (checkbox checked). The Statements panel on the right lists numbered statements with their text and IDs (e.g., 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199). The Search within document and Regex highlighter panels are also visible.

4 coded variables:

1. person
2. organisation
3. concept
4. agreement (binary)

+

additional information to
facilitate grouping of
actors/concepts

*"We believe that **Articles 11 and 13 are ill-considered and should not be EU law** [...] they will subvert their stated purpose while **endangering the fundamental human rights of Europeans** to free expression, due process, and privacy."*

*"We believe that **Articles 11 and 13 are ill-considered and should not be EU law** [...] they will subvert their stated purpose while **endangering the fundamental human rights of Europeans** to free expression, due process, and privacy."*

PERSON: Cory Doctorow

Cory Doctorow

ORGANIZATION: Electronic Frontier Foundation (EFF)

Electronic Frontier Foundation (EFF)

CONCEPT: Article 11 and 13 should be included
in the final proposal

Copyright Directive will be detrimental to
fundamental rights and freedoms

AGREEMENT: 0

1

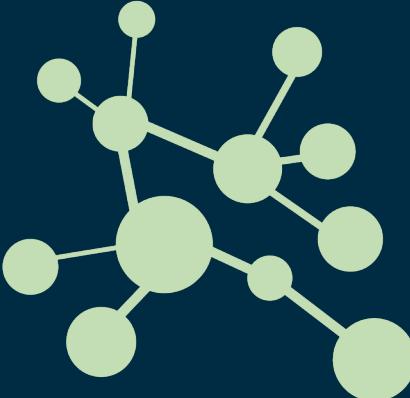
the DNA and R can indeed communicate, great.

but what are **networks**?

why are they **useful**?

```
> # 1. initialize JVM
> .jinit()
[1] 0
> # 2. retrieve the Java-version
> .jcall("java/lang/System", "S", "getProperty", "java.version")
[1] "1.8.0_202"
> # 3. retrieve JAVA_HOME location
> .jcall("java/lang/System", "S", "getProperty", "java.home")
[1] "/Library/Java/JavaVirtualMachines/jdk1.8.0_202.jdk/Contents/Home/jre"
> # 4. retrieve Java architecture
> .jcall("java/lang/System", "S", "getProperty", "sun.arch.data.model")
[1] "64"
> # 5. retrieve architecture of OS (This should have 64 in it if step 4 displays # "64")
> .jcall("java/lang/System", "S", "getProperty", "os.arch")
[1] "x86_64"
> # 6. retrieve architecture of R as well (This should again have 64 in it if # step 4 and 5 display 64)
> R.Version()$arch
[1] "x86_64"
> # 7. load library
> library("rDNA")
> # 8. initialise DNA
> dna_init()
Jar file: /Library/Frameworks/R.framework/Versions/3.5/Resources/library/rDNA/extdata/dna-2.0-beta24.jar
> # open database from rDNA
> set.seed(12345)
> copyright <- dna_connection("copyrightreform.dna", verbose = FALSE)
> # retrieve network matrix
> nw_full <- dna_network(copyright)
(1/5): Processing network options... Done.
(2/5): Filtering statements...
        425 out of 426 statements retained.
(3/5): Compiling node labels...
        56 entries for the first and 45 entries for the second variable.
(4/5): Computing network matrix... Done.
(5/5): Retrieving results.
> |
```

W H Y ?

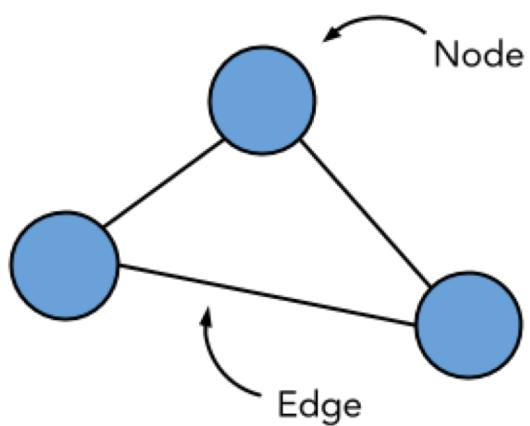


very complex and incomprehensible **amounts of connections** can be made clear and structured

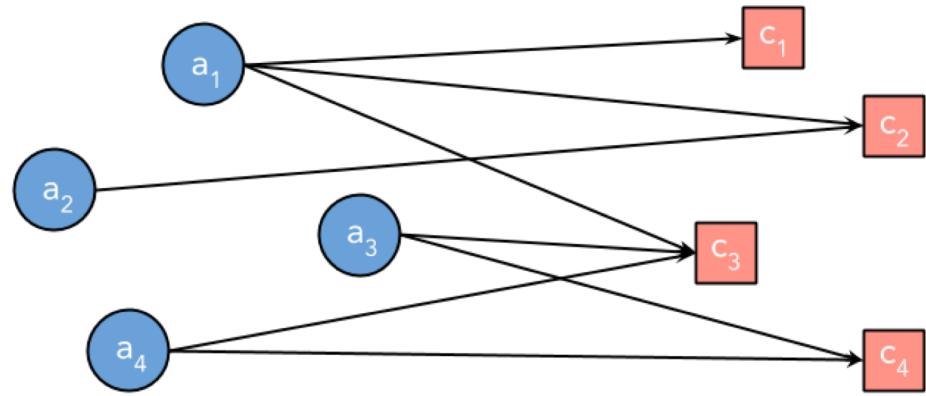
the **relationships between entities** can be measured and interpreted

combined with discourse analysis, the **structure and dynamics of policy debates** can be revealed

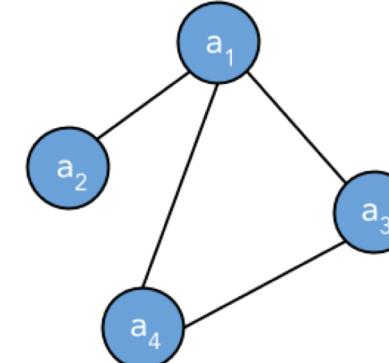
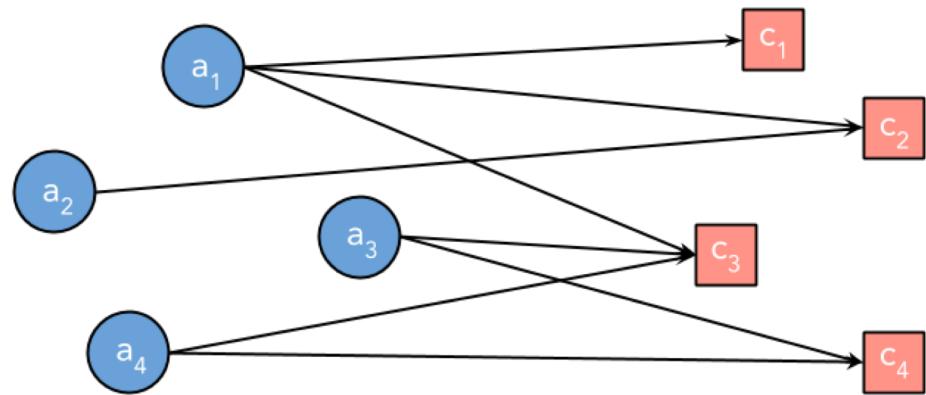
NETWORKS BASICS



At its most simplest form, a network is a collection of points (**nodes** or **vertices**) joined together by pairs of lines (**edges**)



Relations between two different sets of nodes (e.g. '**actors**' and '**concepts**') can be modelled using a **bipartite network**...



Relations between two different sets of nodes (e.g. '**actors**' and '**concepts**') can be modelled using a **bipartite network**...

...from which **congruence** or **conflict** networks can be inferred, allowing to focus on the connections between **actors** (determined by their agreement vis-à-vis **concepts**)

WHEN R TAKES OVER

- The **DNA** can be launched from R to allow data modification/entry
- It is possible to determine which **type of network** (and its related parameters) will be retrieved from the R console

```
#-----  
# RETRIEVE NETWORKS FROM THE DNA  
#-----  
  
# compute congruence network (agreement/disagreement b/w organizations)  
congruence_A <- dna_network(copyright,  
                           networkType = "onemode",  
                           statementType = "DNA Statement",  
                           normalization = "average",  
                           variable1 = "organization",  
                           variable2 = "concept",  
                           qualifier = "agreement",  
                           qualifierAggregation = "congruence",  
                           duplicates = "document")  
  
# convert to igraph object  
net_igraph <- dna_toIgraph(congruence_A, weighted = TRUE)  
  
# retrieve actors attributes  
attr_A <- dna_getAttributes(copyright,  
                           statementType = "DNA Statement",  
                           variable = "organization")
```

Two-mode (bipartite) network matrix

| | Article 11 (Neighbouring Right) should be deleted | Article 11 (Neighbouring Right) should be modified | Article 11 (Neighbouring Right) will add another layer of complexity to licensing deals | Arti (Nei Righ imp free info |
|---|--|---|--|---|
| AFP | 0 | 0 | 0 | |
| Academics | 0 | 2 | 0 | |
| Audio-Visual and Publishing Sectors Organisations | 0 | 0 | 0 | |
| Audio-Visual and Sports Sectors Organisations | 0 | 0 | 0 | |
| Authors and Performers Organisations | 0 | 0 | 0 | |
| Booksellers Association | 0 | 0 | 0 | |
| British Phonographic Industry (BPI) | 0 | 0 | 0 | |
| British Society Of Songwriters, Composers And Au... | 0 | 0 | 0 | |
| CSOs (various) | 0 | 0 | 0 | |
| Centre of the Picture Industry (CEPIC) | 0 | 0 | 0 | |
| Coalition of Innovative Media Publishers | 0 | 1 | 0 | |
| Communia | 0 | 1 | 0 | |
| Creators | 0 | 0 | 0 | |
| Digital Europe | 0 | 1 | 0 | |
| EU Committee on Civil Liberties, Justice and Home ... | 0 | 0 | 0 | |
| EU Committee on Culture and Education | 0 | 1 | 0 | |
| EU Committee on the Internal Market and Consum... | 0 | 1 | 1 | |
| EU Parliament - ALDE | 0 | 0 | 0 | |
| EU Parliament - EPP | 0 | 0 | 0 | |
| EU Parliament - Greens/EFA | 0 | 0 | 0 | |
| EU Parliament - Socialists & Democrats | 0 | 0 | 0 | |
| EU Policy Department A: Economic and Scientific P... | 0 | 0 | 0 | |
| EU Policy Department for Citizens' Rights and Con... | 0 | 1 | 0 | |
| Electronic Frontier Foundation (EFF) | 1 | 0 | 0 | |
| EuroDoc | 0 | 0 | 0 | |

One-mode (congruence) network matrix

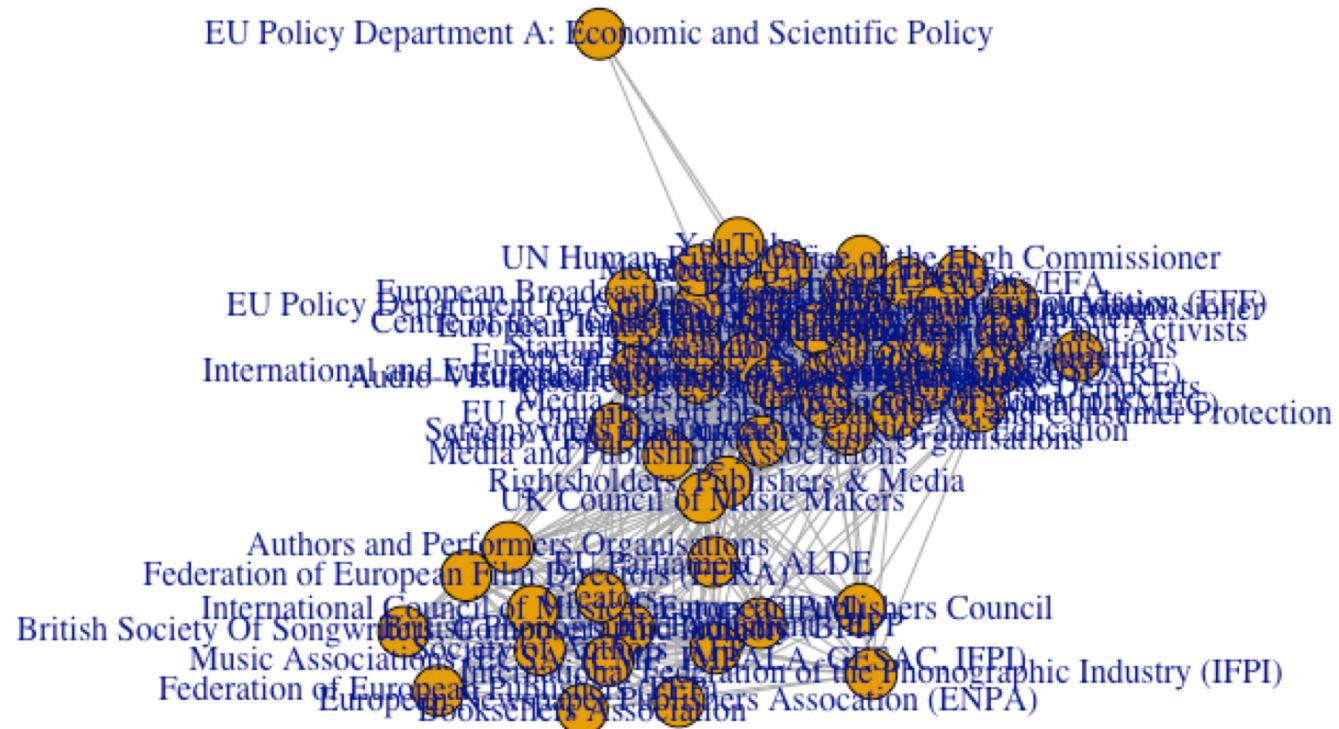
| | AFP | Academics | Audio-Visual and Publishing Sectors Organisations | Audio-Visual and Sports Sectors Organisations | Aut Per Org |
|---|-----------|------------|---|---|-------------|
| AFP | 0.0000000 | 0.0000000 | 0.0000000 | 0.1538462 | |
| Academics | 0.0000000 | 0.0000000 | 0.62857143 | 0.1935484 | |
| Audio-Visual and Publishing Sectors Organisations | 0.0000000 | 0.62857143 | 0.0000000 | 0.4444444 | |
| Audio-Visual and Sports Sectors Organisations | 0.1538462 | 0.19354839 | 0.44444444 | 0.0000000 | |
| Authors and Performers Organisations | 0.1818182 | 0.00000000 | 0.00000000 | 0.0000000 | |
| Booksellers Association | 0.5000000 | 0.00000000 | 0.00000000 | 0.0000000 | |
| British Phonographic Industry (BPI) | 0.4615385 | 0.00000000 | 0.11111111 | 0.2857143 | |
| British Society Of Songwriters, Composers And Au... | 0.2222222 | 0.00000000 | 0.00000000 | 0.0000000 | |
| CSOs (various) | 0.0000000 | 0.64705882 | 0.47619048 | 0.3529412 | |
| Centre of the Picture Industry (CEPIC) | 0.0000000 | 0.48275862 | 0.62500000 | 0.3333333 | |
| Coalition of Innovative Media Publishers | 0.0000000 | 0.55172414 | 0.50000000 | 0.1666667 | |
| Communia | 0.0000000 | 0.46666667 | 0.00000000 | 0.0000000 | |
| Creators | 0.4000000 | 0.00000000 | 0.00000000 | 0.1818182 | |
| Digital Europe | 0.0000000 | 0.46666667 | 0.35294118 | 0.1538462 | |
| EU Committee on Civil Liberties, Justice and Home ... | 0.0000000 | 0.00000000 | 0.00000000 | 0.0000000 | |
| EU Committee on Culture and Education | 0.1428571 | 0.43750000 | 0.31578947 | 0.1333333 | |
| EU Committee on the Internal Market and Consum... | 0.1333333 | 0.60606061 | 0.30000000 | 0.1250000 | |
| EU Parliament - ALDE | 0.3076923 | 0.06451613 | 0.00000000 | 0.0000000 | |
| EU Parliament - EPP | 0.4705882 | 0.00000000 | 0.09090909 | 0.2222222 | |
| EU Parliament - Greens/EFA | 0.0000000 | 0.35714286 | 0.26666667 | 0.3636364 | |
| EU Parliament - Socialists & Democrats | 0.0000000 | 0.46666667 | 0.35294118 | 0.1538462 | |
| EU Policy Department A: Economic and Scientific P... | 0.0000000 | 0.00000000 | 0.00000000 | 0.0000000 | |
| EU Policy Department for Citizens' Rights and Con... | 0.0000000 | 0.30769231 | 0.30769231 | 0.2222222 | |
| Electronic Frontier Foundation (EFF) | 0.0000000 | 0.40000000 | 0.23529412 | 0.1538462 | |
| EuroDoc | 0.0000000 | 0.21428571 | 0.26666667 | 0.1818182 | |
| European Alliance for Research Excellence (EARE) | 0.0000000 | 0.57894737 | 0.48000000 | 0.2857143 | |
| European Broadcasting Union (EBU) | 0.0000000 | 0.22222222 | 0.71428571 | 0.2000000 | |

VISUALISATION WITH IGRAPH

```
plot(net_igraph)
```

VISUALISATION WITH IGRAPH

```
plot(net_igraph)
```



... some transformations are needed!

Transformations:

- simplify graph
- set node size and edge width
- add edge threshold
- assign node colours according to actor type
- plot graph with additional aesthetic tweaks

```
#-----
# VISUALISATION WITH IGRAPH
#-----

# simplify graph (remove loops)
net <- simplify(net_igraph, remove.multiple = F, remove.loops = T)

# compute node degrees (weighted) and use this value to set node size:
deg <- igraph::graph.strength(net, mode="all")
V(net)$size <- deg

# set edge width based on weight:
E(net)$width <- E(net)$weight

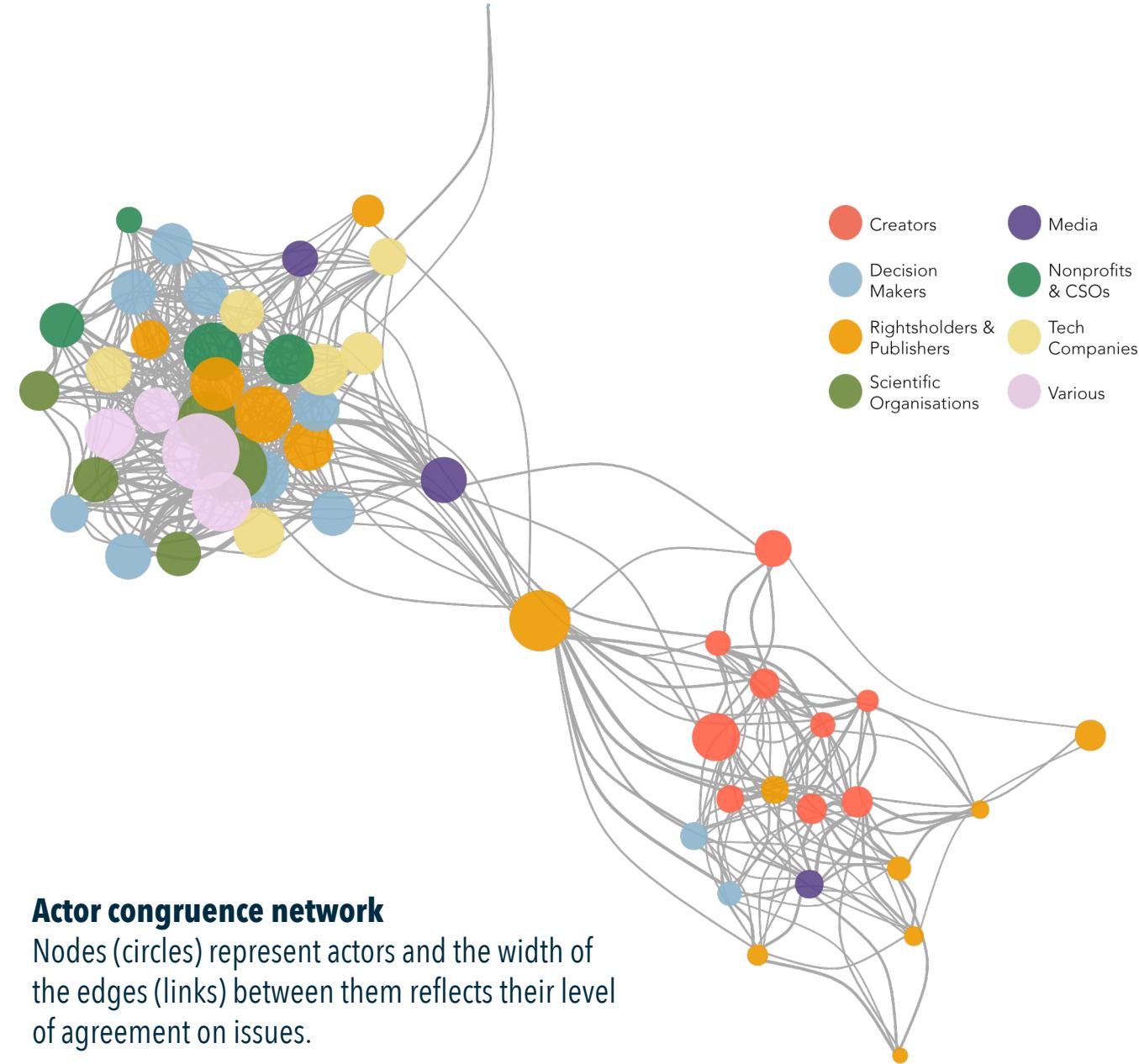
# edge threshold to remove low-intensity ties
summary(E(net)$weight)
net <- igraph::delete.edges(net, which(E(net)$weight <0.30769)) # median value of edges

# set attributes to igraph object (retrieved from the previous attribute object from the DNA)
net <- set_vertex_attr(net, "type", index = V(net), as.character(attr_A$type))

# generate colors based on actor type:
V(net)$color <- V(net)$type
V(net)$color=gsub("Creators","tomato",V(net)$color)
V(net)$color=gsub("Decision Makers","lightskyblue3",V(net)$color)
V(net)$color=gsub("Media", "mediumpurple4",V(net)$color)
V(net)$color=gsub("Nonprofits & CSOs", "seagreen4",V(net)$color)
V(net)$color=gsub("Rightsholders & Publishers", "orange2",V(net)$color)
V(net)$color=gsub("Scientific Organisations", "darkolivegreen4",V(net)$color)
V(net)$color=gsub("Tech Companies", "lightgoldenrod2",V(net)$color)
V(net)$color=gsub("Various", "thistle2",V(net)$color)
colors.vect <- as.factor(V(net)$color)

# plot new version of graph
plot.igraph(net,
            vertex.label=NA,
            edge.color = "darkgrey",
            vertex.color = adjustcolor(colors.vect, alpha.f = .9),
            edge.curved=.5,
            vertex.frame.color = NA,
            layout=layout_with_fr)
```

et voilà!



COMMUNITY DETECTION

- There are dozens of community detection algorithms (cluster optimal, fast greedy, edge betweenness, multilevel, random walk, etc.)
- They try to find dense subgraphs in graphs, by optimising some criteria and usually using heuristics
- Using the cluster optimal algorithm, 3 communities were identified, with respectively 20, 22 and 13 actors

COMMUNITY DETECTION

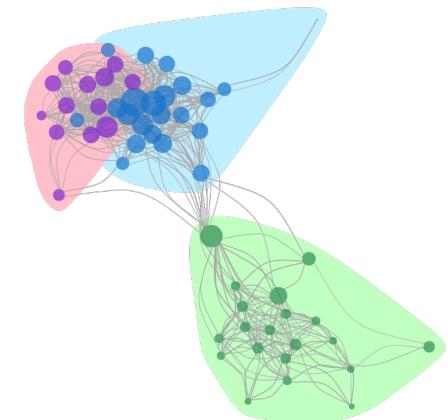
- There are dozens of community detection algorithms (cluster optimal, fast greedy, edge betweenness, multilevel, random walk, etc.)
- They try to find dense subgraphs in graphs, by optimising some criteria and usually using heuristics
- Using the cluster optimal algorithm, 3 communities were identified, with respectively 20, 22 and 13 actors

```
#-----
# COMMUNITY DETECTION
#-----

# community detection
clp <- cluster_optimal(net)
sizes(clp)

# set vertex community attribute according to membership
V(net)$community <- clp$membership

# plot graph with communities
cols <- adjustcolor(c("seagreen4", "dodgerblue3", "purple3"),
                     alpha=.9)[membership(clp)]
plot(clp, net,
      col=cols,
      mark.border="black",
      mark.col=c("darkseagreen1", "lightblue1", "pink"),
      edge.curved=.5,
      vertex.label=NA,
      edge.arrow.size=.2,
      vertex.frame.color = "NA",
      layout=layout_with_fr)
```



(SOME) NETWORKS PROPERTIES



Centrality

the number of edges connected to an individual node



Transitivity

a measure of the clustering in a network, based on the relative number of triangles in the graph



Modularity

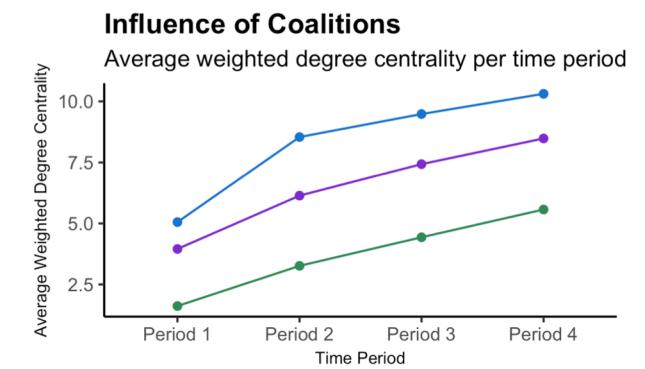
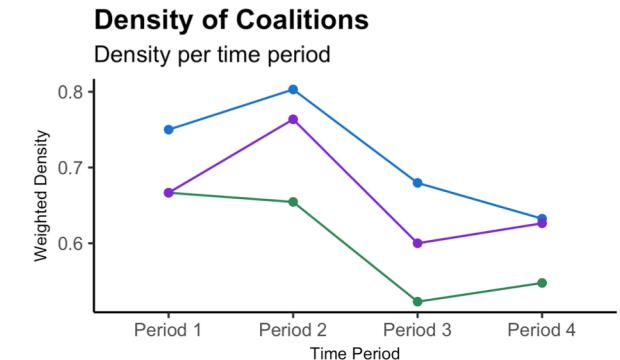
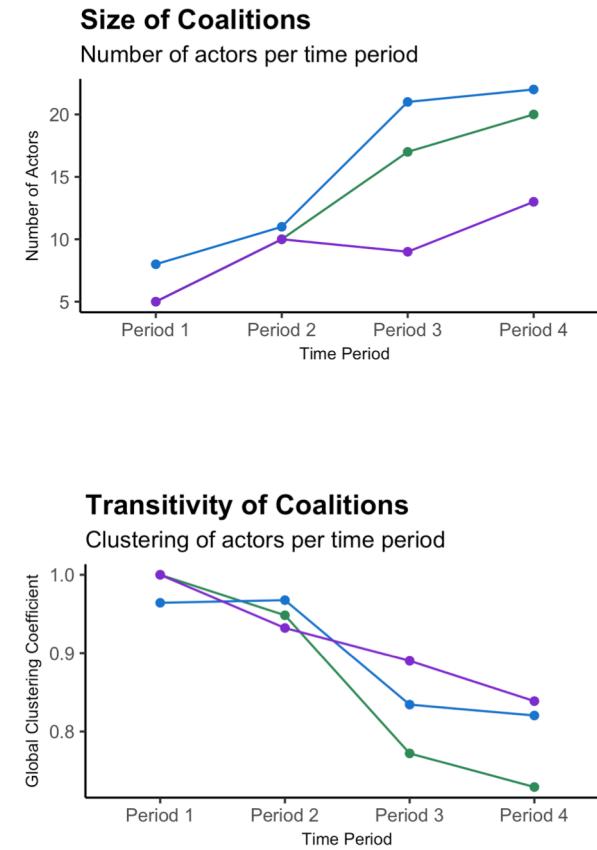
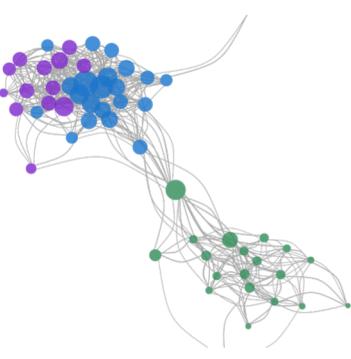
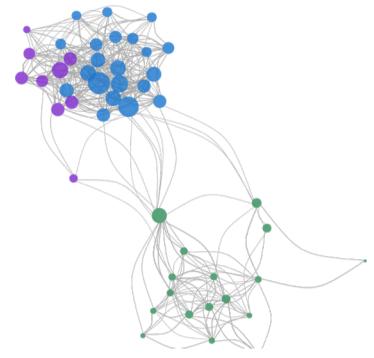
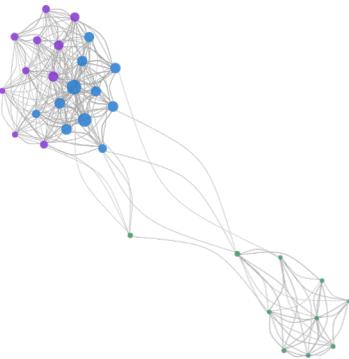
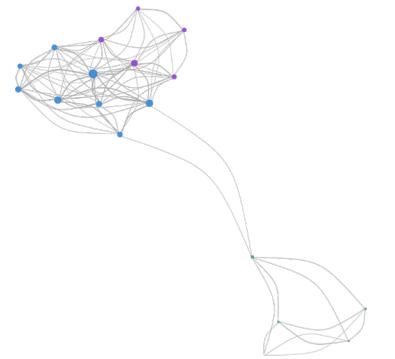
a measure of the structure of networks, designed to measure the strength of division of a network into modules



Density

a measure of the connectedness in a network, defined as the actual number of ties expressed as a proportion of the maximum possible number of ties

DYNAMIC EVOLUTION OF NETWORKS



LEARN MORE

Head over to github.com/brunosj/ to consult:

- replication materials
- thesis document + poster
- other projects!

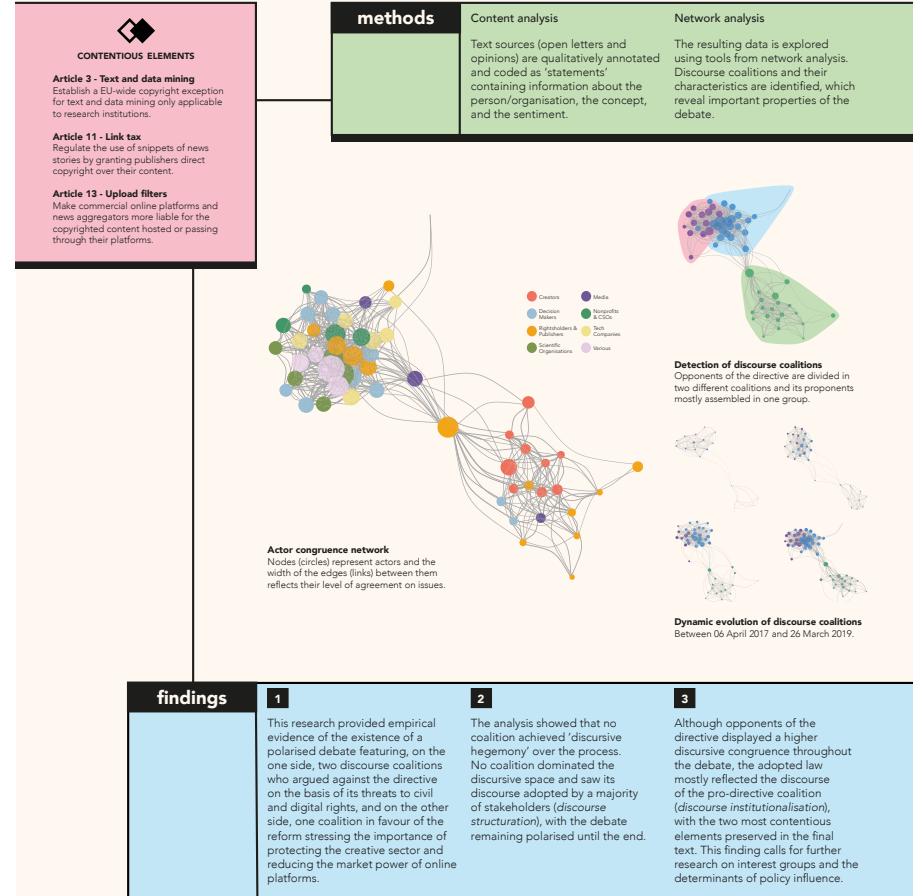
us and them

MAPPING DISCOURSE COALITIONS IN THE EU COPYRIGHT DIRECTIVE DEBATE

The European Parliament approved the most recent Copyright Directive in March 2019, a reform aiming to make copyright law fit for the information age. Its adoption ended a 3-year legislative process punctuated by a fierce debate

between its proponents and critics. Exploring the different ways copyright legislation is being justified and negotiated, this research combines Hajer's argumentative discourse analysis framework with quantitative social

network analysis to comprehensively map the discourse coalitions engaged in this debate, scrutinise their characteristics, and appraise their influence over the policy process and outcome.





THANKS !



github.com/brunosj/

Bruno St-Jacques
15 June 2019
bpstjacques@gmail.com