



# Culture Analytics and the Humanities: Literary and Cultural Data

Culture Analytics Tutorials



# Overview

- Folklore
  - Problems in classification
  - Mapping, GIS and thick cultural description
- Old Icelandic literature
  - Challenges with highly inflected languages
  - Character Networks as social networks
- Things I will not have time to address
  - Challenges in the study of literature
    - Influence and translation networks (meta-data)
    - Reading at scale
    - Topic modeling as search engine



# Fast furrier transform



# Graf - Laplacian



+



=





# Folklore

- “Folklore emerges from the dialectic tension between the individual and tradition”
- Folklore is predicated on the study of variants of expressive forms
- Folklore has a long history of classifying 10s of thousands of stories as a means for being able to work at the macro-, meso- and micro-levels

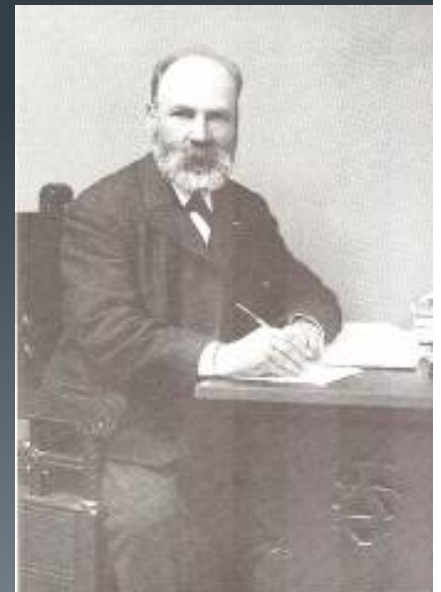
# Challenges

- Data acquisition, data structur(ing), data access
  - A pipeline
- Classification, Pattern Discovery and Navigation
- Given a very large, poorly labeled folklore collection, how do I:
  - Discover resources (stories) that will help me address a research question?
    - E.g.: I am interested in ghosts. How do I find stories about ghosts?
  - Discover patterns in the collection that might lead to new research questions?
    - E.g.: Are there any geographical distribution patterns that might point to a prominent belief associated with a particular area?
    - In other words, is that haunted house really haunted?



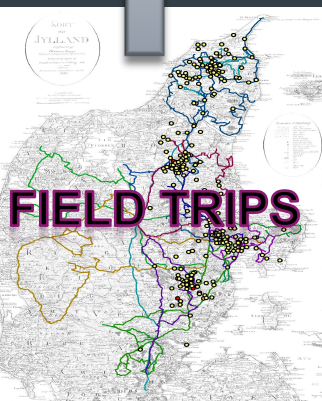
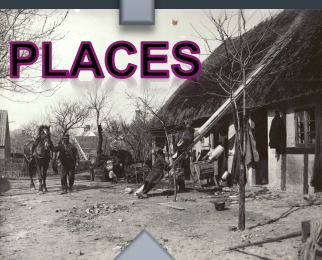
# Target Corpus

- Danish school teacher, Evald Tang Kristensen (1843-1929)
  - Actively collected from 1865-1923
- Stories collected
  - from ~4500 informants
  - in ~2200 places (mostly in Jutland)
  - mentioning ~12,000 places (also mostly in Jutland)
- 24,000 manuscript pages
  - Covering ~210 fieldtrips
  - 6,000 additional loose leaf pages
  - 2,000 additional pocket diary pages
- 80 published volumes
  - ~100,000 entries
  - 12 separate indices

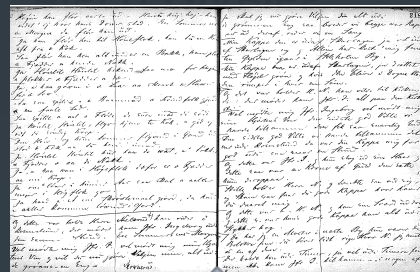


# Understanding the Collection

## PEOPLE

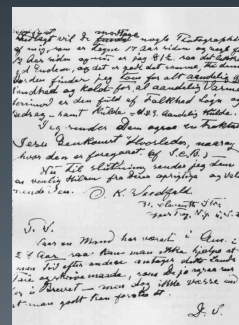


## FIELD DIARIES

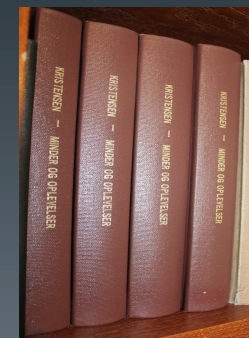


ETK

## PUBLISHED WORKS



LETTERS



MEMOIRS



# Making the collection machine actionable

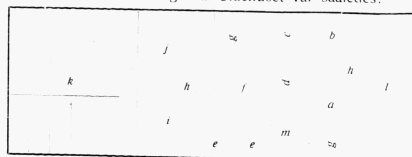


# Challenge: Digitization and Archiving

- Scanning
  - Condition of materials
- OCR
  - Language
  - Fonts and printing
- Indexing
  - Tagging – TEI
  - Named Entity Detection and Resolution
    - Natural Language Processing (NLP)
- Semi-supervised and unsupervised methods



Gaarden, men maatte hente Vand ved en Dam, et Par Boseskud vesten for. Der var en Have med lidt Frugtbuske og Stikkelsbærriis. Stuehuset laa i Øst og Vest, og Laden laa for Norden i Gaarden. I den fandtes ogsaa Faarehus og Grisehus, men Kohuset var i Østrenden af Stuehuset. For Østen i Gaarden stod Tørvestakken, og andre Huse var der ikke. Indretningen af Stuehuset var saaledes:



a. Aftægtsstue. b. Deres Kjøkken dertil. c. Spisekammer. d. Seng. e. Senge. f. Ejeren's Stue. g. Borde. h. Skorstene. i. Kjøkken. j. Spisekammer. k. Kohus. l. Høgtulv. m. Forgang.

Alle Husrum er nu jævned med Jorden, og det Stuehus, jeg har været inde i, er ogsaa brækket ned for nogle faa Aar siden. Jeg har ligget i det et Par Nætter. For Tiden er der en meget pæn ny Plantorboliig.

Endnu vil jeg tilføje, at Jens Albertsens ene Søn er Lærer, og den anden har i flere Aar været Missionær i Indien, men er nu kommen hjem.

Min Oldefader havde følgende Born:

1. *Jens Christensen*, kaldet *Jens Mikkelpborg*, født 1760. Han blev bosat i Mikkelpborg i Vind og døde som Aftægtsmand 5. April 1837, 78 Aar gl. Hans Kone var en Datter af *Niels Skolde* og døde samme Aar som Manden. Han har efterladt sig en talrig Slægt, der til Dels endnu boer i Vind.

2. *Christen Christensen Garsdal*, døbt 12. Juli 1761. Han døde 3 Uger gl.

3. *Christen Christensen Garsdal* (senere kaldet Skaderis). Dette er altsaa min Bedstefader. Han blev døbt 10. Sønd. eft. Trin. 1764 og døde 16. Juni 1831, 67 Aar gammel. Han lærte Skrædderprofessionen og blev 11. Oktober 1793 gift med *Else Katrine Knudsdatter i Bursig i Bur Sogn*. Hun er da min Bedstemoder. De blev viet i Ulfborg Kirke. Hendes Forældre var *Knud Christensen* og

*Inger Sig.* hun var døbt 9. Febr. 1772 og døde 6. Maj 1856, lidt over 84 Aar gl. Hun kaldtes almindelig gamle *Else Skaderis*, og jeg kan lige mindes hende fra et Besøg, jeg sammen med mine Forældre aflagde hos hende i Skaderis. Det var vist nok 1850. Imens min Fader var Hjælpe lærer paa Holmsland, gik hun hver Sommer den lange Vej ned for at besøge gamle, da vi Besøg siden.

Jeg har f. Broder Knud og da det er skrive det he

For kort hjemme, at

erføer af Fader

en Rejse did kom

lykkelig kan jeg ikke

dende haver I en halv Sn

han travlt n og andre Sn

var han ned Dette var de

han gik hjem Søndag om

gen, og fra Sengen og

Ende paa h genen. I F

Vattersot, ti at det var

men formoc Et meget va

os om at gj dommen ha

begravet de borg d. 21.

vi det for b Moder

mer gaar d

731. Go daw-i då,  
ka I kjennd mæ en rå?  
a sku ha en stång,  
å dænd sku væ lång,  
vi hå en tække (o: tække)  
han æ så gal, te hans rø  
732. Her hvil'er han så simpel  
den fordrukne bimpel,  
al verden var han til spo  
slet ingen mennesker gjo  
Gud ske lov, han er dod  
733. Söwel (o: Sevel) post kjo  
de mo I trow,  
han kjor te Struer,  
dæær fek han æen fuld s  
da han kjor hjem,  
wa han et slæt så slem,  
så kam han i hans seng  
o sow æ rus hæen.  
734. Først gav han dem havre  
så blev de mavre,  
så gav han dem byg,  
så blev de syg,  
så gav han dem rug,  
så slog de op i et trug,  
så gav han dem sød mælk af en kop,  
så slog alle Per Jensens grise op.  
735. Pof, pof, py—y—y,  
mi muer koft grød,  
fæer de fætrød,  
hå I så kjend mååg,  
hun koft dem i lud,  
å såldt dem mæ krud,  
a vild åld-er smååg em.

T. Kristensen.

## 1. At holde Kede paa Søndagen.

1. De var en Præst i Fellske, han arbejdet lisse wal som ande Vonder å war bådte Buund å Præst. Men han håd dænd Svååghjejed, te han kund et wal hånst Ree o æ Dawu, å så håd han Limmer eund i en Kammer, å damm flot han jæn aw få hwæ Daw. Men så håd æ Rowen jæn Daw tæjen (tauen) jæn å fæjj mæj, å hon håd jat en te fæjj Si, dænd Gång hon håd en brugt. Så blow han jo wifd o et,

å de blow Søndde, å han wifd  
Så sejer han te fæ fæjj o æ  
studejret i Daw, men da de wa  
te han wifd nejer å groof. S  
go te Kjerf, å da han sejer bæ  
de war Søndde. Så stufd han  
å stufd åw ha nåwæd å low  
ve å fog Kæf, å så taw han e  
dæær i æ Kållgrynd, får han  
Gænd, får han stu jo loos nåw  
å fofd nejer i æ Kållgrynd, å så  
hænd en da åp o æ Kof å en E  
å dæær fæf æ Hund fat o en,  
han jo et mejer fæf, å han mo  
o æ bår Bejen å fomm i æ K  
Så stu han jo te å begomnd  
a wæ et, å a wæ nåwæd, I wæ  
i æ Stowuf, de wæ a, de wæ  
nåwæd te en Par Høwæjer, de i  
fæder i æ Hunds Row. Amen.  
håd jo et fæt løst. Stowider  
Her figes: groof, men i Grind

modtage en mont til bruden til nålepenge. Af alt det indkomne tilhører den største mont bryllupsmanden. For flere år tilbage fulgte en pige af den nærmeste slægt med til hest og forrettede brudesengs-indbydelsen. Egen. Nis Nielsen, Nørborg.

155. Bydemændene kaldes de to ungarle, der går omkring og indbyder til bryllup. Hos præsten lyder indbydelsesformularen således: »A här en flittig hilsen från ungkål N. N., liksom ässe från hans kjerrest N. N., at Di vel komm te kjerken næstkom-mend toosde å gjeer djer brundviels å felle hjäm mæ däm å wæ djer indbudden gjæst, så lænng Di lester«. Efter at den ene har sagt disse ord, siger den anden: »De ær ässe minn öwr«. Nordby. Addit. 365 4to. Fr. Jacobsen.

156. Byderemse fra Vejle-egnen: »Jeg har en venlig og kjerlig indbydelse til eder fra [Mads Andersen] og hustru på Jerlev mark. Da det har behaget den alwise og gode Gud at stifte et kristeligt og kjerligt ægteskab imellem ungkarl [Johan Thomsen] af Højen og pigen [Elise Nielsen] af Ammitsbøl, så har de da besluttet — næst Guds hjælp — at lade denne deres hellige ægteskabsforbindelse foretage næstkommande fredag, som indtræffer den 23. oktober. Det er derfor deres venlige begjæring til eder, at I vil lade eder indfinde i Mads Andersens hus og gård for at ledsage brudeparret til Jerlev kirke og der være til stede ved en kristelig kombination (!) eller brudevielse, og når disse hellige forretninger i Guds hus ere til ende, da at ledsage brudeparret til brudehuset for at lade jer opvarte med et æres-måltid, som for jer og flere gjæster vil blive frembåren, og der-efter at deltage i forskjellige uskyldige fornøjelser, så længe I selv synes, og lysten vedværes«. Efter afleveringen af denne ramse havde bydesvenden at tilføje i bondemålet: »Vil I så wæ så gue å sto mæ æ båd te«.

Foresagt mig af min farbroder Søren Lavrsen i Store-Lime, som selv var bydesvend. — Det var på den egn spillemandens sag at forfatte byderamsen, og ovenstående er forfattet af en meget bekendt landsly-spillemand i Ammitsbøl, Jens Kristian.

Lærer Karl Lavrsen, Give.

157. I Tinglew boede en for sin gjerrighed berygtet peber-svend ved navn Ross. Han ejede nogle brogede gardiner, som aldrig brugtes undtagen ved bryllup, ti så brugtes de til at klæde sal med. En gang kom et par nabopiger og bad om at låne dem. »Ja, dem skal I sku få straks, men æ sku sands en lidt fæst«. Derpå tog han dem, lagde dem på en blok, tog sin økse og hug dem i tusende stykker til pigernes store forskrækkelse. »Se så«, sagde han, »no her sku den lånen åp, no kan I tæjj





JP.U.64.

MINERVA MIKROFILMA/s

## EVALD TANG KRISTENSEN

DAGBØGER

1. BIND

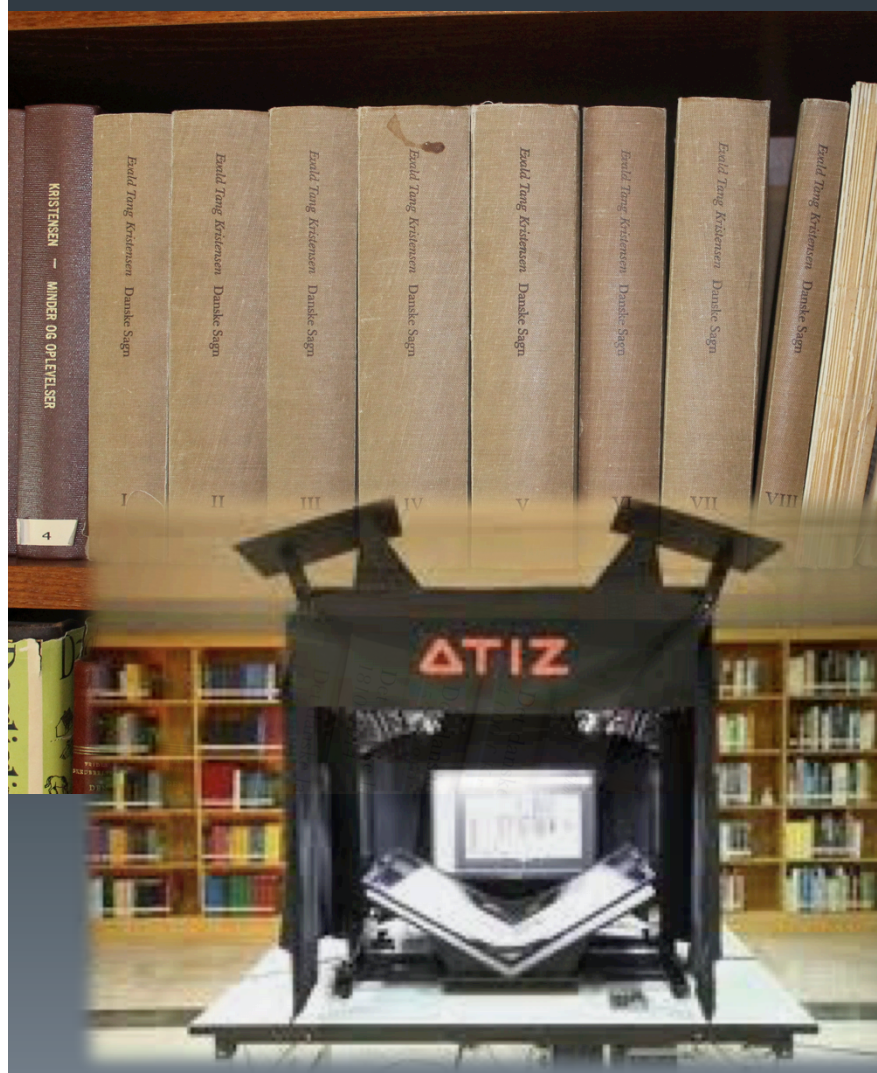
BLAD 1-342

HEFTE 1- 15

DFS 1929/16

1871. April.	
1-2 b.	Mrs. Pedersen, Lønd.
2 b-3 a.	Jens Christian Madsen, Rind.
3 a-10 a.	Ane Johanne Jensdatter
10 a.	Mads Møller, Lønd.
10 b-14 b.	Evel Thomassen, Rind.
14 b-16 b.	Kjærulff, Rind.
15 a-16 b.	Eise Skole, Bræ.
16 b-17 a.	Ane Jensdatter, Lønd. Hede.
17 a-18 a.	Kristen Mørings, Rind.
18 a-19 b.	Kristen Mørings, Tjørning.
19 b-20 b.	Morten Petersen, Tjørning.
20 b-23 b.	Karen, Kirst. Hansen, Tjørning.





## Anmærkninger til 2<sup>det</sup> Hæfte.

Nr. 17 har jeg i 3 Opskrifter, haade fra Sidsel Jensdatter, Kirsten Sønderager og Abelone Joh. Kristensdatter. Deres Ord stemmede ikke nær overens og kunde bedst stilles sammen for at fuldstændiggjøre hverandre; dog er der kun taget grumme lidt Hensyn til Abelones Vers, da de vel have anderledes Ord og Udtryk, men i Hovedsagen intet andet Indhold end Sidsels. — Her ere nu Opskrifterne sammenstillede, og Sidsels Vers betegnede med a, Kristens med b. Nr. 18 har en Slutning som Ribolds Vise. Nr. 19 synges til samme Tone som Nr. 4 i 1ste Hæfte. I Nr. 23 har jeg anført Gadens Navn, som det blev udtalt for mig, udenat indlade mig i Gissninger. Ligesaa i Nr. 25, 26 behandler et stedligt Sagn. Det henlægges af Folk paa Eggen til en Gaard i Gjellerup Sogn, der endnu kaldes Fuglsang, og Skytten skal have boet paa Herningsholm. Noget lignende gjælder om Nr. 27; thi en Høj i ovennævnte Sogn paavises som den, hvorved Bondesønnen sov, og Hanen, der vækkede ham skal have været Lundgaards. Af Nr. 30 har jeg udeladt det meste for dets Værdiløshed, og ligesaadan har jeg gjort ved 31, fordi den har været trykt.

Sidsel Jensdatter har sunget følgende for mig: Nr. 17, 13, 19, 22, 23, 24, 26, 27; Kristen Sønderager følgende: Nr. 17, 20, 25; Mette K. Larsdatter Nr. 20; Kristen Jensen i Lund Nr. 21; Maren Pedersdatter Nr. 28; Jens Vonneborg Nr. 29; Peder Andreas Nielsen i Have Nr. 30; Ane Nielsen fra Tørring Nr. 31; Lavst Bolle Nr. 32, 33 og min Fader Nr. 34.

Angaaende Tonerne maa jeg endnu tilføje: Nr. 20 b har Mette K. Larsdatter sunget; Nr. 23 har jeg fra Abelone J. Kristensdatter; 24 fra Jens Vonneborg; 25 a fra Hans Schuster i Lyngby ved Grenaa og 25 b fra Sidsel Jensdatter.

### VI. Religiøse sagn, legender m. m.

lejested, ikke større, end som en lille mand kunde have ligget der. Men folkene mente, at det var Jerusalems skomager, som havde haft husly i gården den nat.

Således er fortalt af Svend Jensens datters søn, som Anders Svendsen, som nu ejer gården.

709. De faldne engle. a. Bjærgmænd, ellevæsen, nisser, trolde, dværge og havfruer og hvad alle disse overnaturlige væsener nu hedder, er blevne til på følgende måde. Da englengene havde syndet, styrtede Gud dem fra Himmelen ned til jorden, og de, der da faldt i moserne, blev til ellefolk; de, der faldt på markerne, blev til nisser og trolde; de, der faldt på de høje steder, blev til bjærgmænd og dværge; de, der faldt i havet, blev til havfolk. For Kristi fødsel regerede disse væsener, men ved kristendommens indførelse begyndte de at tabe deres magt, og efter reformationen er de helt forsvundne.

710. De faldne engle. b. Det var den gang, da Verden herre stødte englengene ned af himlen, da faldt der nogle på bjærgene, og de blev til bjærgfolk, dem, der faldt i moradserne på elle, de blev til ellefolk, dem, der faldt på den flade jord, de blev til underjordiske mennesker, og dem, der faldt på højene, de blev nisser.

711. Hundedyrker. Der sagdes, at Bonaparte var ikke død, som folk skrev i bladene, men han var kommen til nogle mennesker, der kaldtes *Hundedyrker*, og de havde aldrig, men lå i dvale to timer hver middag. Det gjorde derimod Napoleon ikke, og da gik han om imellem dem. Sådan fortalte min bedstemoder om Bonaparte, og hun havde hun al tid megen snak om, men hvor hun havde den videnskab fra, ved jeg ikke.

Skulde dette navn ikke stå i forbindelse med navnet *Hunnere*?

712. En spraglet orm. Det hændtes en gang, at en pige gik fra by og undervejs kom forbi en spraglet orm. Hun løb da hjem og tog en kjæp, og lige så mange uldtråde med de samme farver, som ormen



# OCR Software

- Tesseract (<https://github.com/tesseract-ocr>)
  - Slightly opaque but highly customizable
  - Lots of new language training data
- Abbyy
  - Proprietary
  - Lots of language support
  - Fraktur engine
  - Extremely easy to use



Nikkelsen, Nils Mikael, Sognefoged Lyngby  
 Ad. 4104 b - o o b. (Top. 1249).

f. 22/ 1 1838.

Strandfogedga

( 17/ 2 1844

Via

den overgaar

selv 1925

DS. III 6

6104,

M. O. T.

Friederichsen, Kristense, Plet. (Top. 2433), aflyglskme.

(f. 30/11/1897 x 13932) Shurup, kom som 19-årig til at tjene  
 i gård i Lyngum. G. med Ole Peter Mogensen fra  
 Plet. - Modtil Maren Olesdatter, Lyngum, se denne).  
 of Johanne Marie Olesdatter  
 Kristense (død 1893)

Da. 5708 b - 10a. 2191<sup>2</sup> Alv. 1497; 1504 - ~~08~~, ~~1508~~. 1508<sup>4</sup>

Ref. III 2: 174, 269.

J F. XI (1890), 18a, 20a, 21, 22B, 28a, 40a, 43, 46F, 47.  
 Jd. V 776. 49c, 57, 59B, 64a, 85.

2191, (1890)

† Mo. III 303, 306 J F XI, 342. J F VI 308.



# Challenge: Classification, Search and Navigation



Ewald Tang Kristensen Danske Sagn II

Ewald Tang Kristensen Danske Sagn III

Ewald Tang Kristensen Danske Sagn IV

Ewald Tang Kristensen Danske Sagn V

Ewald Tang Kristensen Danske Sagn VI

Ewald Tang Kristensen Danske Sagn VII

Danske Sagn Ny række

Jyske Almueliv

Danske bonde- og Mark...

**Jyske  
Almueliv**

# Danske Sagn Ny række

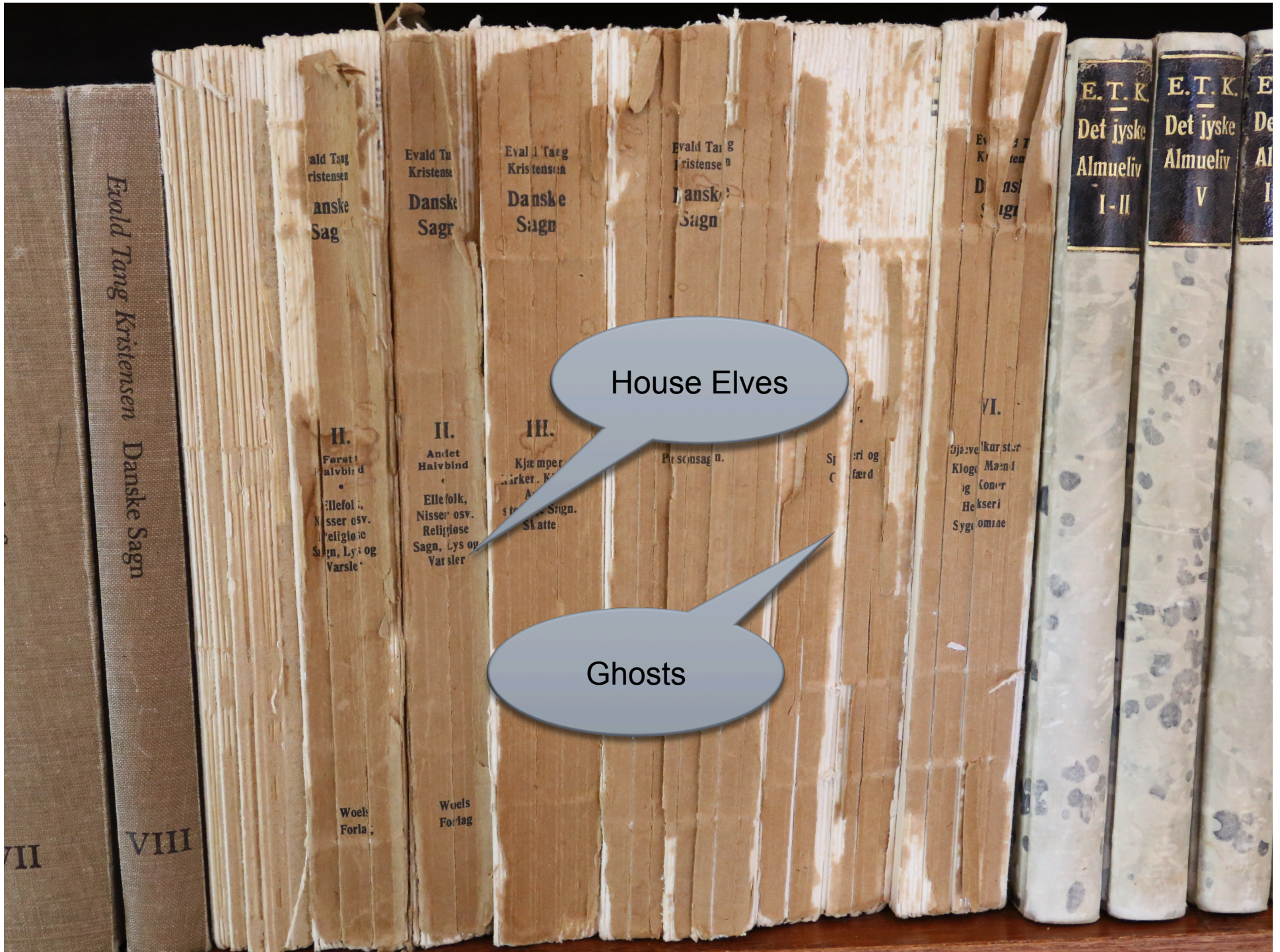




Hidden Folk

Witches and  
their Sport

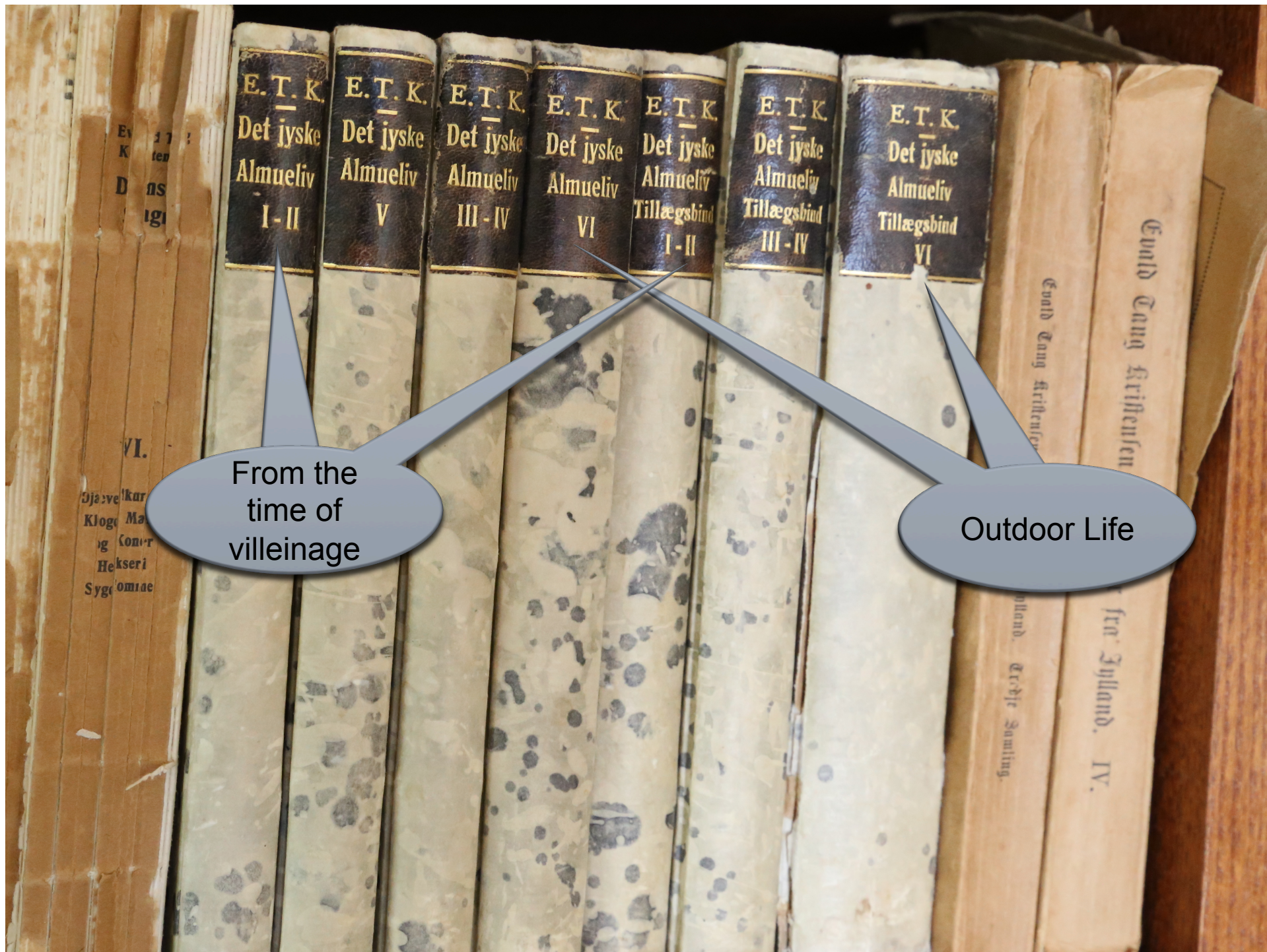




House Elves

Ghosts





From the  
time of  
villeinage

Outdoor Life

# A story...

It was the old counselor from Skaarupgaard who came riding with four headless horses to Todbjærg church. He always drove out of the northern gate, and there by the gate was a stall, they could never keep that stall door closed.

They had a farmhand who closed it once after it had sprung open. But one night, after he'd gone to bed, something came after the farmhand and it lifted his bed straight up to the rafters and crushed him quite hard. Then the farmhand shouted and asked them to stop lifting him up there. "No, you've tormented us, so now you'll die..."

I heard that's how two farmhands were crushed to death. He wanted to close the door and then they never tried to close it again.



# Classification

- Most non-digital collections are constrained by costly, manually applied classification systems that
  - Are usually one text ↔ one classification
  - Predicated on research questions current at the time of the development of the classification system
  - Do not play well with other classification systems
  - Case in point: The story example was classified as one about manor lords!





# Goal

- Our goal is to develop flexible search methods that
  - Allow the researcher to emphasize aspects of the collection related to a research question
  - Allow for search and discovery
    - across the entire corpus
    - outside the corpus



# Information flow and Networks

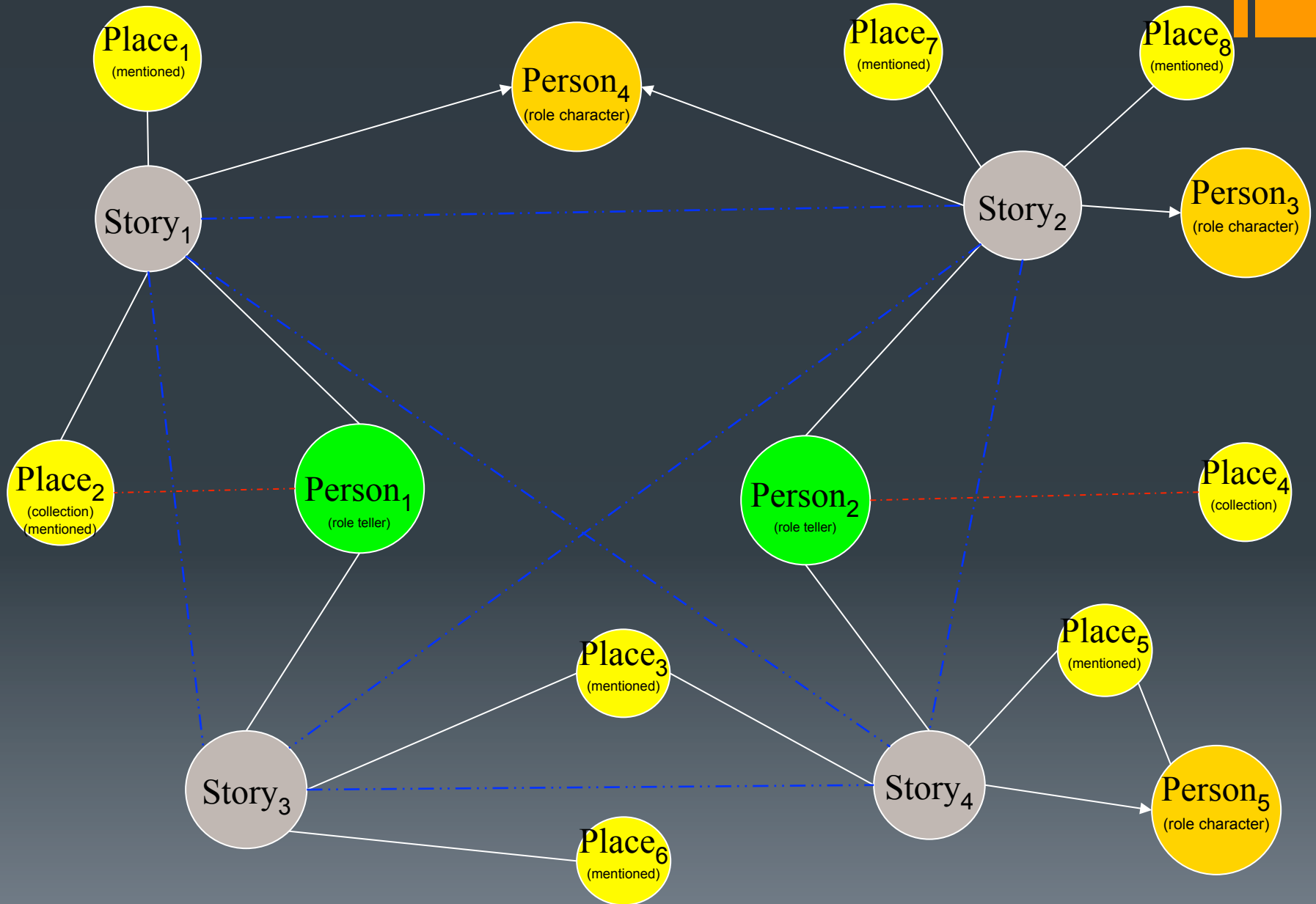
- Information flows very poorly across disconnected resources
- By connecting stories together, information can flow across the resources
- What are the possible connections between stories?
  - Topics (existing and newly derived)
  - Keywords (how does one derive keywords?)
  - Shared features (ontological organization)
  - Storytellers (who told the story?) and collectors (who collected the story?)
  - Named entities (place names and personal names)
  - Dates (dates of collection, dates mentioned)

# The “Story Space Navigator”

- Multi-modal network classifier
  - Modes of the network: People, Places, Story features
- Stories
  - Represented as a weighted feature vector
- Places
  - Can attach features to places (administrative units, political aspects, local church, etc)
- People
  - Can attach features to people (biographical data, education, financial data, work and training, etc)

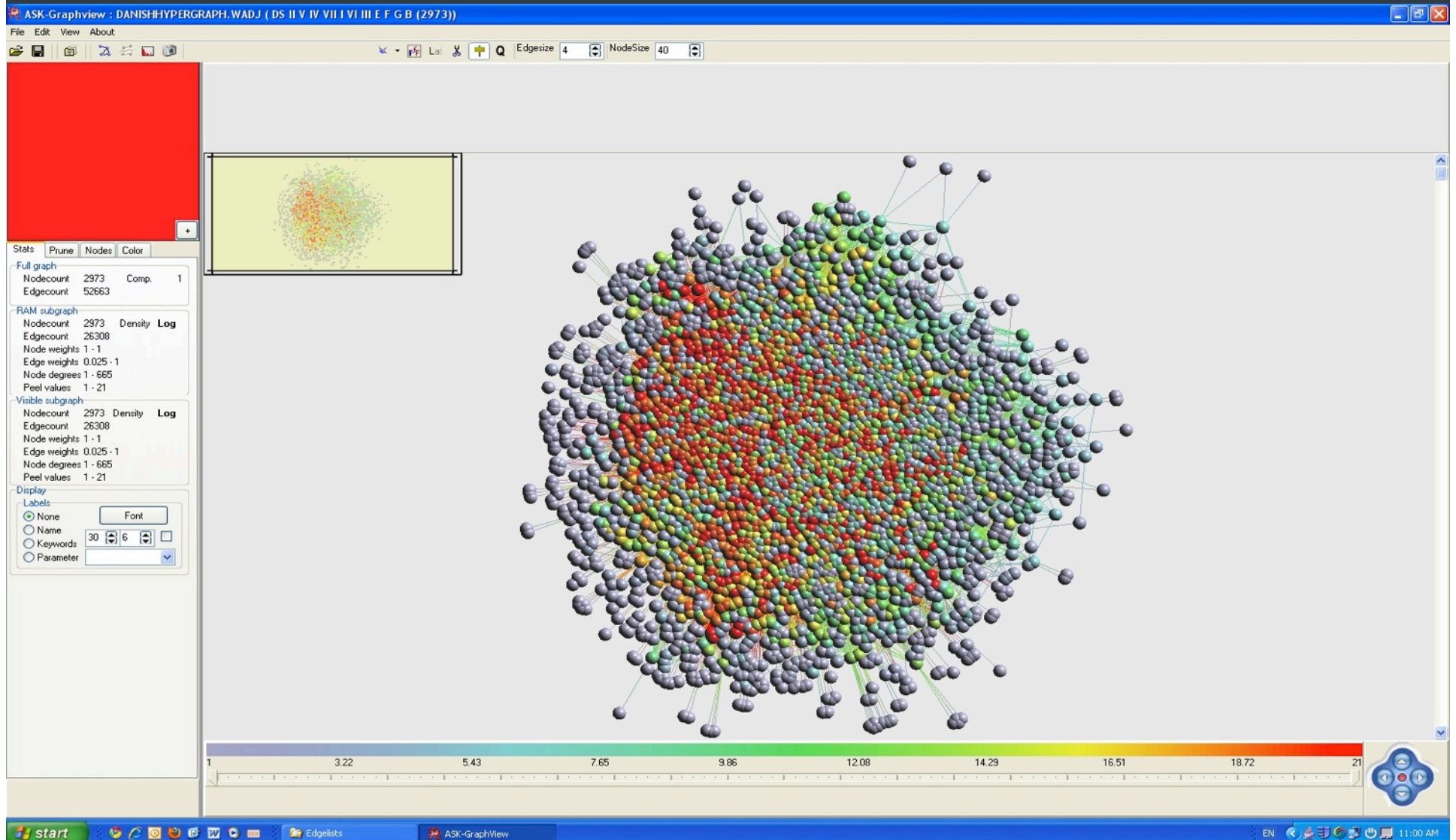
# Connecting the dots...

26

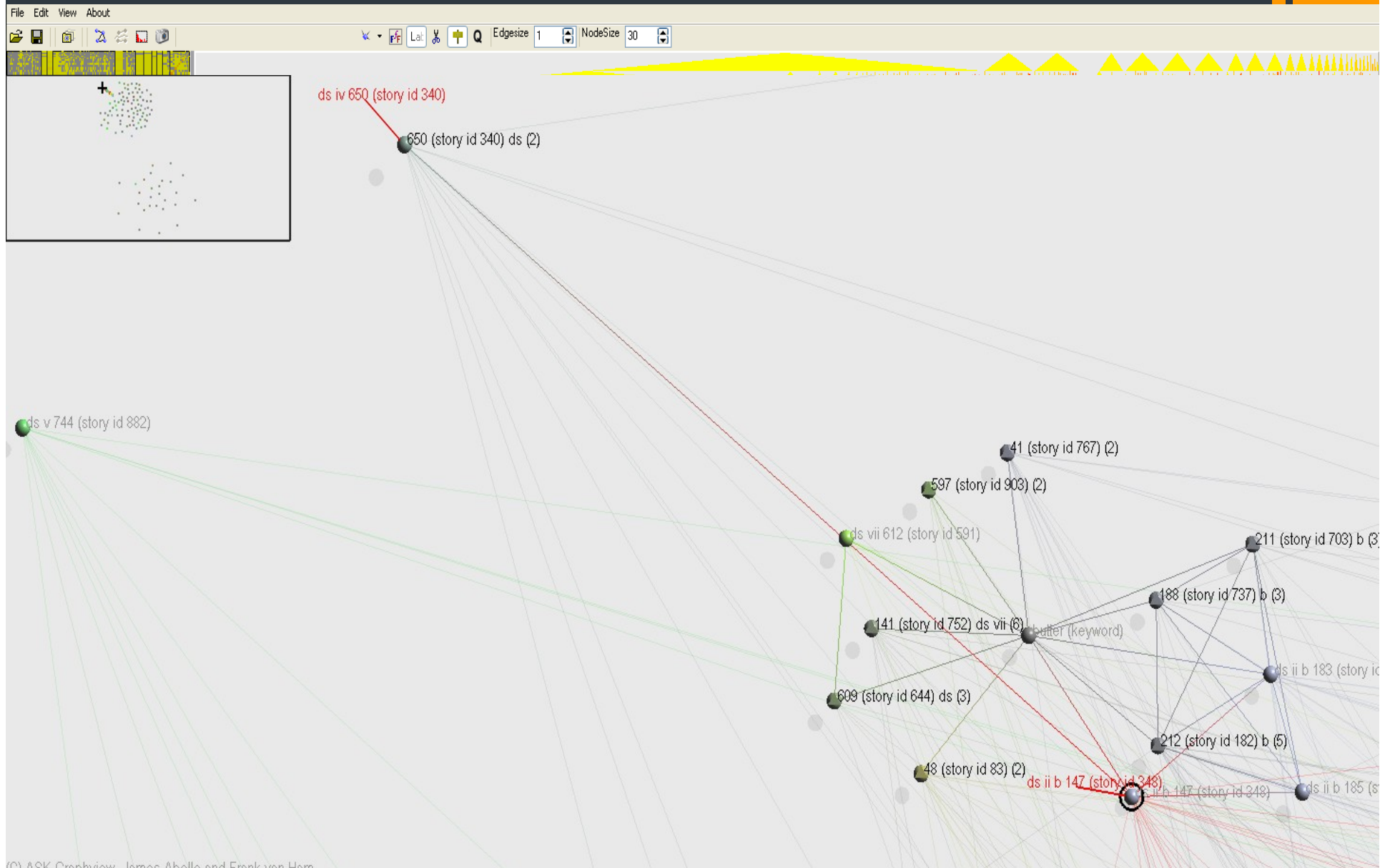




# Story Hairball



# Structure in the network



## Helps us

- Reclassify the story to make it discoverable with other ghost stories
- And discover a story about a homicidal house elf at a farm in Egå...

When they got home, the farmhand was happy because now he'd gotten something to use for fodder, and afterward *nis* could go and feed the animals just as much as he wanted to. Then they got another farmhand, and he didn't want to let *nis* go on like that. But the farmhand got lifted up in his bed and all the way up to the rafters, so he lay there dead when people got up the next morning.

# New research question

- What is the relationship between ghosts and house elves?





## Another approach to classification

- What happens where/when classifiers “fail”?
- What is failure (in the Humanities)?
- Aren't the borderlines some of the most productive regions?
  - Which stories are the disruptors?



# The Tell-Tale Hat

Per Overlade was out one evening shooting hares. It was up on Kræn Møller's field. Kræn was in the process of moving his farm, and the old farm had not been completely disassembled yet, and Per intended to hide amid the old frame that was still standing and shoot a hare or two. But when he gets there, he sees an old man who is sitting in there with a red cap on who nods to him. Per gets scared and doesn't dare go in there, and so he doesn't catch any hares.



# Original Classification

- Story classified as story about hidden folk
- Could also have been classified as a story about:
  - Hunting and life outdoors
  - Law breaking
  - Witches and their Sport
  - House Elves (nisser)



# Why is this story important?

- Commentary on change in agricultural organization in early 19<sup>th</sup> century Denmark
- Commentary on the constraints on hunting
- May help answers about the belief in house elves and stories about changes in rural economic organization
- Nothing about this story suggests that it was properly labeled as a story about hidden folk





# What does this let me do?

- <http://mimno.infosci.cornell.edu/naivebayes/>
- Stories on the margins
  - Are hard to classify stories rhetorically important?
  - Do they address complex and ambiguous issues?
  - How many marginal stories do I find?
  - Are there classes of informants who tend to tell these stories that slip through the original classifier's cracks?
- Reclassifying affords new perspectives on a collection

## Witches classified as Cunning men



hare heks hekse mælken malke dér koen skød kat smør kjærne mælk hekseri kjærnen forhekset kirke smør kors oven forgjort hegse kreaturerne sön låne kniv fortræd kalven bedstemoder tænkte ringe mette mødte stiltiende køer ane kjælling moder magt folks ligesom lig chr skjære broder kastede kirken begravet knap forhegset kreatur kalv bjælken hendes bedre nørre vesten skyld barnet møde datteren uheld hver bæk hoppede brædet tjene sengs skik huu gårdmand fremmed ban sort hende bedste hun bange søgte kirkegården brændt lært sidst ganske loftet karen mund skidt fejl pigen disse altså fløj står gjore søster ende lagt led kun kærne vor sætter råd død sine pige stalden kone mands fået indtil store dage gård kreaturer når omsider dagen måde faldt straks intet tit præsten gammel får aften sprang længe dag huset døren bag svarede gennem stykker gården gamle man för karlen løb gjort deres dette hus rigtig syg endnu under eller går over hed konen alle tre blev efter den sig var med dem kommer nok gang fra lidt kan hen gjøre manden inden hos ude kommen mig hans faa fader stykke ild siger dig hjemme sat præst per han tilbage mine igjen natten par netop nede ord sidder nej derfra saa paa ham imens dermed hesten vendte kjørte slemme niels brændte sorte andersen læse gaa gaaet ligge viste hestene skaffe jeg kurere ogsaa maa mærke videre sager meu saadan blevne mænd skoven jakobsen bestemt finde stedet stillede midt hau blevet henne land marked arm skilling maatte peder smed hvide altsaa strags ilden vinduet allerede stue bedstefader betale vognen vejle mistænkt øster vand iver Jens vorherre temmelig rejst papir mælholm forstaa pastor kunnet mørkt krog gaardmand samlede kornet levende sandt alen taget kjed bede cyprianus klæderne fisk smeden førte ridende kristian lomme vendt balle kolding klogere porten kammer timer bygget raabte brændevin djævelen hentet storm tilfældet bande banke samlet betalt straf salmebog befalede helbrede times ring binde stjal siges svin hals søgt saks heste bæster neden slaaj jørgen brug krybe sold benet penge besked pind lavst veje tåning kjørt bedsted lindeballe øjnene ravn sillesthoved gjør søren beholde skåret plove fingre kureret læste garn staa kjørende posen daler løse flyve gaug håning helnæs dermed stiller frø kro thomas knude flyver holstebro jokum kroen bøger nar øje smidstrup herredsfogden staa stille pose laden bæsterne standsede vinden vise gadbjærg pengene stjæle løste træet mus standse nordmanden nordmand tyven stjålet kusk

27. Der var to brødre i Saltum, den ene hed Døve-Lars, og den anden Døve-Ole, og de var kloge begge to og kunde se det, andre kunde ikke se. Kristen Nielsen, Mosehuse.

32. Enevold Hvid kom ved en stor del af hans klogskab på den måde, at han satte alle bogstaver op, ligesom når vi sætter tal op og regner med, og der var dem der havde set, at han havde alle 23 bogstaver stillet op for sig, og han multiplicerede og adderede med dem, og deraf regnede han sig til hans videnskab og spådom. J. B. og M. H., Lille-Tåning.

76. Min første mand han var en dag ude i marken at plove, og han havde stude for. Så tykte han, der gik så meget en grov stor lus på den ene, og han vilde da prøve at prikke den ihjel med et halmstrå, men det kunde han ikke. Så tog han noget jærntøj, men a kan ikke huske, enten det var hans træsko eller noget af ploven, og slog den ihjel med. Ved han fik den slået ihjel, så gav han sig til at plove igjen, men tog ikke uden en ornfer, så blev han syg og kunde ikke blive ved længere. Nu måtte han så spæn- de fra og gå hjem, men kunde snart ikke magte at flytte benene og måtte krybe meget af vejen. Lav han kom hjem, sagde han til konen, at hun skulde gå over til deres nabo, det var én, de kaldte Søren Krænsen, og han var hend farbroder. Så kommer han jo om og skal gjøre noget ved Jens, for han var noget klog, og så siger han: »Du skulde ikke have gjort sådan ved den lus, du skulde have skåret den i fire parter med din kniv, så skulde du have set spil.« Så véd a ikke mere om det, for Jens han kerte sig, men Søren Krænsen sagde, at hvis han ikke snart havde fået råd, så havde han død. Det var i denne her gård, det tildrog sig. Karen Frederiksdatter.

156. Til Præjst-Jens i Gjødvad kom en gang en mand, som boede der i nærheden, og klagede sin nød, et skarns menneske havde om natten overskåret hans plovås, (den gang var en plov for største delen af træ), og han bad ham nu om at sige ham, hvem gjerningsmanden var. Præjst-Jens var kjendt i mange miles om- kreds for at være en klog mand, der havde gået igjennem den sorte skole og havde Cyprianus. Han lovede at vise ham gjerningsmand- en, og da det var blevet aften, gik de begge to ud til ploven. Præjst-Jens gjorde med en stok en rundkreds i jorden omkring den, her skulde manden stå og måtte ikke gå over kredsen, inden Præjst-Jens kom tilbage. Da han var gået, begyndte noget som gloende lyngkuipper at dandse rundt omkring. Sraks så han en mand komme løbende alt hvad han kunde, kun iført det halve klæder, og tilstod, han var gjerningsmanden, han vilde til at gå i sin seng og havde lige fået sin ene strømpe af, da han styrtede ud af døren og ned til ploven. — Præjst-Jens var en berømt kvaksal- ver, kjendte mange lægedomsurter, som han brugte til at helbrede syge kreaturer med, og alle havde den største respekt for hans kunster. J. Jensen, Refshalegård.

158. Peder Ginderup havde en stud, der var syg, og det kom da af, at den var forgjort. Så var Jokum i Skals i det samme kommen hen til Per Ginderup, og så siger han: »A har en stud, der har wol i øjnene, kan du trække mig det ud?« Ja, såkorn de jo ud til studen. »Håå«, siger Jokum, »kan du ikke se, hvad han fejler, har du noget i din flaske?« Om den manglede noget for det. Ja, det var da tydeligt. Så gav de ham noget af vandet. Så siger Per Ginderup: »Kan du gjenne skademaunden til at komme?« Nej, det hverken kunde han eller vilde han have med at gj^re. »Nå, ja så kan a selv.« Det var nok i høslætten. Så var der en maud, der gik og slog hø, han smed hans hølé og rendte så stærkt han kunde efter Lvvel bro til, og kom derind, te tungen hængte ud af halsen af ham, så forpust var han. Så vilde Per Ginderup jo have at vide, hvad han skulde. Han foregav, at han vilde snak- ke med ham og lod, som han kom i al skikkelighed, og der var ingen ting i vejen. Så gav han ham en dram og skjældte ham halsen fuld og bad ham om at holde næsen ved sig selv for efter- tiden. Nu kunde han gå med det. Den mand, der havde forgjort studen, var fra Pederstrup udflyttere. Lavat Jensen, Bjærggrav.



# Impact on Folklore

- Archival collections are rapidly moving into the digital realm requiring scalable methods for indexing and classification
- Research questions no longer align with pre-existing classification schemes
- Pattern discovery across tens of thousands of stories may help us discover trends that were otherwise hard to discern



# Impact on Humanities

- Cultural expressive forms are reemerging as a powerful lens onto histories, both local and global
- Rapid methods for complex, multi-membership classification can help with
  - The discovery of research resources in a target corpus
  - The analysis of latent cultural patterns in large corpora

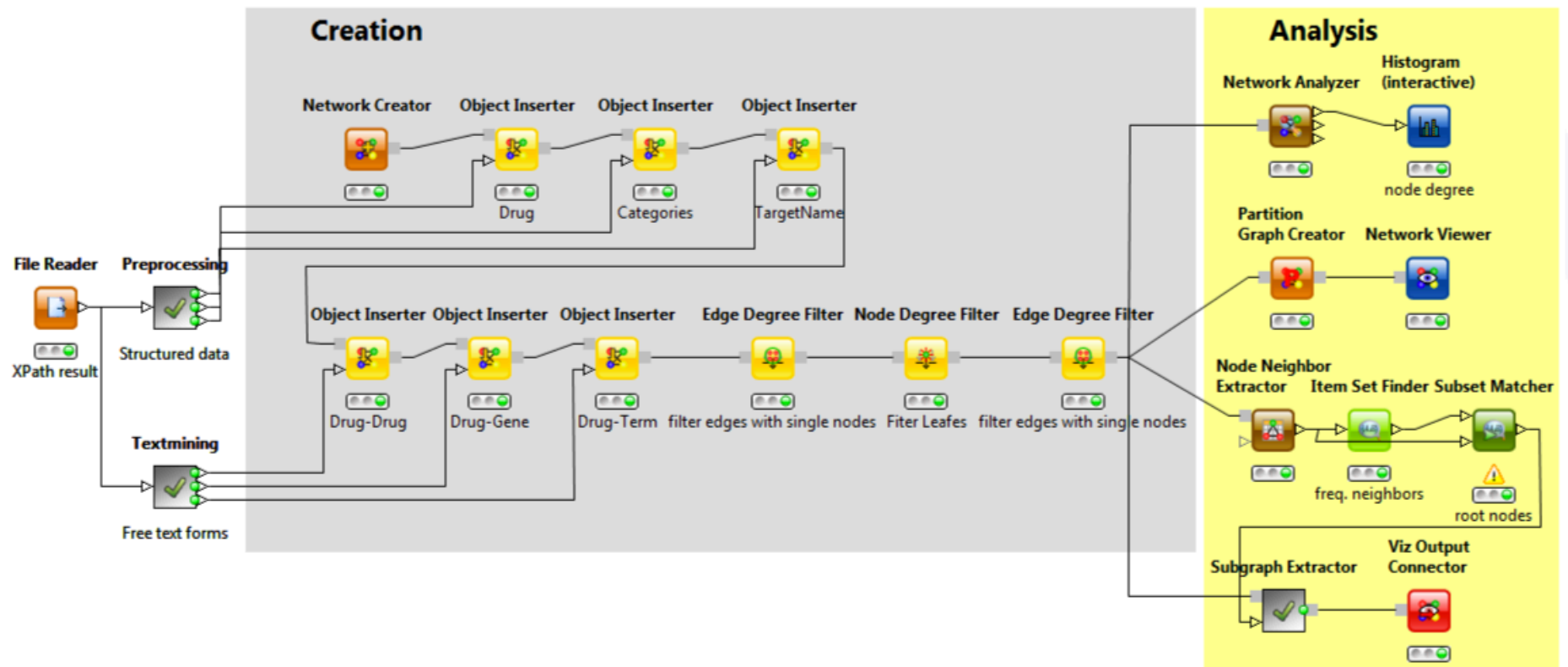


# Machine learning environments

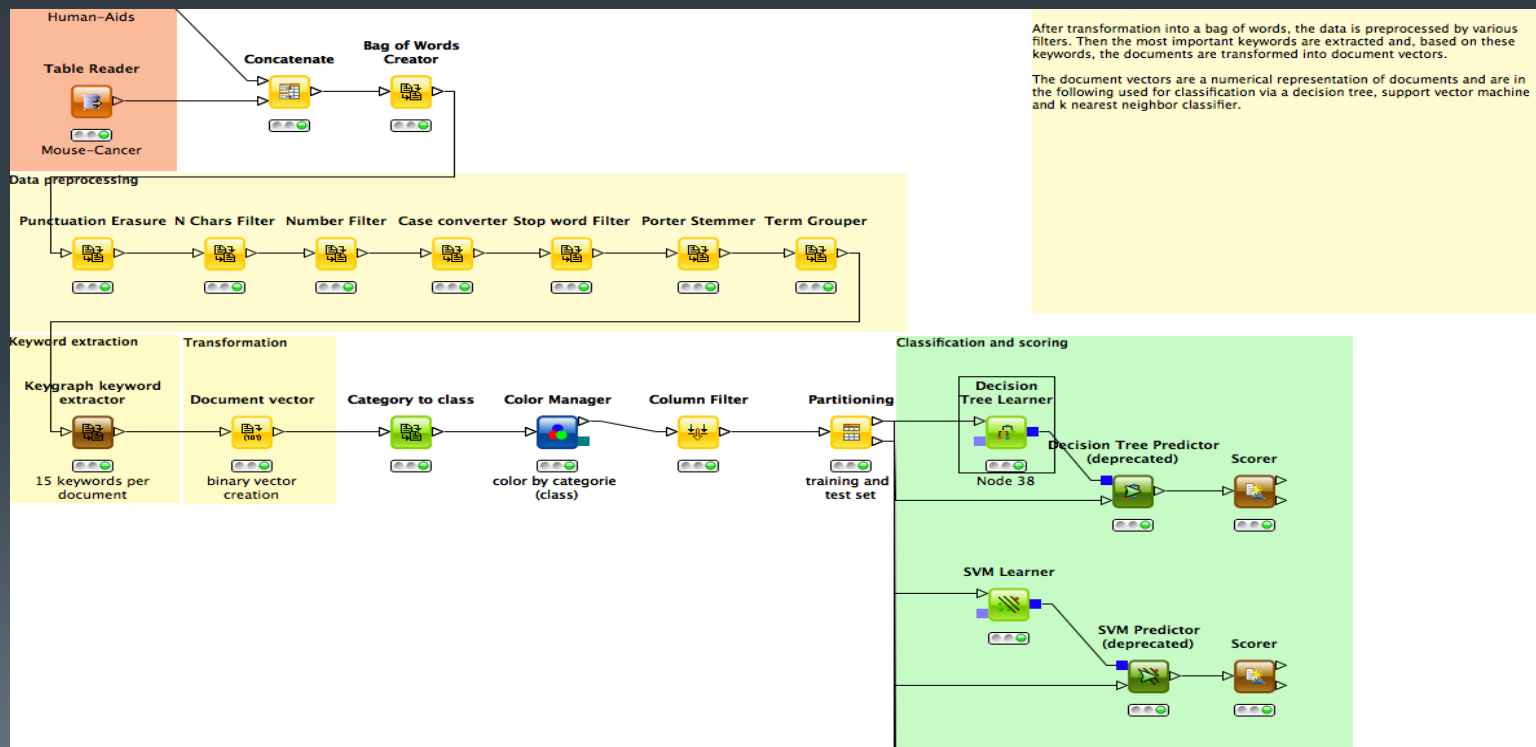
- R / R-studio
  - Open source
  - Many many libraries / new code all the time
  - R-studio provides nice interface
- Matlab
- Knime
  - Open source
  - Very easy to use
  - R extensions
  - Weka learners
- RapidMiner
  - Similar to knime
  - Proprietary / Academic license available
  - Multiple modes for use
- Gensim
  - Open source
  - Very straight forward for topic modeling

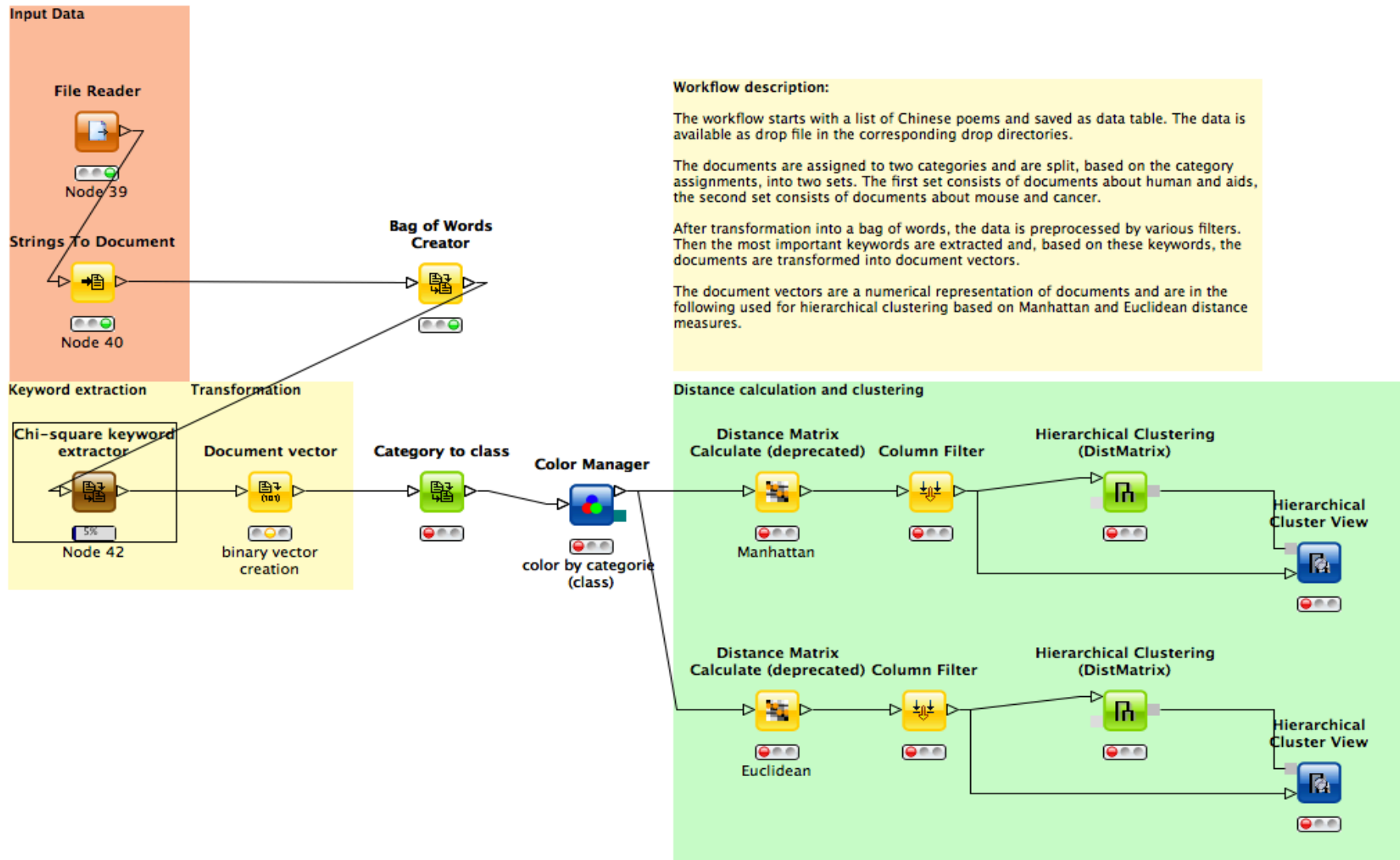
# Sample workflows

- Network creation and analysis



# Supervised and Unsupervised Learning

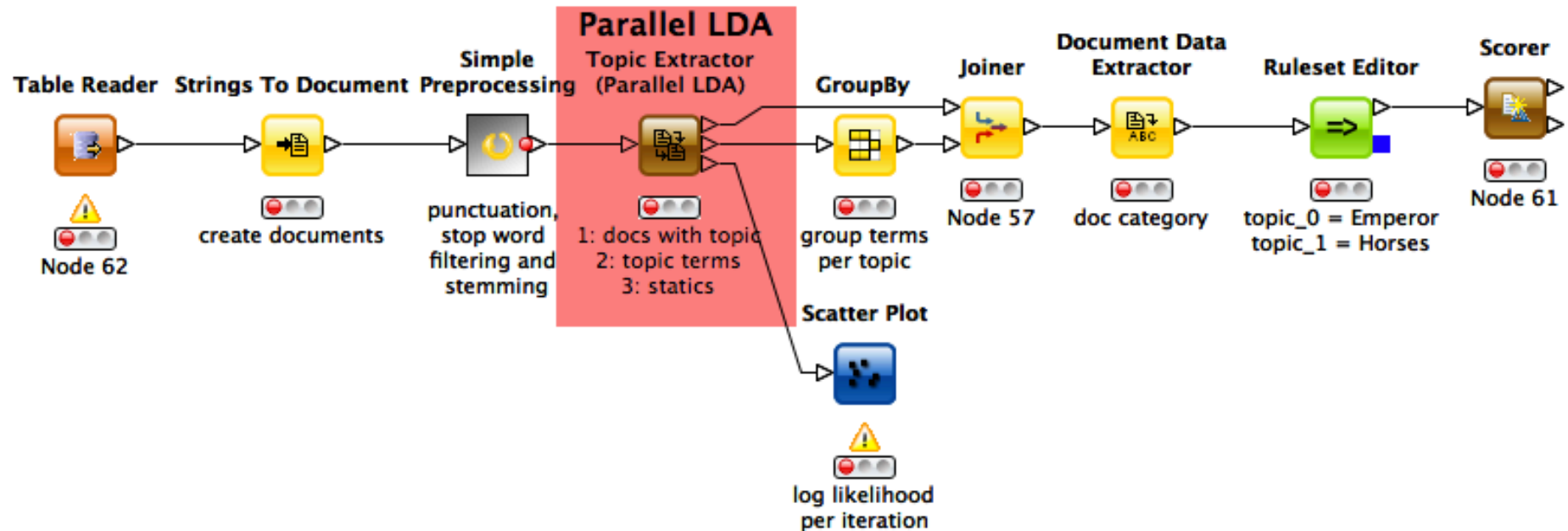






# Topic Modeling (via Mallet)

Topic Modeling using Parallel LDA





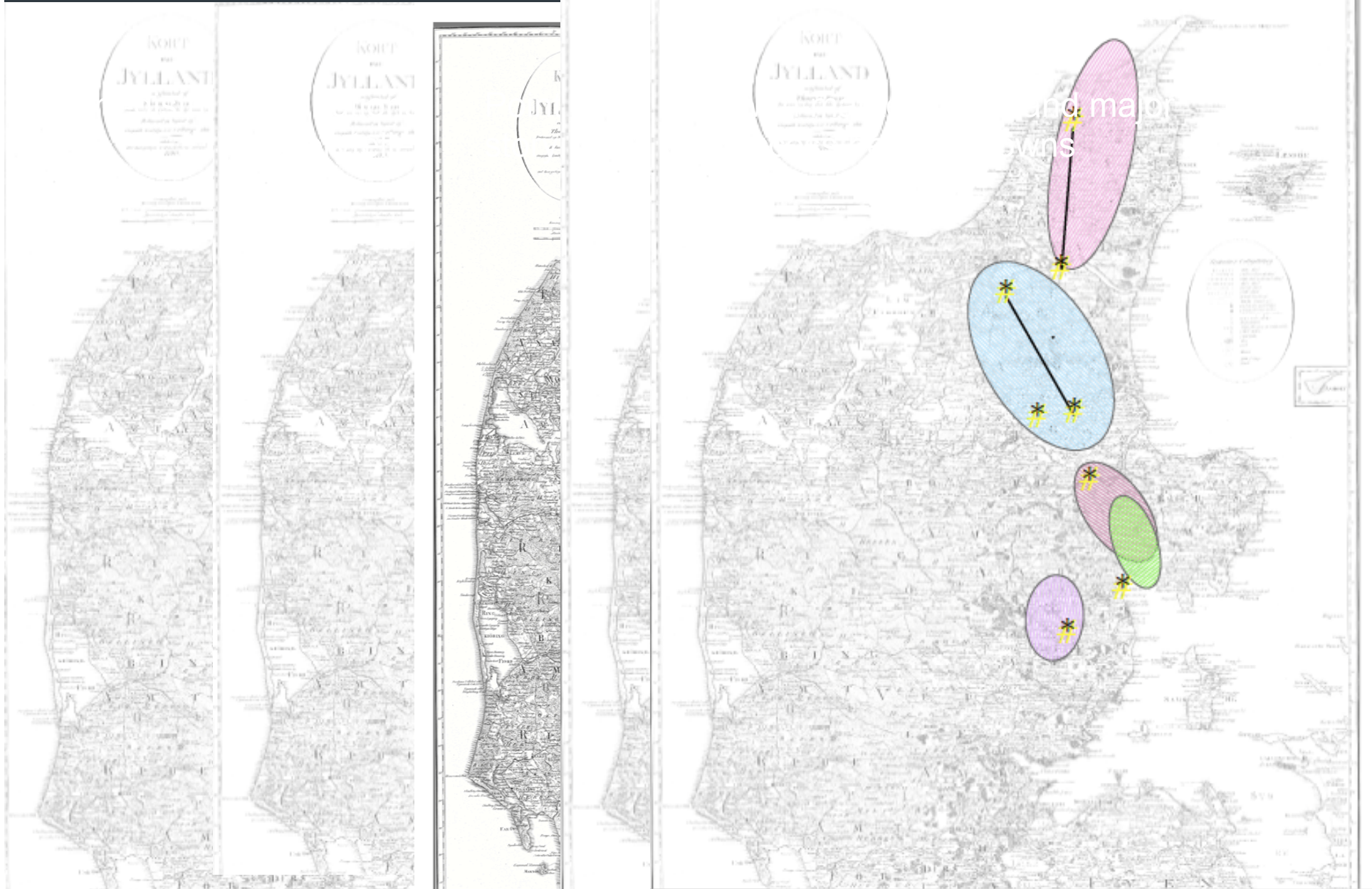
# Geographic Pattern Discovery

- Recognize that stories are often about things that happen embedded in time and space
- Comment on our relationship to places
- Can one use this new model of the collection to understand how people use stories to comment about their interactions in space?



# Storytellers and their Stories

- What was the geographic “narrative reach” of 19<sup>th</sup> rural Jutlandic storytellers?
  - Break out place name distribution by storyteller
  - Standard Deviation Ellipse Distribution
    - Discover patterns for places mentioned in storyteller repertoires
  - Discerning patterns in the underlying place name distribution





# Geographic Information Systems (GIS)

- ArcGIS
  - Proprietary
  - Industry standard
  - Cryptic but preposterously powerful
- QGIS
  - Open source
  - Huge community
  - Based on ArcGIS ideas
  - Very easy install / learning curve is unusual

# TROLLFINDER



Geo-SEMANTIC VISUALIZATION OF  
SUPERNATURAL OCCURRENCES





# TrollFinder: Geo-semantic Folklore Browsing

- Can we develop a method that allows us to find story topics by geographic area?
- Where are the areas that, even after correcting for population density, seem to be highly related to a particular topic?
- Conversely, can we develop a method that allows us to search for topics given an area in which we're interested?
- Can we develop a system that allows a researcher to decide what type of searching they want to do, be it based on keywords, topics, spatial criteria or a combination of these and other approaches?



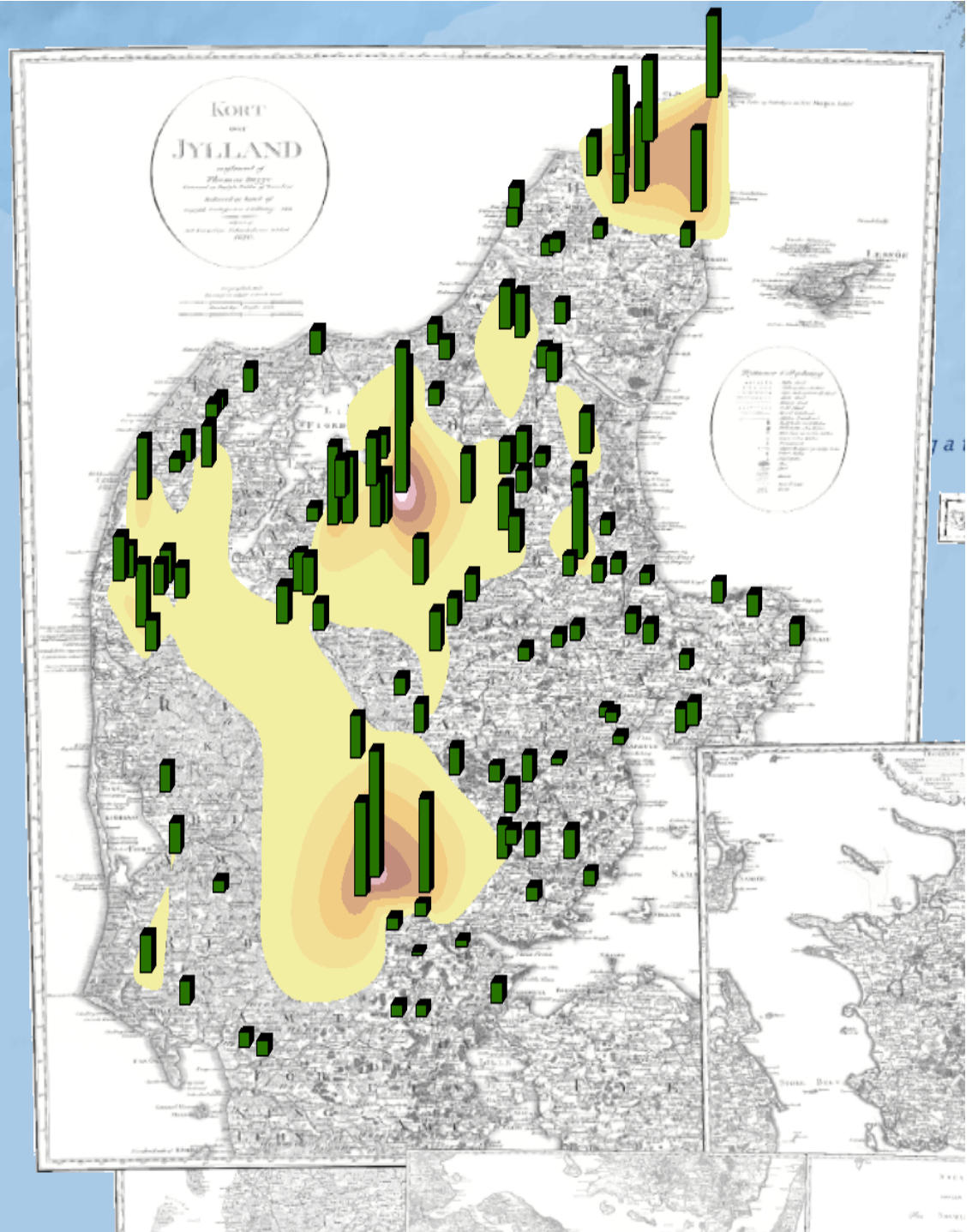
# Sample questions

- Were witches (trolls) closely associated in the legendary material with a specific area or areas?
- If so, are there reasons why they were witches so commonly associated with a particular area?
- Or:
- Is there a close spatial relationship between mentions of cunning folk and ministers?
- If there is, what may lie behind this close relationship?

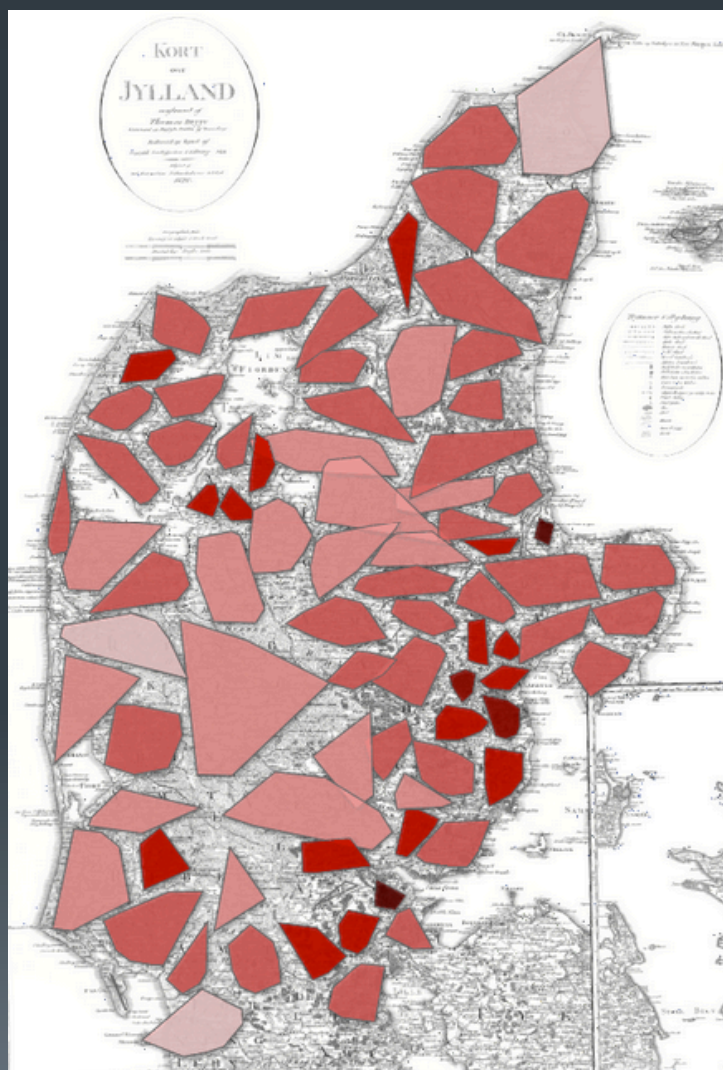
# TrollFinder 1.0

(CMN/LREC 2012)

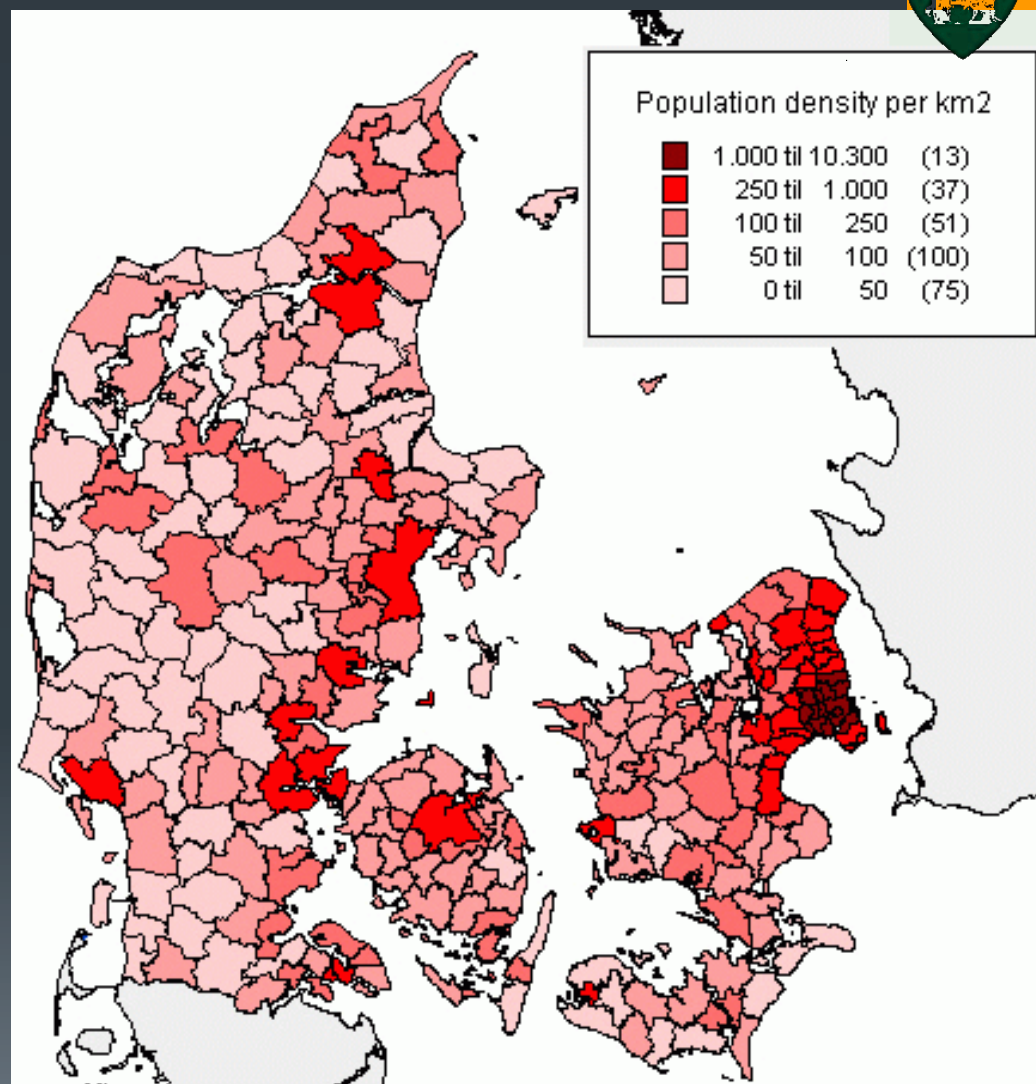
Heat map for the topic  
“Heks” [witch] with story  
frequency histograms



# Normalizing by population density

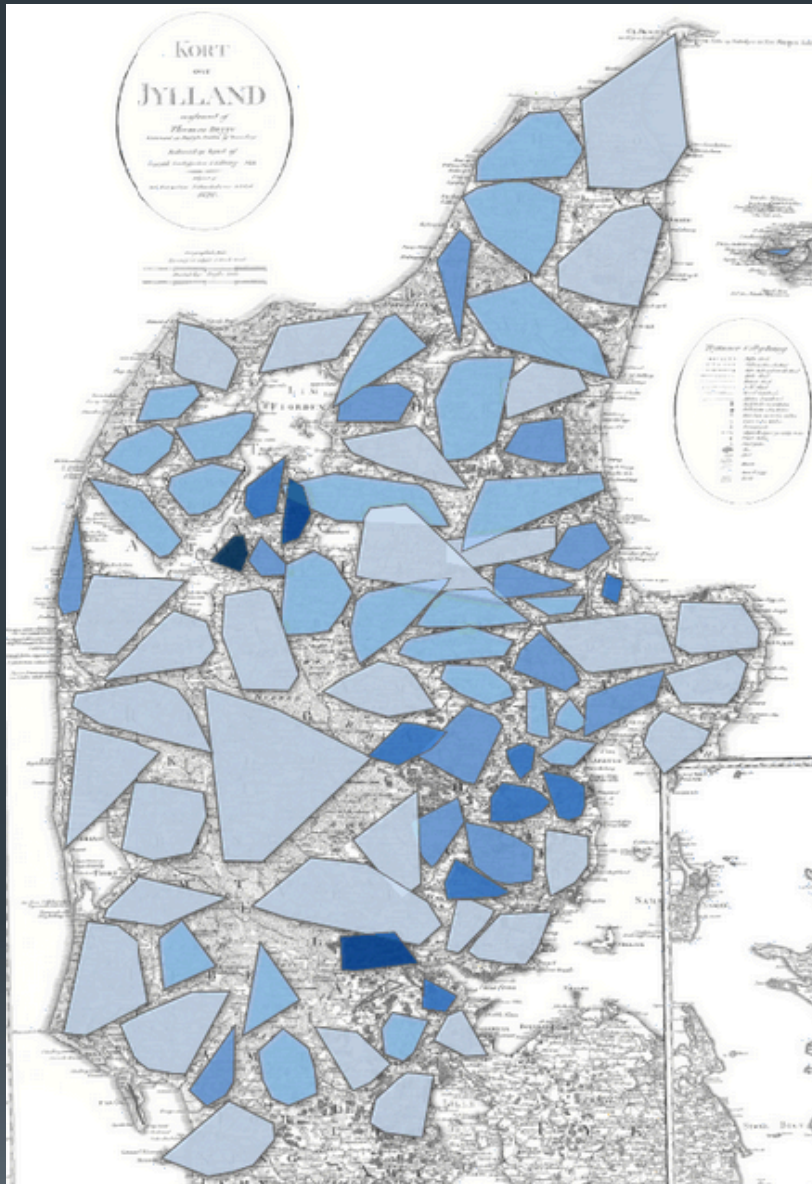


Population density based on 1901 census

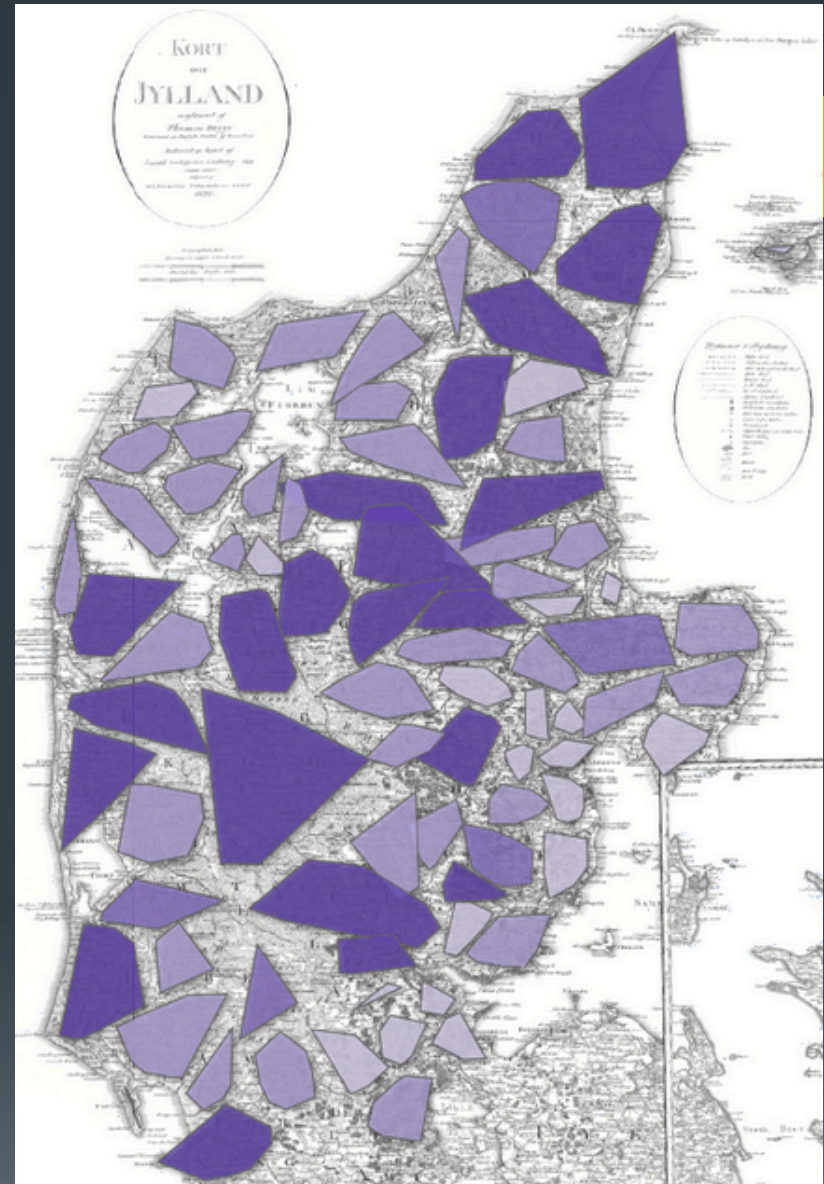


Population density circa 2000





**Story density:** story-to-place mentions per km<sup>2</sup>

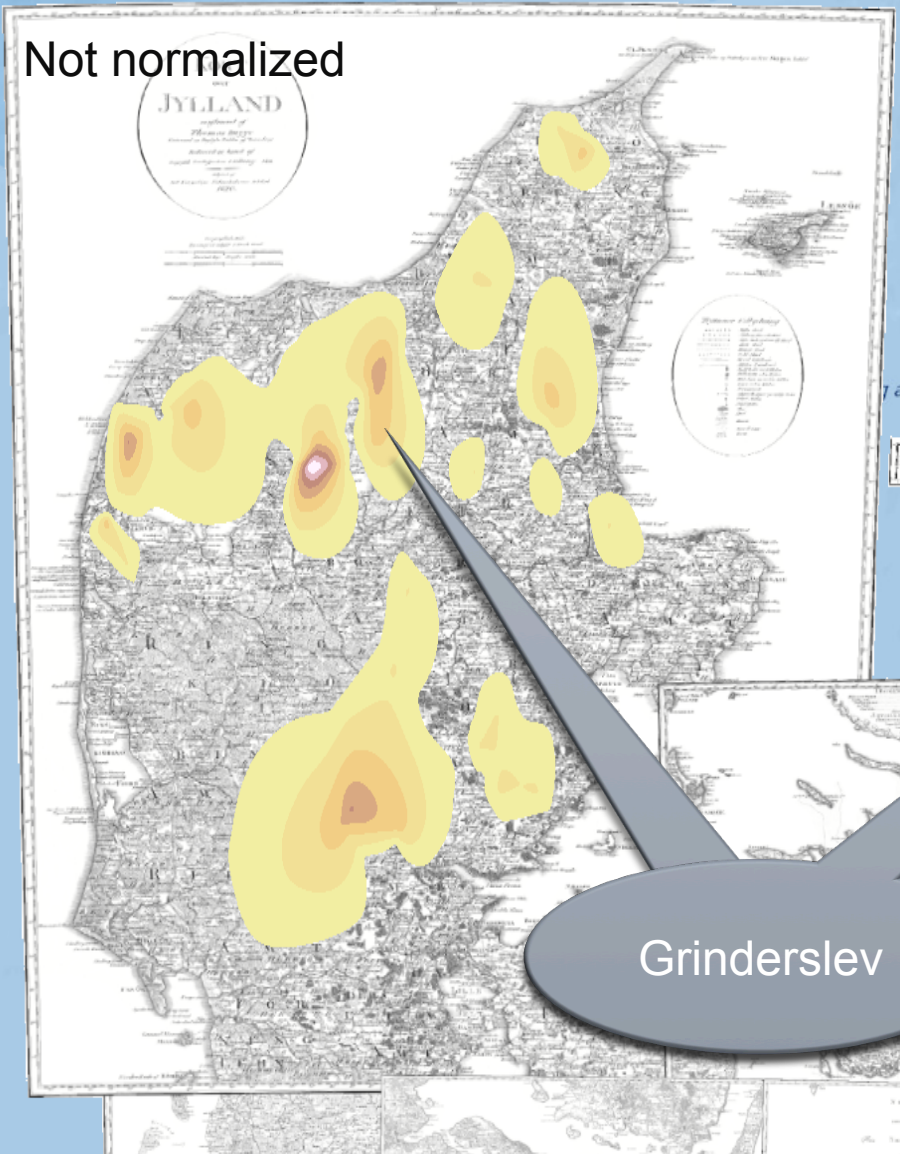


**Normalized story density:** places mentioned/person/km<sup>2</sup>

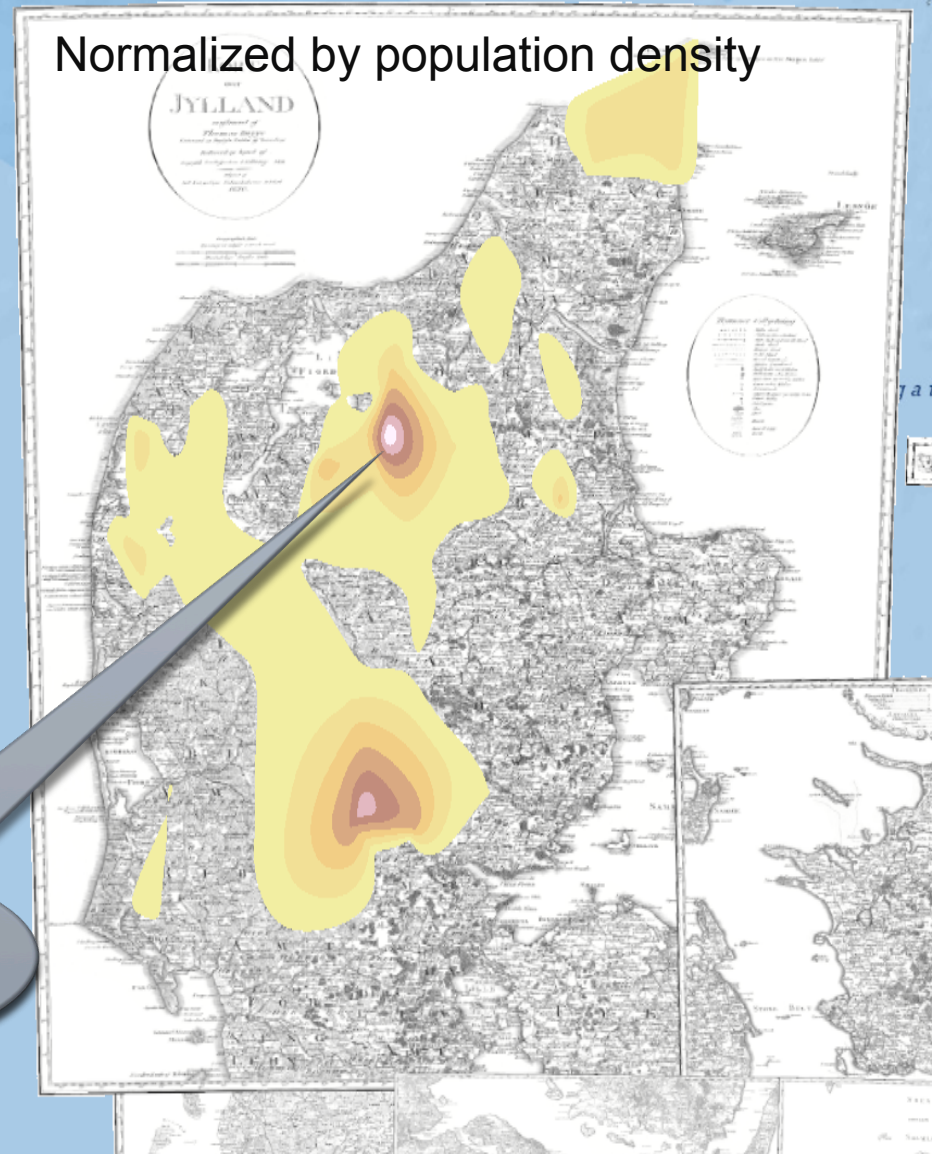
# Effects of population density-based normalization for keyword “witch”



Not normalized



Normalized by population density



Grinderslev



# Grinderslev? Really?



# From the Danish Folklore Nexus...



Browser window: Danish Folklore: The Evald Tang Kristensen Collection

URL: <http://projects.cdh.ucla.edu/danishfolklore/index.html>

Search: dannet

Navigation: HOME | Fieldtrip 6 | Search Stories | Story 88 | Story 177 | Place 947 | Story 889

PLACE DETAILS for DS\_VII\_304 told by Mads Christensen

STORY DETAILS for DS\_VII\_304 told by Mads Christensen

Map: Modern Aerial | Modern Street | Historical: ☐

50 km | 50 mi

Story data for DS\_VII\_304

- Order told: 0.00
- Recorded during fieldtrip: None recorded
- Place recorded: **Vester Grønning**
- Field diary page: N/A
- Associated keywords: cat, church, dancing, ear, east, horse, monastery, old, riding, road, spring, St Hans,
- Places mentioned in story: Troms kirke, Grinderslevkloster

DS\_VII\_304

INFORMANT BIOGRAPHY

PLACES

STORIES TOLD by INFORMANT

570 - 0.00 - DS\_VII\_304

English translation of published version: DS\_VII\_304 told by Mads Christensen

The witches are supposed to gather on St. Hans eve at Breum Spring east of the Grinderslev monastery church by the road which goes over to Langesgård. That was before they'd go off to Troms church. They were supposed to dance there and there was one who sat and played on an old horse's head. When they had good cats, they could also ride on them over there. Because of that, everybody in the neighborhood always marked their cats in their ears, because when the cats were marked, then the witches couldn't take them.

FIELDTRIPS

Actions or events: Witchcraft

MANUSCRIPT VERSION / PUBLISHED VERSION / SCHOLARLY ANNOTATION



# Why Grinderslev??

- Site of a well-known Augustinian monastery, *Grinderslev kloster*, founded in the twelfth century.
- The monastery was built near a holy spring, *Breum kilde*, but was abandoned in the aftermath of the Reformation.
- The spring at Breum was subsequently associated with witchcraft
- In 1686, Anne Madsdatter and her sister were burned at Breum, the last witch burning in Denmark (Bruun, 1920).
- Although this episode is well known in the study of Danish witchcraft, the persistent relationship between the area surrounding Grinderslev and stories about witchcraft has not been recognized previously, suggesting a topic for further, in-depth inquiry.
- Only a few of the stories mention Grinderslev (but rather places near Grinderslev such as Breum)



# Trollfinder 2.0 (WitchHunter)

## Topics and Heatmaps

- Challenge 1: Given a set of labels, can we develop geographic representations of the concentration of those labels (topics) in a given area, and project multiple classifications into a single map?
- Challenge 2: Given a large number of stories that include information on places collected and places mentioned, can we devise metrics for understanding the “conceptual distance” between the “teller” and the “told”?





## Multiple Topics and Heatmaps

- Develop a system for rapidly displaying stories and the places mentioned in those stories
- Allow for multiple layers for human interactive time exploration
- Allow for drill-down to access underlying stories

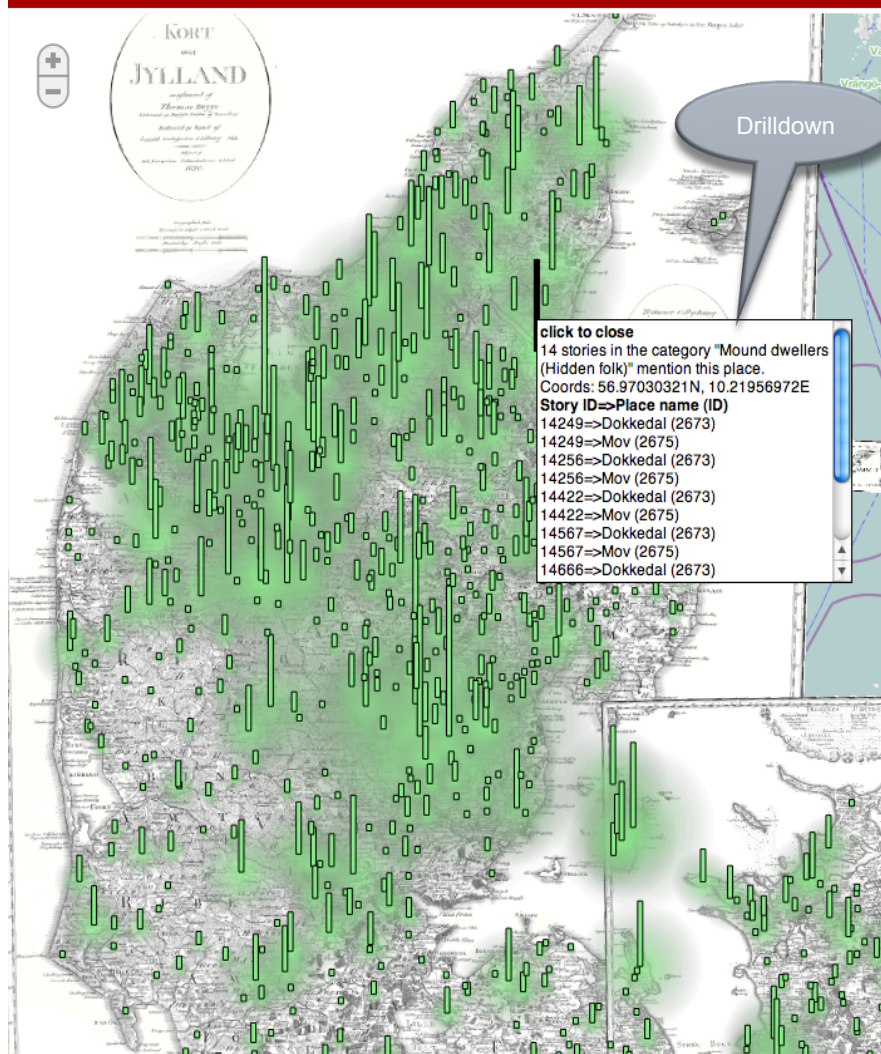
# Hidden folk distribution



## WitchHunter

Mapping the Evald Tang Kristensen Collection

Map data CC-BY-SA [OpenStreetMap](#), Fugro Aerial Mapping A/S



- ☒ Historical maps
- ☐ Aerial images
- ☒ Display place/topic co-occurrence markers
- ☒ Display place collected->mentioned vectors (↑)
- ☒ Display visualizations of spatial statistics
- ☒ Display stories related to map objects

### Story topic indices:

- ☒ Mound dwellers (Hidden folk)
- ☐ Elves
- ☐ Household spirits
- ☐ Traveling monsters
- ☐ Water spirits
- ☐ Wiverns and small creepy-crawlies
- ☐ Werewolves and nightmares
- ☐ Religious legends
- ☐ Death portents
- ☐ Lights and portents
- ☐ Heroes and their sport
- ☐ Churches and their construction. Monasteries, holy springs, etc.
- ☐ Legends about farms and towns
- ☐ Diverse place legends
- ☐ Legends about treasure
- ☐ Small kings and their feuds. Kings. Enemy invasions
- ☐ Manor lords, ladies and mistresses
- ☐ Ministers
- ☐ Diverse people
- ☐ Robbers, murderers and thieves
- ☐ Strandings
- ☐ Plague and illnesses
- ☐ Diverse hauntings and conjuring of revenants
- ☐ Female revenants
- ☐ Revenants on ill-gotten land
- ☐ Revenants in diverse places
- ☐ About the Devil and being in league with him
- ☐ Cunning men and women and their activities
- ☐ Witches and their sport
- ☐ Human and cattle illnesses
- ☐ Concerning earlier agricultural practices
- ☐ From the time of villeinage
- ☐ Houses and life in them
- ☐ Social gatherings and parties
- ☐ Outdoor life
- ☐ Our forebears' way of thinking and spiritual life

# Ministers and Cunning Folk



## WitchHunter

Mapping the Evald Tang Kristensen Collection

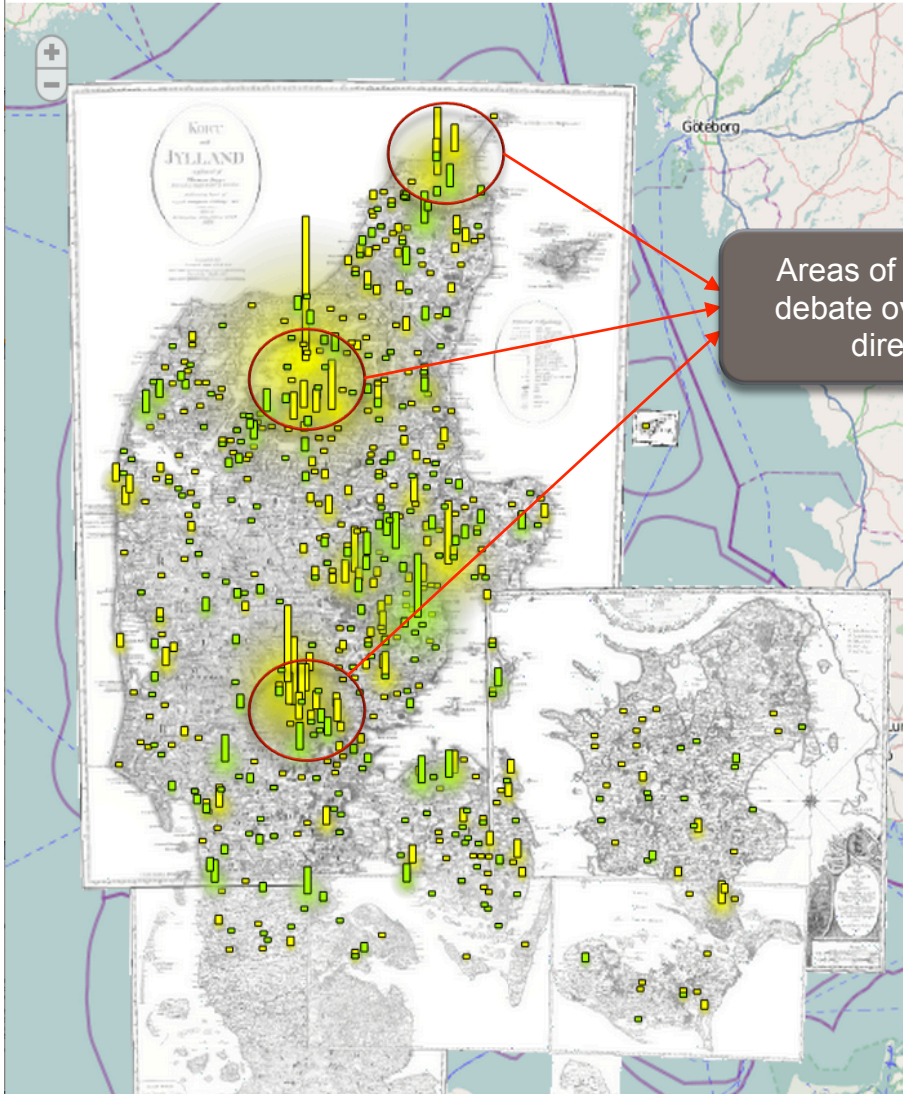
Map data CC-BY-SA [OpenStreetMap](#), Fugro Aerial Mapping A/S

- ☒ Historical maps ☐ Aerial images
- ☒ Display place/topic co-occurrence markers
- ☐ Display place collected->mentioned vectors (↑)
- ☒ Display visualizations of spatial statistics
- ☒ Display stories related to map objects

### Story topic indices:

- ☐ Mound dwellers (Hidden folk) ■
- ☐ Elves ■
- ☐ Household spirits ■
- ☐ Creepy-crawlies ■
- ☐ Nightmares ■
- ☐ Lights and portents ■
- ☐ Heroes and their sport ■
- ☐ Churches and their construction. Monasteries, holy springs, etc. ■
- ☐ Legends about farms and towns ■
- ☐ Diverse place legends ■
- ☐ Legends about treasure ■
- ☐ Small kings and their feuds. Kings. Enemy invasions ■
- ☐ Manor lords, ladies and mistresses ■
- ☒ Ministers ■
- ☐ Diverse people ■
- ☐ Robbers, murderers and thieves ■
- ☐ Strandings ■
- ☐ Plague and illnesses ■
- ☐ Diverse hauntings and conjuring of revenants ■
- ☐ Female revenants ■
- ☐ Revenants on ill-gotten land ■
- ☐ Revenants in diverse places ■
- ☐ About the Devil and being in league with him ■
- ☒ Cunning men and women and their activities ■
- ☐ Witches and their sport ■
- ☐ Human and cattle illnesses ■
- ☐ Concerning earlier agricultural practices ■
- ☐ From the time of villeinage ■
- ☐ Houses and life in them ■
- ☐ Social gatherings and parties ■
- ☐ Outdoor life ■
- ☐ Our forebears' way of thinking and spiritual life ■

Areas of significant  
debate over church  
direction





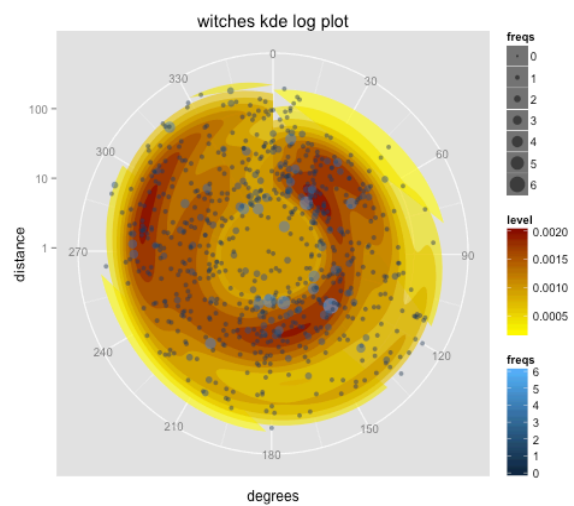
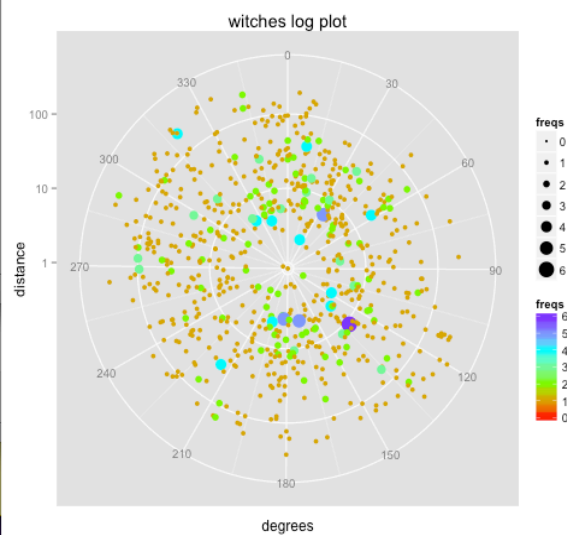
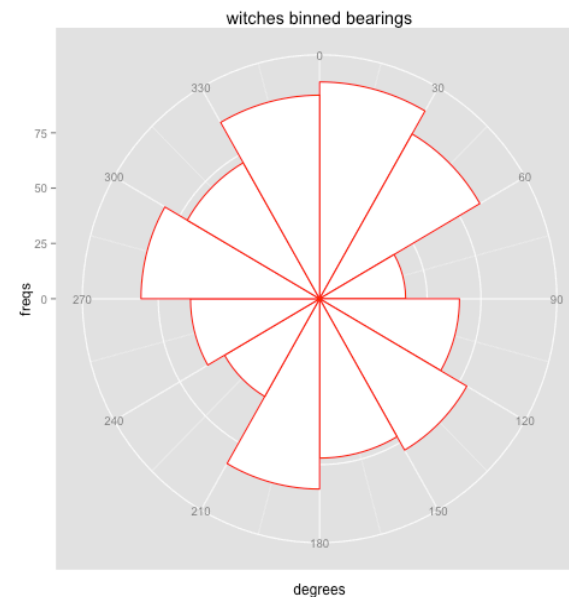
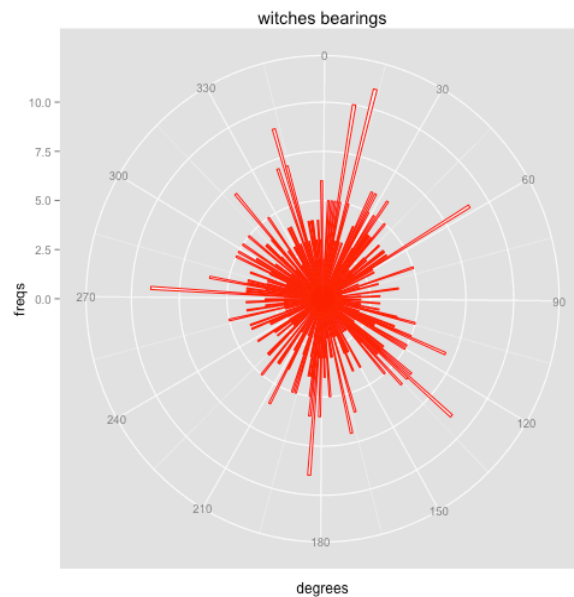
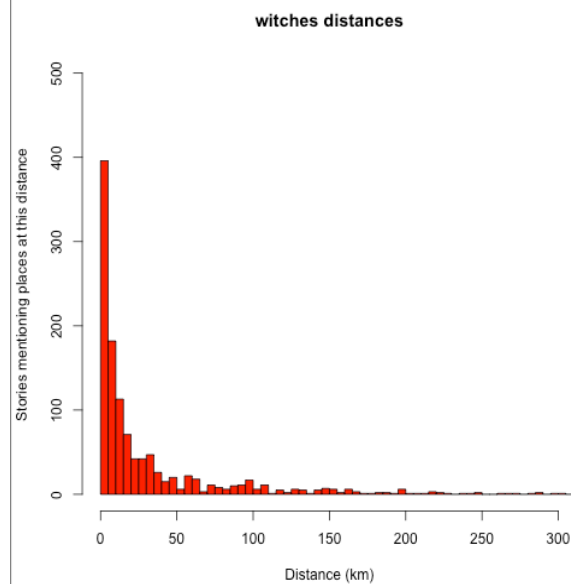
# Conceptual Maps

- How do people conceive the distance between the “inside” and the “outside” for various topics?
- What is the distance between a storyteller and the stories they tell by topic?
- How can one visualize the relationship between stories and the places they were told by topic?



# Distance from “home”

Distances, bearings, and frequencies of mentioned places for the most recently selected topic:

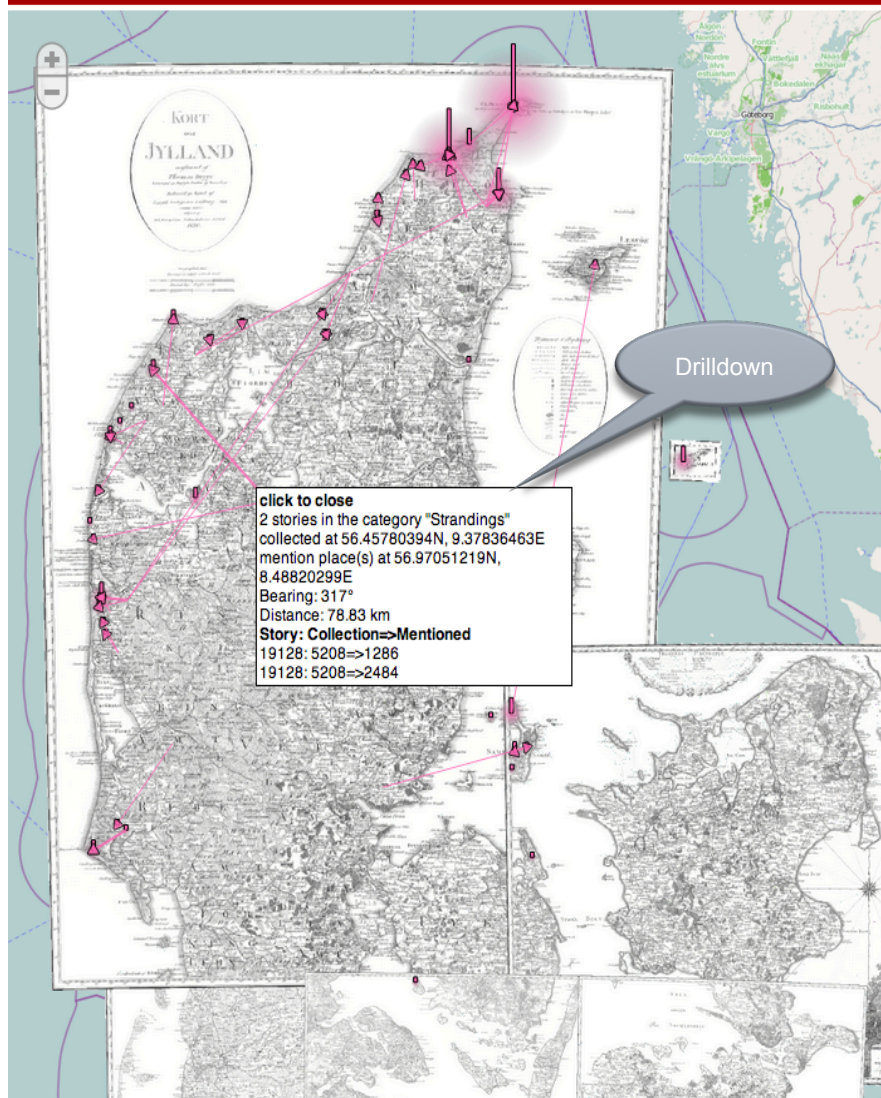


# Visualizing direction

## WitchHunter

Mapping the Evald Tang Kristensen Collection

Map data CC-BY-SA [OpenStreetMap](#), Fugro Aerial Mapping A/S



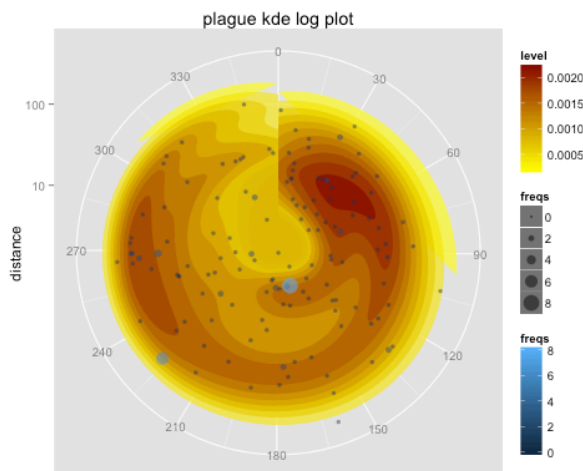
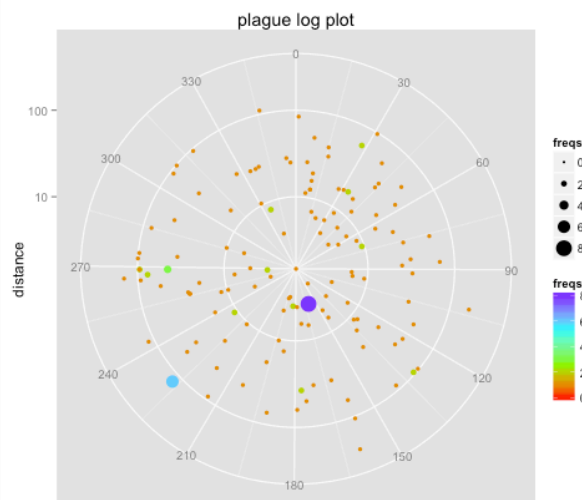
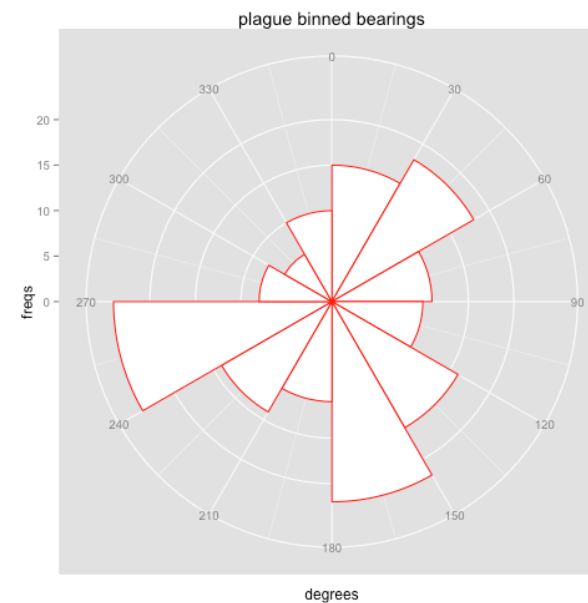
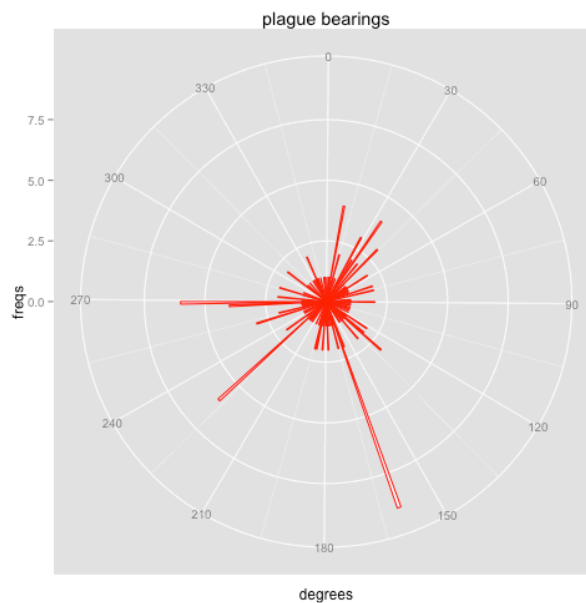
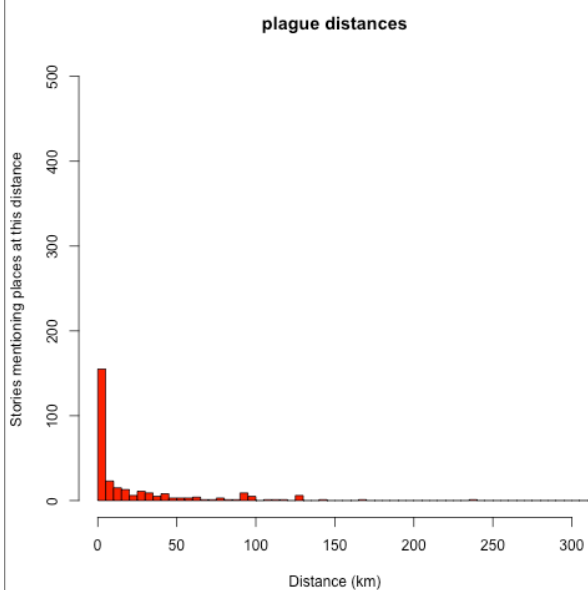
- ☒ Historical maps ☐ Aerial images
- ☒ Display place/topic co-occurrence markers
- ☒ Display place collected->mentioned vectors (↑)
- ☒ Display visualizations of spatial statistics
- ☒ Display stories related to map objects

### Story topic indices:

- ☐ Mound dwellers (Hidden folk) ■
- ☐ Elves ■
- ☐ Household spirits ■
- ☐ Traveling monsters ■
- ☐ Water spirits ■
- ☐ Wiverns and small creepy-crawlies ■
- ☐ Werewolves and nightmares ■
- ☐ Religious legends ■
- ☐ Death portents ■
- ☐ Lights and portents ■
- ☐ Heroes and their sport ■
- ☐ Churches and their construction. Monasteries, holy springs, etc. ■
- ☐ Legends about farms and towns ■
- ☐ Diverse place legends ■
- ☐ Legends about treasure ■
- ☐ Small kings and their feuds. Kings. Enemy invasions ■
- ☐ Manor lords, ladies and mistresses ■
- ☐ Ministers ■
- ☐ Diverse people ■
- ☐ Robbers, murderers and thieves ■
- ☒ Strandings ■
- ☐ Plague and illnesses ■
- ☐ Diverse hauntings and conjuring of revenants ■
- ☐ Female revenants ■
- ☐ Revenants on ill-gotten land ■
- ☐ Revenants in diverse places ■
- ☐ About the Devil and being in league with him ■
- ☐ Cunning men and women and their activities ■
- ☐ Witches and their sport ■
- ☐ Human and cattle illnesses ■
- ☐ Concerning earlier agricultural practices ■
- ☐ From the time of villeinage ■
- ☐ Houses and life in them ■
- ☐ Social gatherings and parties ■
- ☐ Outdoor life ■
- ☐ Our forebears' way of thinking and spiritual life ■

# The Black Death and relative directionality

Distances, bearings, and frequencies of mentioned places for the most recently selected topic:

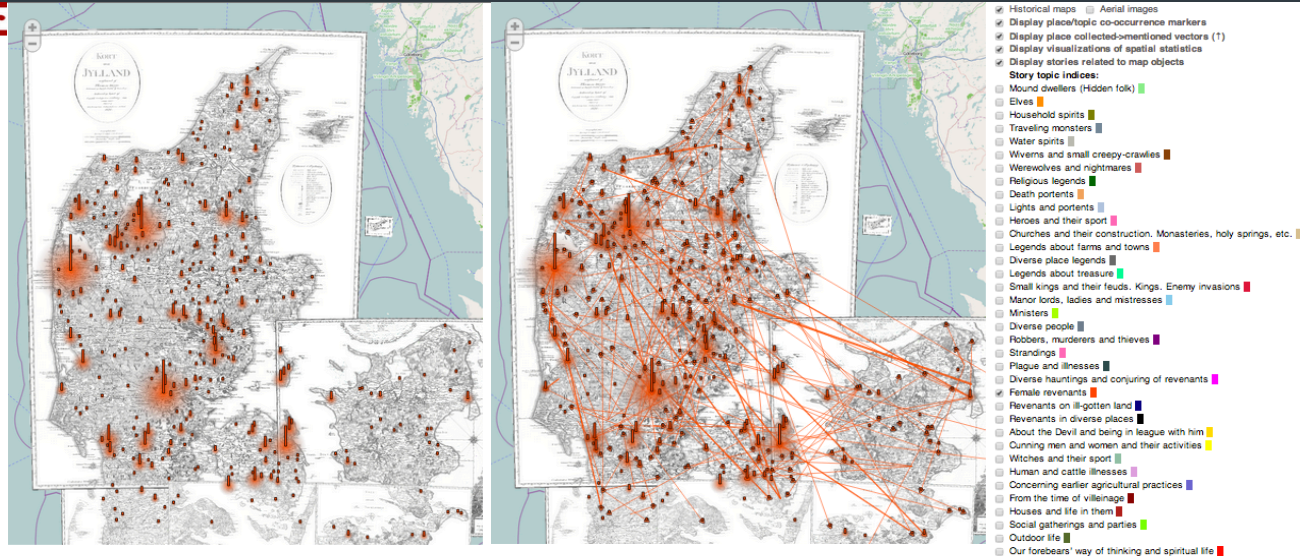




# Bringing it all together: Toward a Folklore Macroscope

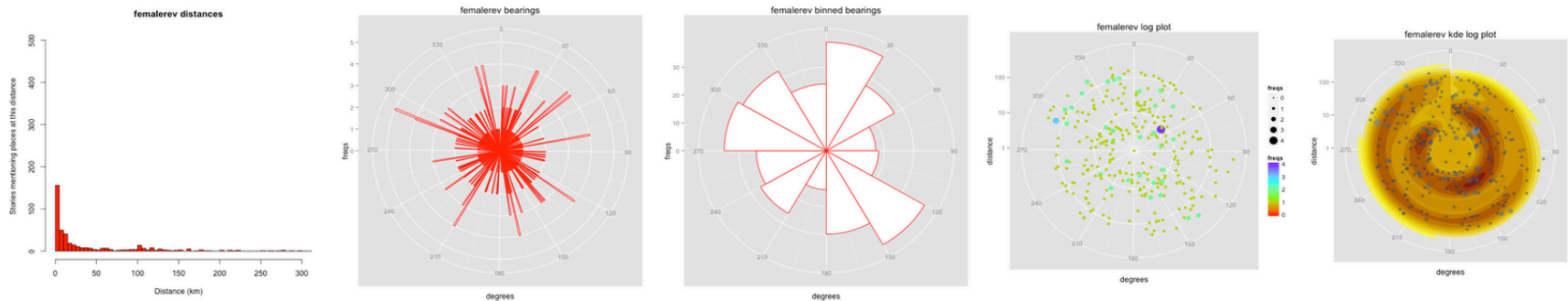


**WitchHunter**



Map data CC-BY-SA [OpenStreetMap](#), [Fugro Aerial Mapping A/S](#)

Distances, bearings, and frequencies of mentioned places for the most recently selected topic:



Stories for most recently selected graphic

**Story 19324:** DSnr\_04\_0\_01213  
informant: 30 (M. Balle), collection place: 4685 (København, 55.68488326N, 12.553387781E)  
Place mentioned: 1812 (København, 54.89912876N, 11.1239032E)  
ETK index: Pest og sygdomme  
Danish publication text: 1213. Naar man pøder en /Eblekvist i et Pæretæe eller omvendt, vil en frugtsommelig Kvinde, der spiser af et saadt Træes Frugt, føde et Barn, som er baade Pige og Dreng. M. Balle, København.

**Story 19273:** DSnr\_04\_0\_01162  
informant: 43 (Professor H. V. Rasmussen), collection place: 4685 (København, 55.68488326N, 12.553387781E)  
Place mentioned: 1812 (København, 54.89912876N, 11.1239032E)  
ETK index: Pest og sygdomme  
Danish publication text: 1162. Man hinder et lille Barn en Føltand om Halsen, naar det skal hare Tænder. Bliver denne hængende, vil Barnet leve; og om- vendt, det doer, naar den falder af. H. V. Rasmussen, København.

**Story 19294:** DSnr\_04\_0\_01183  
informant: 43 (Professor H. V. Rasmussen), collection place: 4685 (København, 55.68488326N, 12.553387781E)  
Place mentioned: 1812 (København, 54.89912876N, 11.1239032E)  
ETK index: Pest og sygdomme  
Danish publication text: 1183. Naar en Ko bander en, og man ikke spytter, faar man Bleg- nor paa Øjet. H. V. Rasmussen, København.

**Story 19321:** DSnr\_04\_0\_01210

Story Drilldown





# Folklore Macroscope

- <http://etkspace.scandinavian.ucla.edu/macroscope.html>



## 2. Old Icelandic Language and Literature

- Challenges with highly inflected languages
- Character networks as social networks: Icelandic Family Sagas as model



# Old Norse Morphology

- Highly inflected
- Difficult to search
- Difficult to generate basic statistics
- Very difficult to use machine learning techniques with consistency



# IceMorph

- Challenge: given very small training data, can one achieve competitive accuracy on a morpho-syntactic tagging?
- Classic problem in Humanities:
  - Very low amounts of funding mean one needs to be very efficient at finding where one wants to / needs to apply effort
- We discovered that, with minimal tagging, we could achieve corpus-wide POS tagging accuracy of 96% and MS tagging accuracy of 84%



# Some comparisons (i.e. the obligatory inscrutable table)

Method	POS EXPERT	POS GOLD	MS EXPERT	MS GOLD
Bayes-base	95.43%	79.25%	80.67%	48.34%
Bayes-80/20	85.71%	75.14%	62.37%	43.24%
HMM-dV-base	93.85%	25.60%	75.82%	13.62%
HMM-dV-80/20	93.68%	34.74%	82.11%	18.75%
HMM-rV-base	96.11%	71.58%	79.92%	53.98%
HMM-rV-80/20	96.84%	73.16%	84.21%	54.86%
CRF-1-base	89.75%	36.58%	78.07%	11.54%
CRF-1-80/20	87.30%	46.07%	77.78%	16.55%
CRF-2-80/20	84.13%	48.69%	56.08%	17.24%



# IceMorph

- <http://tango.bol.ucla.edu/andreas09/index.html>
- And the obligatory equation:

For a given  $t \in T$  and observations  $o_1, \dots, o_n$  we find the most likely state sequence by solving

$$V_{n,k} = P(o_n | k) * \max_x (t * V_{n-1,x})$$

for a given element  $x$  in the sequence.



# NLP tools and taggers

- Stanford CoreNLP
  - <http://stanfordnlp.github.io/CoreNLP/>
- LingPipe
  - <http://alias-i.com/lingpipe/>
- Open Calais
  - <http://www.opencalais.com/>
- MALLET
  - <http://mallet.cs.umass.edu/>



# Character Networks as Social Networks: Facebook for Vikings



## Egil's expedition with Hakon's men

The trail forks here... The farmer who lives beneath the ridge is named Arnald and he's a friend of ours. We will go and stay with him, and you should go up on the ridge. When you get there you'll soon see a big farm where you are sure of a place to stay. A very wealthy man named Armod Beard lives there. We will meet up again early tomorrow morning and go to Eideskog in the evening. A farmer lives there, a good man named Thorfinn (Scudder 2001, 137)

# Sagas and Social Networks

## ■ Sagas

- > 300 characters in many family sagas (e.g. Egils saga, Njals saga)
  - ES n= 320
  - NS n= 707
- These characters may appear in other sagas in varying roles (e.g. Egil in Laxdoela saga)
- These sagas also intersect with historical sources (e.g. Landnámabók)

## ■ Social Networks

- Measures of social connectedness may help us understand the complexities and nuance of social interaction understood by saga authors and audiences
- Changes in the social network over time may shed light on social decision-making in the fictional world of the saga

## Network model of the sagas

- Saga social network is a metafictional representation of idealized fictional interactions created for aesthetic and rhetorical ends
- What happens to one's understanding of the saga, or all of saga literature, if one can view the social landscape from the perspective of any saga character?
- Does a social network model of the saga(s) help address research questions concerning social complexity in the saga literature?

# Challenge questions

- How can SNA help us understand consanguineous, affinal, and fostering relationships in the saga?
- How are relationships coded in the saga?
  - Is this consistent across individuals and groups?
  - To what extent are relationships based on structural (social) balance theory (Heider 1946), and to what extent are they based on calculations of status (Leskovec, Huttenocher and Kleinberg 2010)?



# Data preparation

- Extracting network data from sagas
  - Nodes = characters
  - Edges = interactions
- Developing a dynamic model of networks
  - Use chapters as first level temporal distinction
  - Use pages as a second level temporal distinction
- Appropriately valuing connections
  - Acknowledging that not all connections are equal (e.g. friendship vs. enmity)
  - Acknowledge that not all connections are equally “weighted” (e.g. lethal hostility is weightier than nonlethal hostility)



- Edgelist
  - List of characters
  - List of interactions
- List of characters
  - Each character is listed once with a unique identifier
- List of interactions (simple, based on anthropological concepts and SNA concepts, non-ontological)
  - Twenty types of actions
  - Nine types of relationships
  - Twelve types of events
  - Thousands of geo-coded placenames

# Actions

- descent
- marriage or fostering
- allegiance or friendship
- request or offer assistance
- provide or discover information
- intervention
- invitation
- giftgiving
- betrothal
- death\_non-hostile
- inheritance or succession
- placed in command
- accusation or summons
- insult
- lying
- threat
- challenge
- hostility\_non-lethal
- hostility\_lethal
- conversation

# Relationships / attributes

- Gender (male/female/corporate)
- Affinal-cosanguineous
  - Parent
  - Sibling
- Affinal
  - Conjugal (spouse)
  - Fostering
- Kinsman (unspecified)
- \*Neighbor
- \*[Enemy]
- \*[Friend]



# Types of graphs

- Networks of Characters
  - agent x agent
- Networks of Motifs and Themes
  - event x action
  - agent x event
  - agent x action
- Networks of Places
  - event x location
  - agent x location
- Networks of Interactions
  - agent x agent
- Networks of Genealogical relations
  - agent x agent
- Unions of these graphs

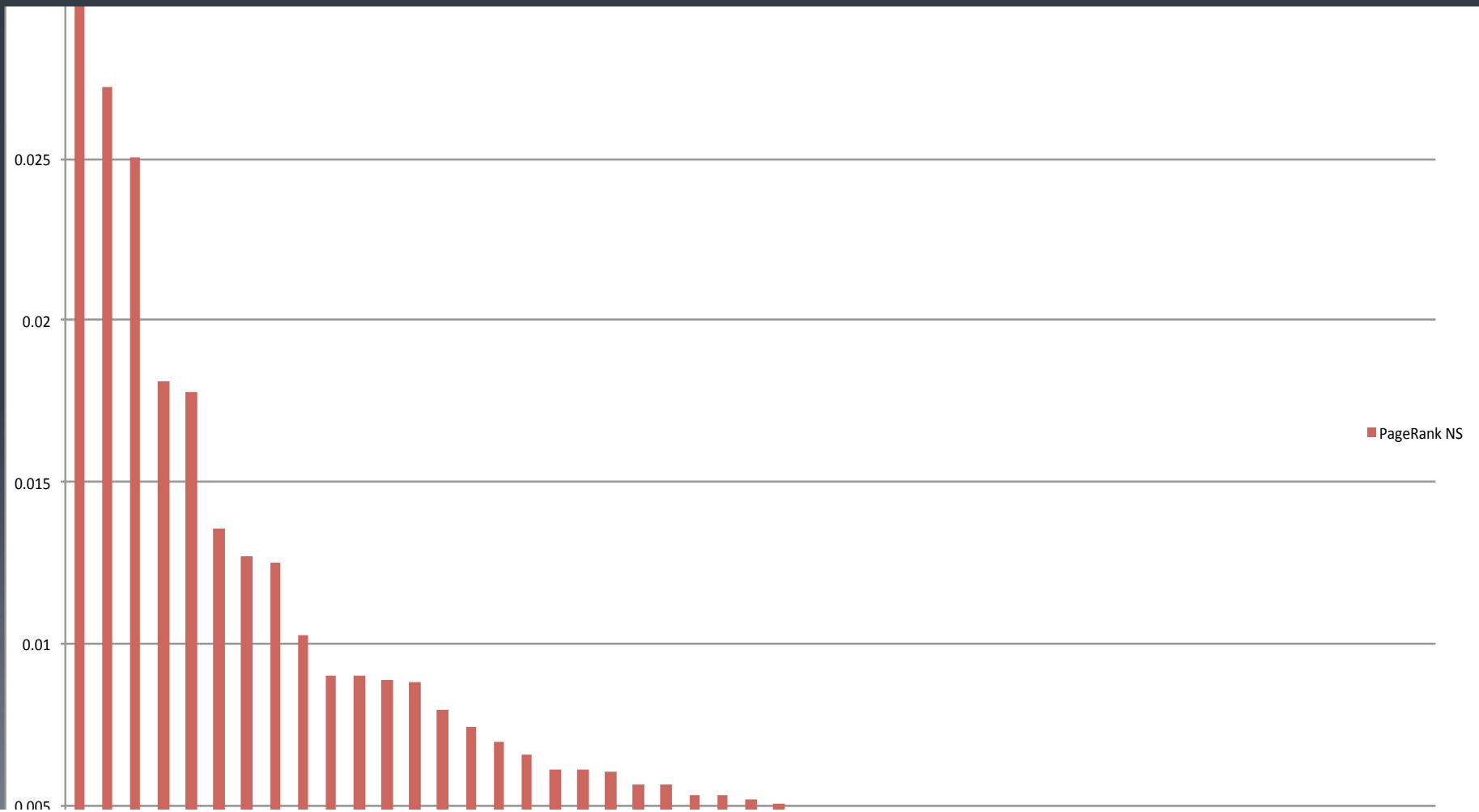
# Egils saga general stats

- Extracted Egils saga network by hand
- Some characteristics of Egils saga network:
  - 320 characters (most individual, although some corporate—e.g. the Finns)
    - 258 male / 58 female / 4 corporate
  - 1352 interactions
    - 431 genealogical interactions
    - 316 friendly interactions
    - 151 hostile interactions
    - 454 ambiguous or unvalued (inter)actions
- Some general statistics of Union graph
  - Number of components: one major component with 318 characters, one minor unconnected component with 2 characters (Bragi Boddason and King Bjorn of Sweden)
  - Graph density: .0083 (total number of edges / total number of possible edges)
  - Avg. path 4.45
  - Avg. total degree 3.51
    - in-degree 1.75
    - out-degree 1.75

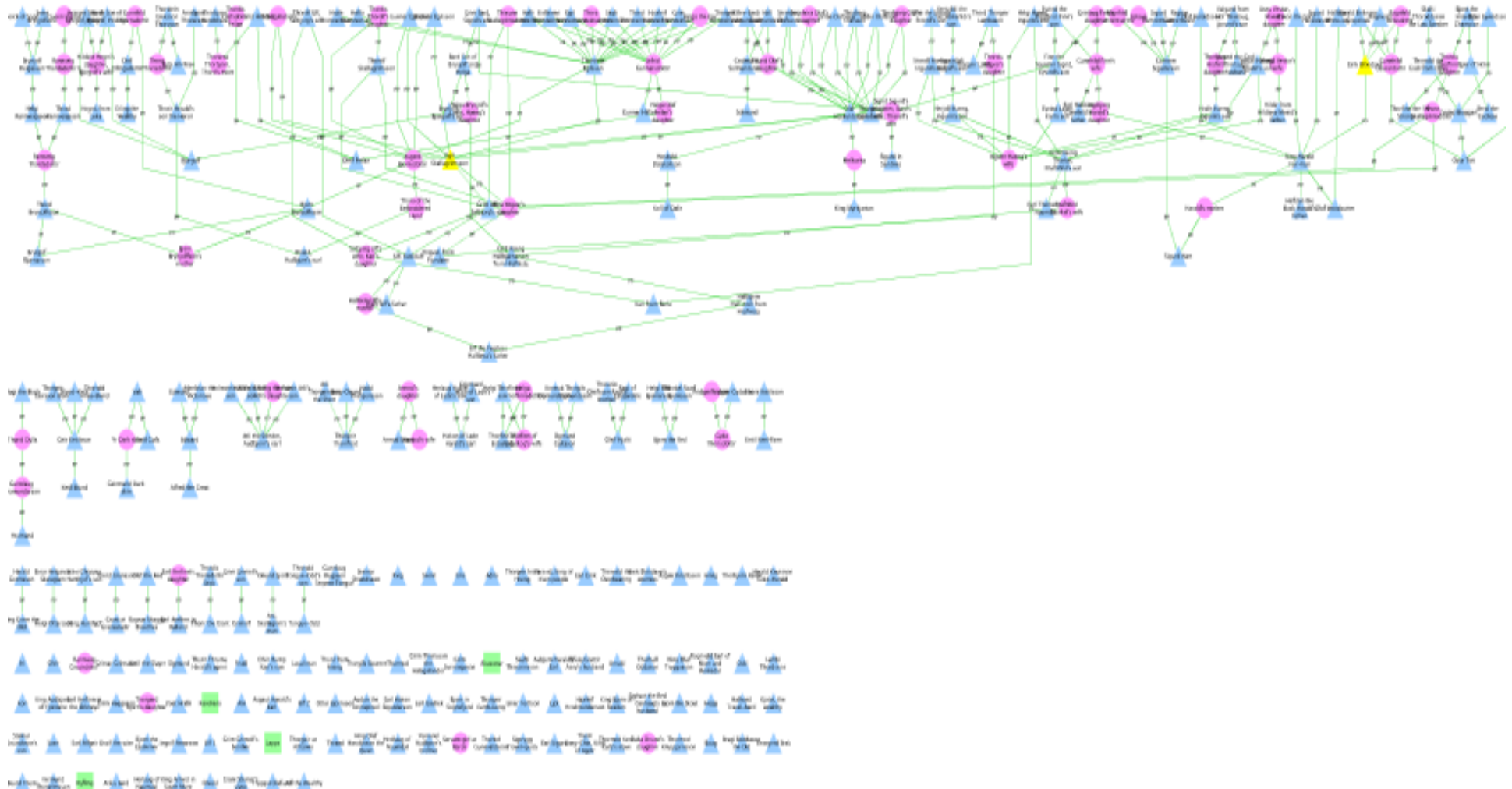
# Njals saga general stats

- Extracted Njals saga network by hand
  - More complex extraction than ES
    - Added “lives at”, “events”, and geo-located places
- Some characteristics of Njals saga network:
  - 707 characters (most individual, although some corporate—e.g. the Finns)
    - 565 male / 102 female / 47 corporate
  - 3055 interactions (including 648 conversations)
    - 448 genealogical interactions
    - 776 friendly interactions
    - 584 unfriendly (or downright hostile) interactions
    - 1247 ambiguous or unvalued (inter)actions
- Some general statistics of union graph
  - Number of components: one major component with 681 characters,
  - Graph density: .006 (total number of edges / total number of possible edges)
  - Avg. path 5.07
  - Avg. total degree 2.58
    - in-degree 1.29
    - out-degree 1.29

# PageRank in Njals saga— all interactions (fairly typical power law)

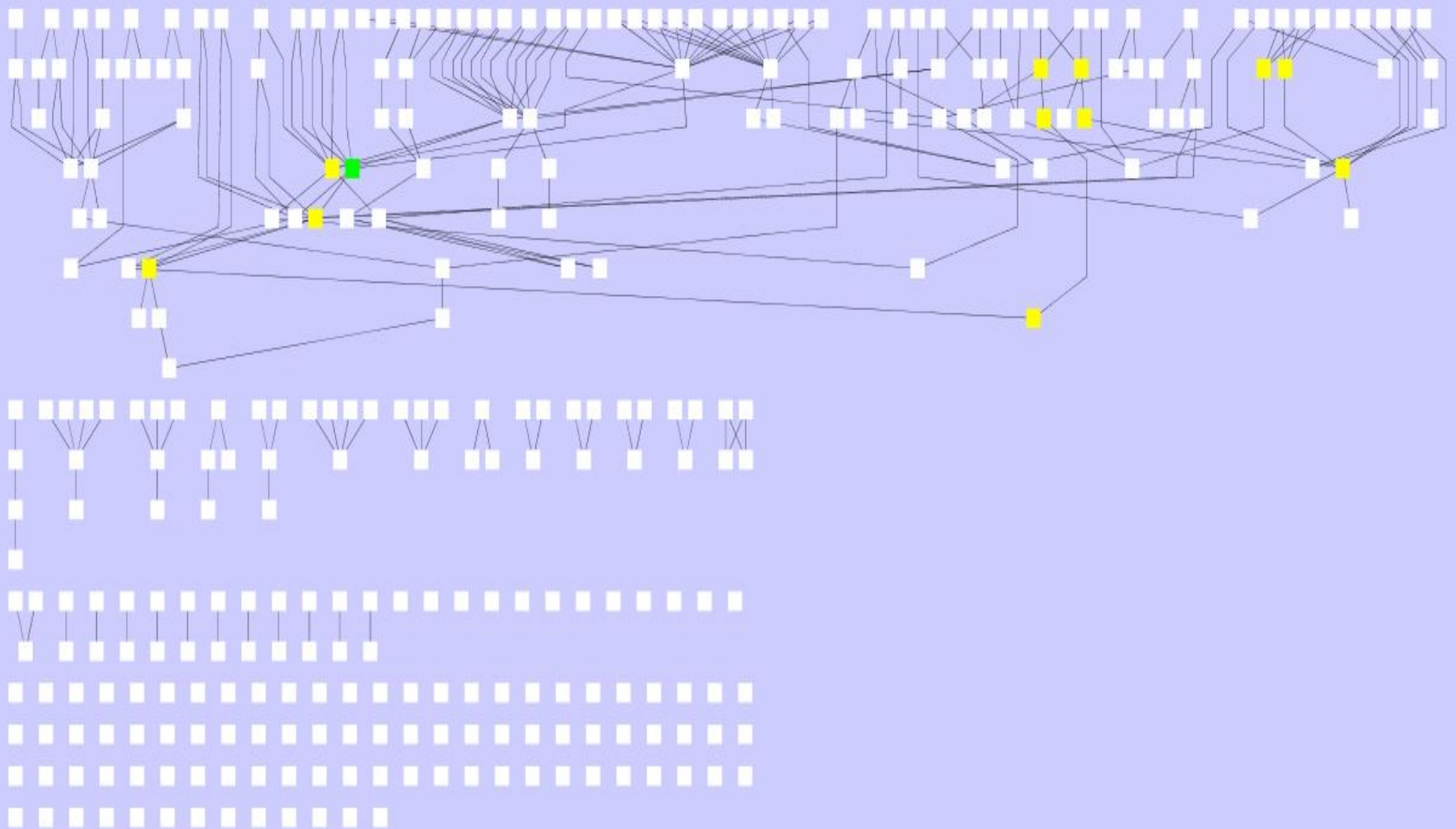


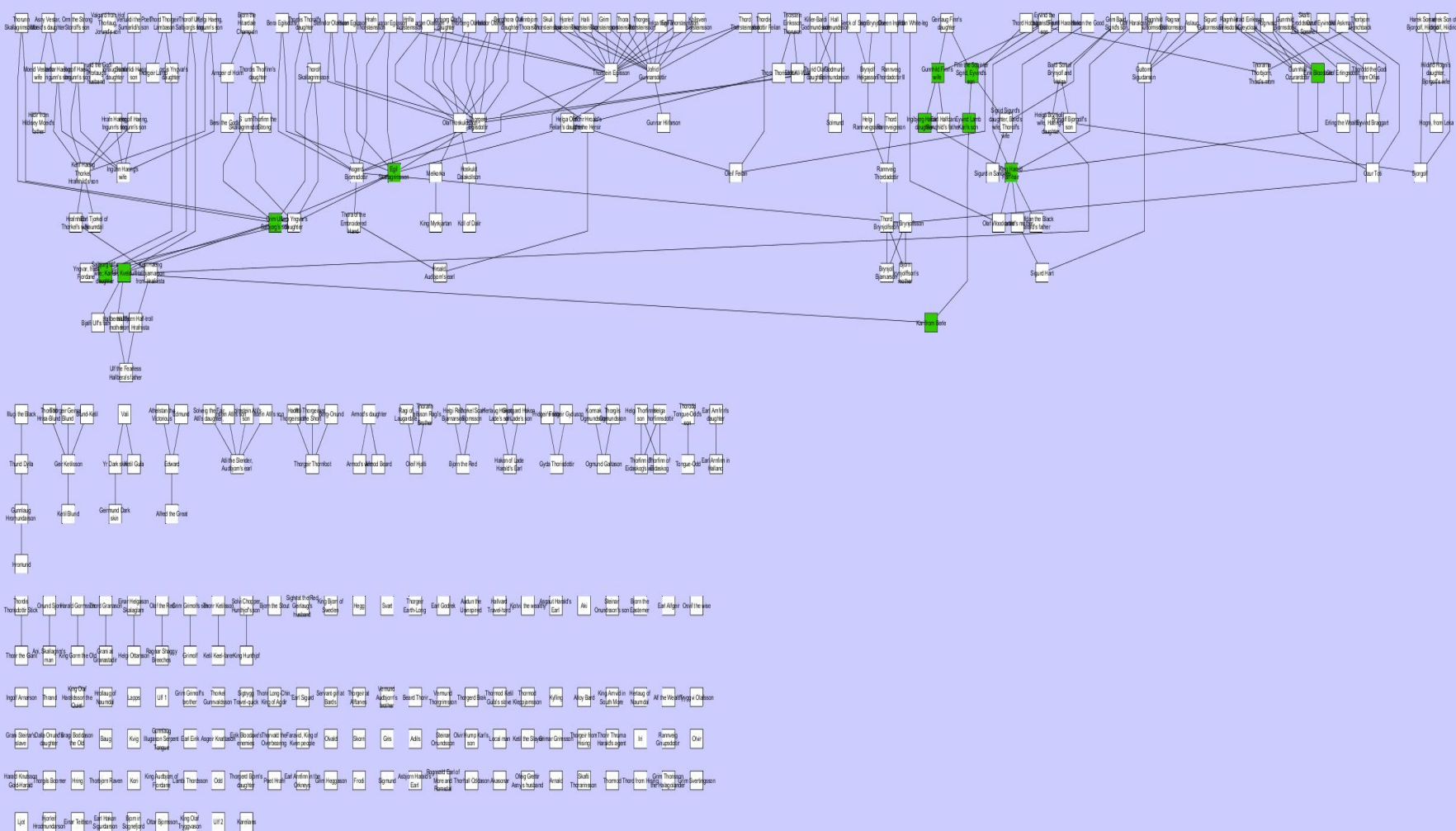
# Egils saga--genealogy

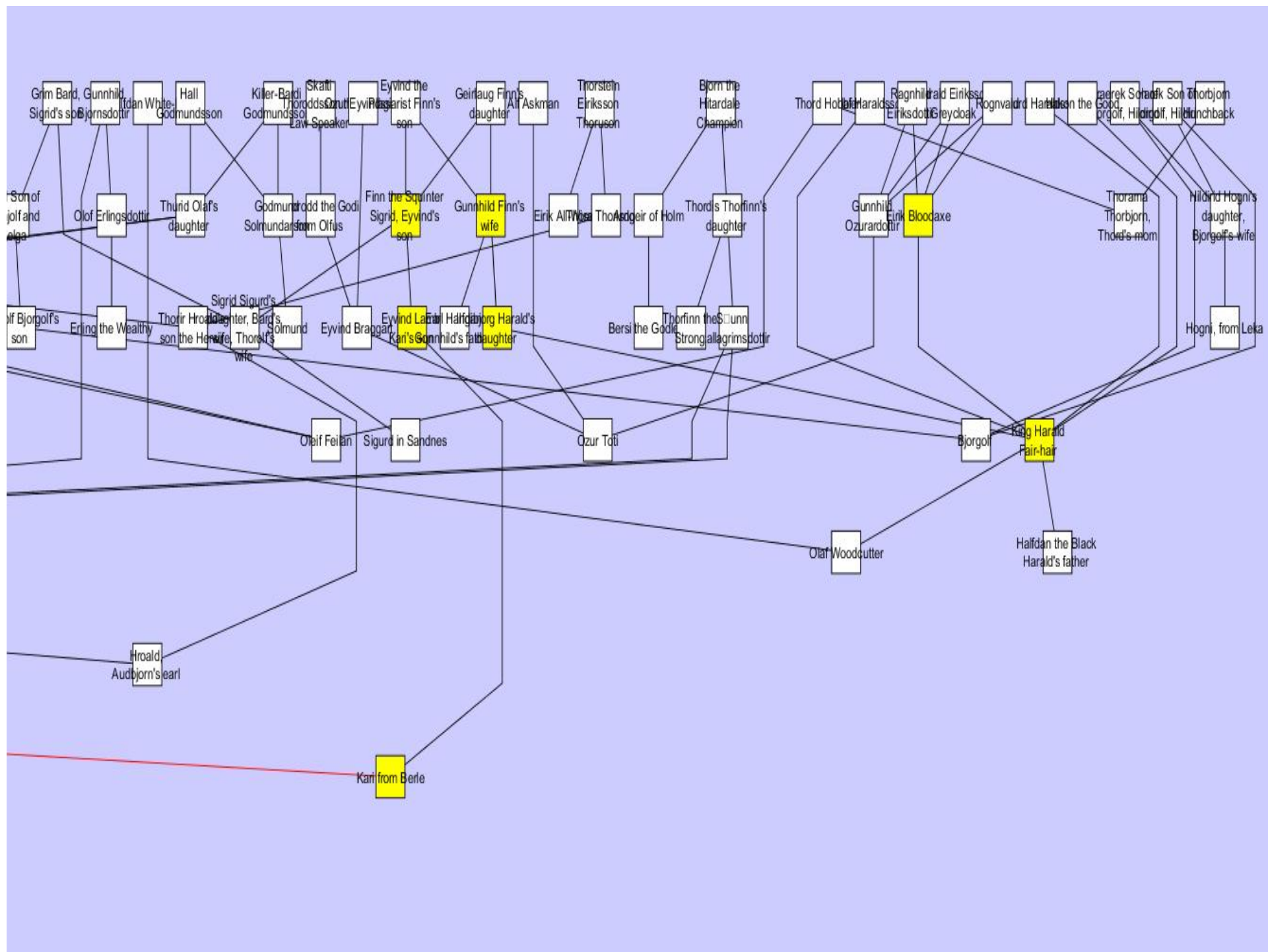




# Egils saga—no fosters

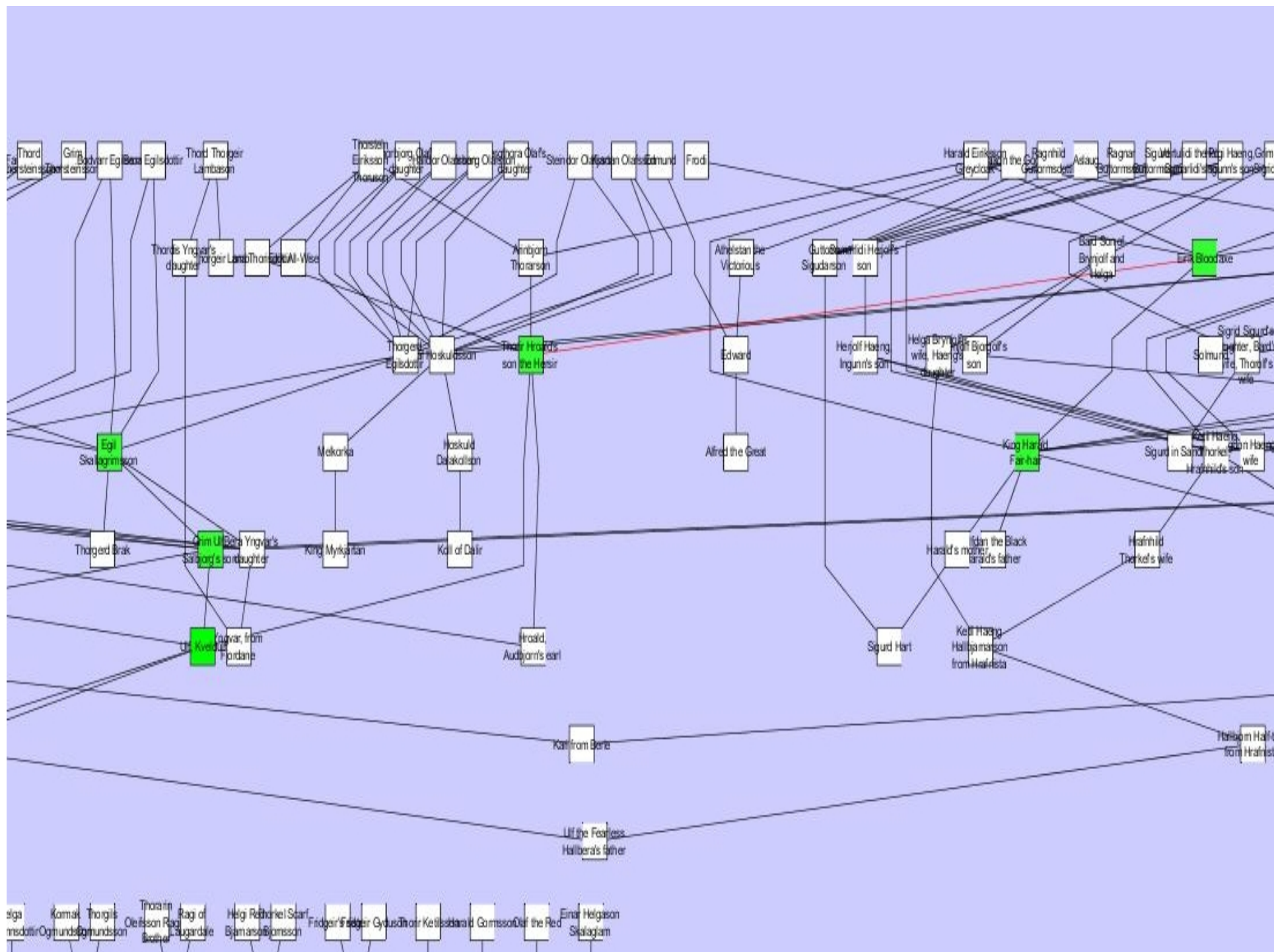




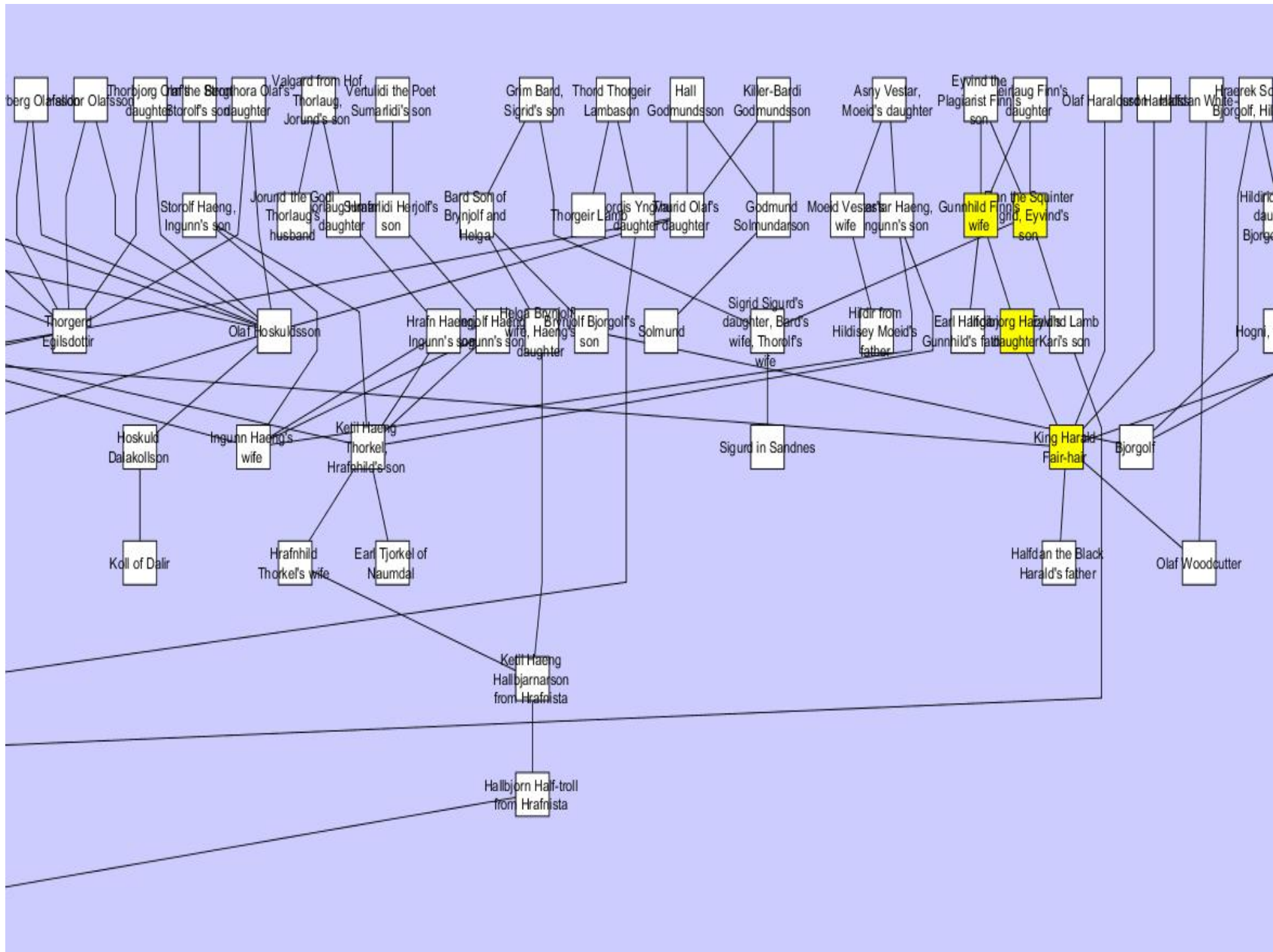






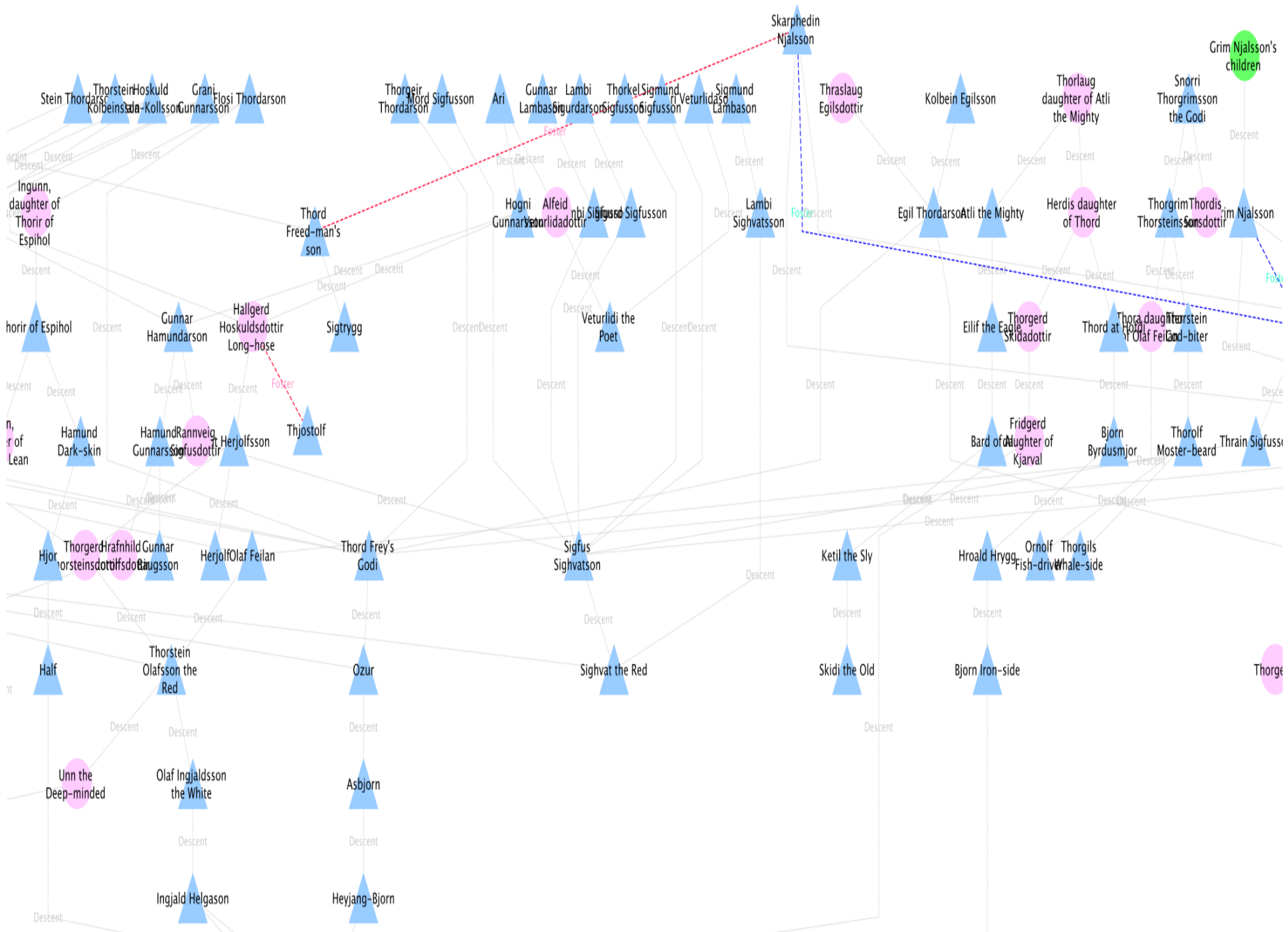


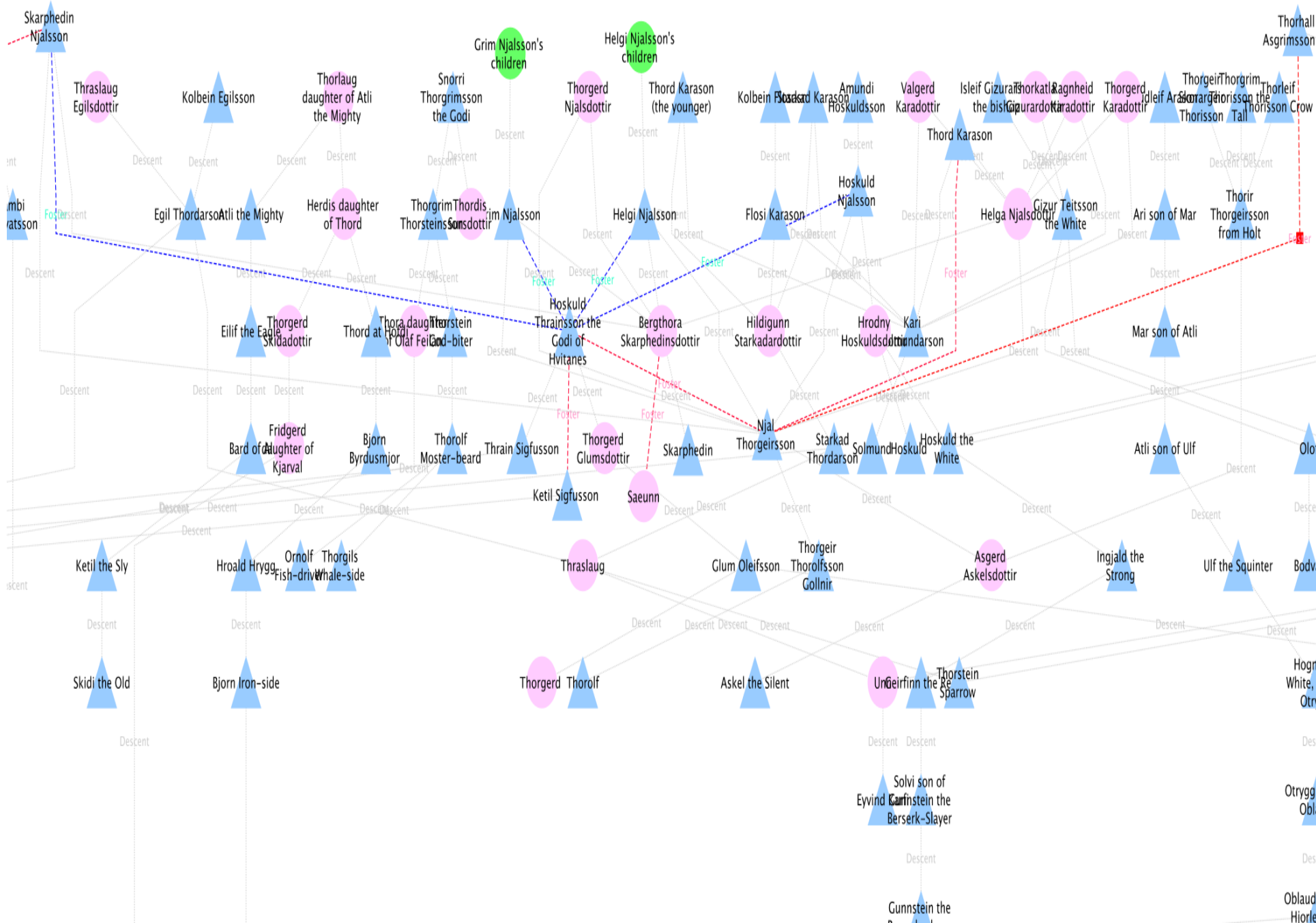




# Njals saga--genealogy







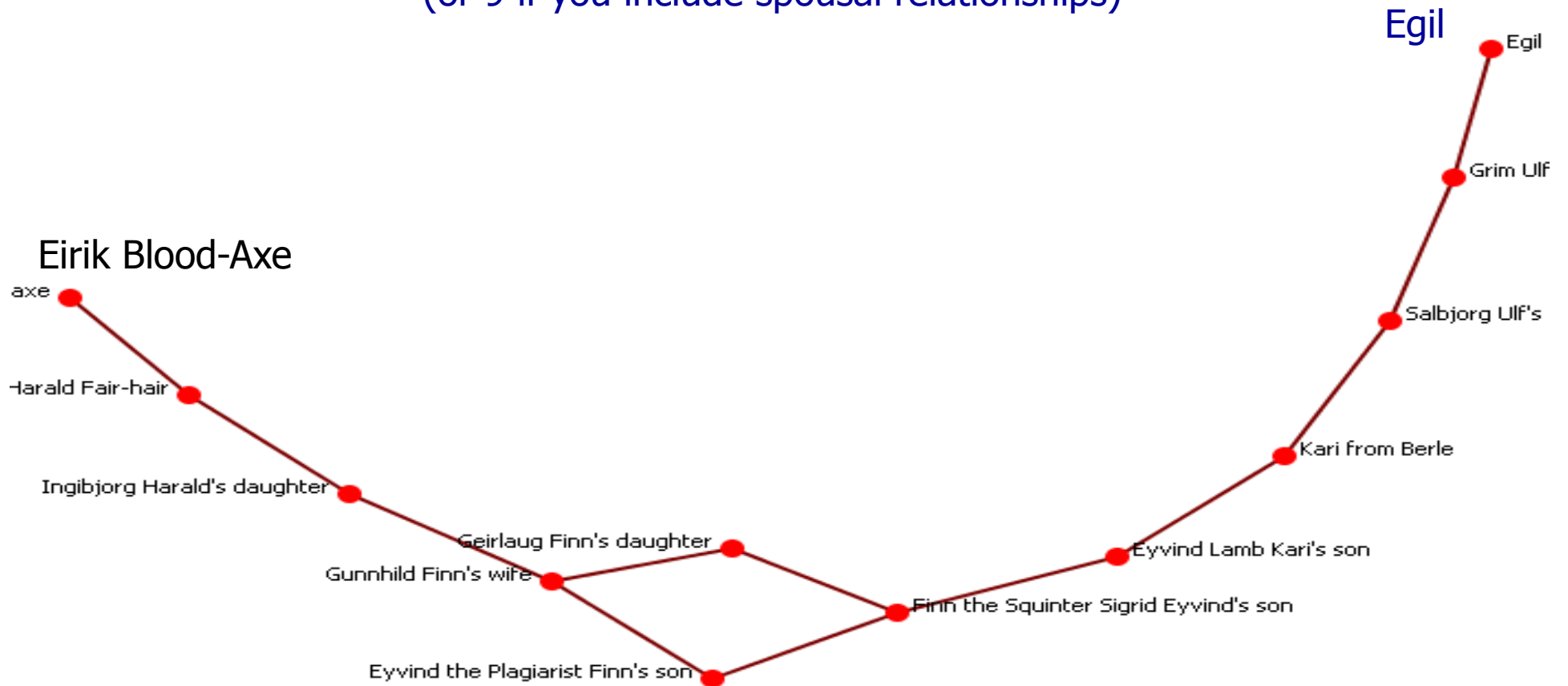


# Kinship Shortest paths

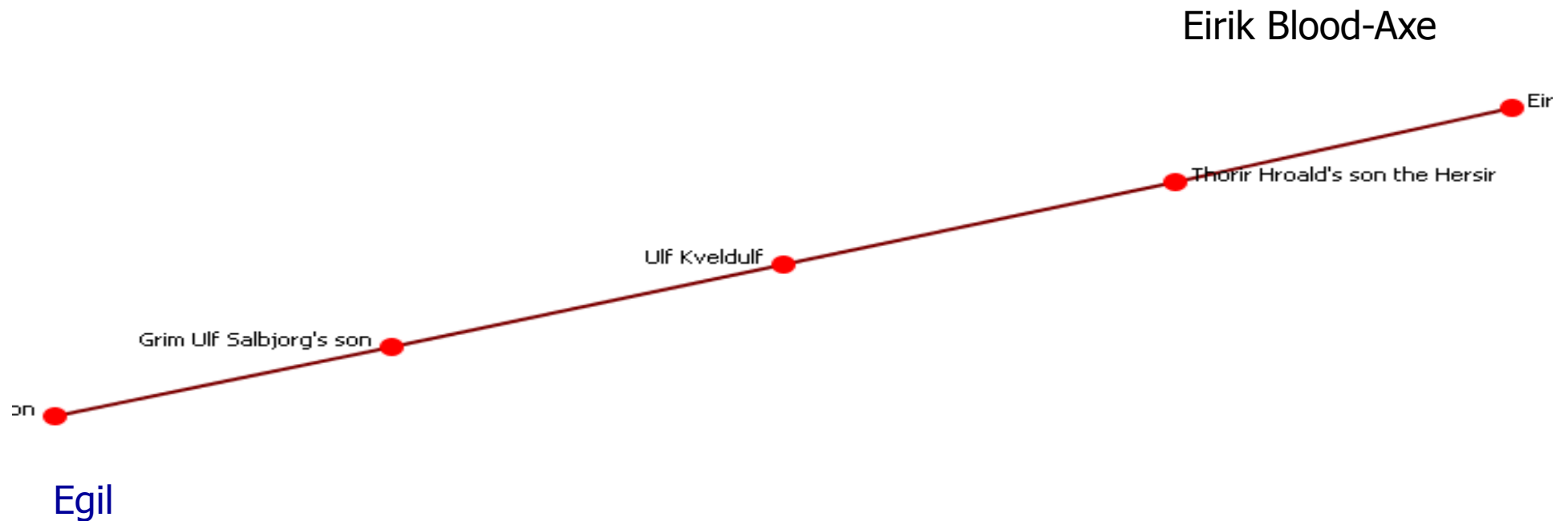
- Since the saga is about blood feud and
- Since many of these feuds revolve around problems in the valuation of kinship vs friendship and
- Since fostering significantly changes the pathways in a network by adding only a small number of edges
- It might be worthwhile to trace shortest path distances between saga antagonists purely on the kinship graph
- What happens to the ES and NS graphs between antagonists when fosters are added?
- Is fostering a good strategy to curtail or interrupt feud?

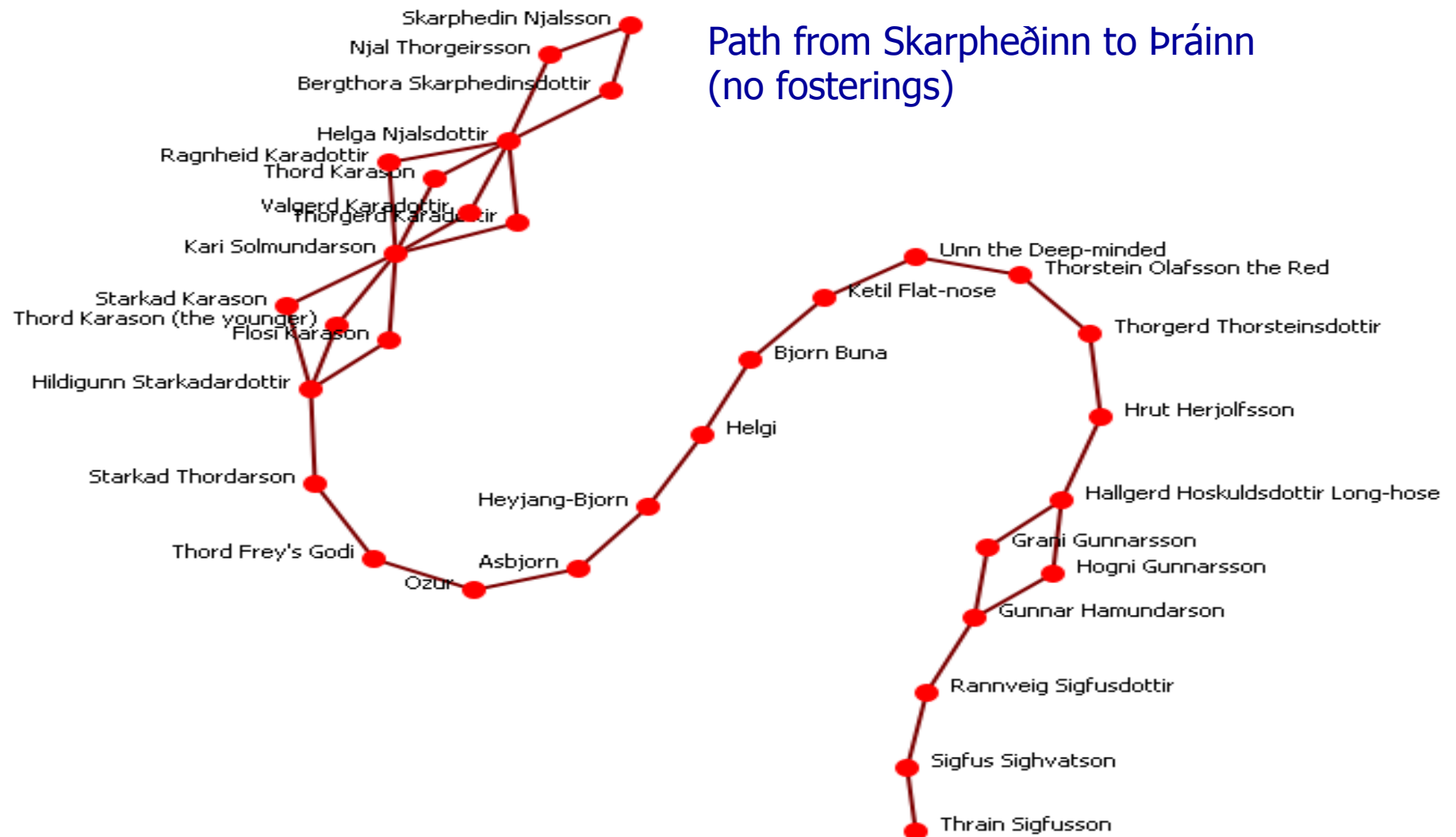


Path from Egil to Eirik of 10  
(or 9 if you include spousal relationships)



## Path from Egil to Eirik of 4 (includes fosterings)





## Path from Skarpheðinn to Þráinn (with fosterings)



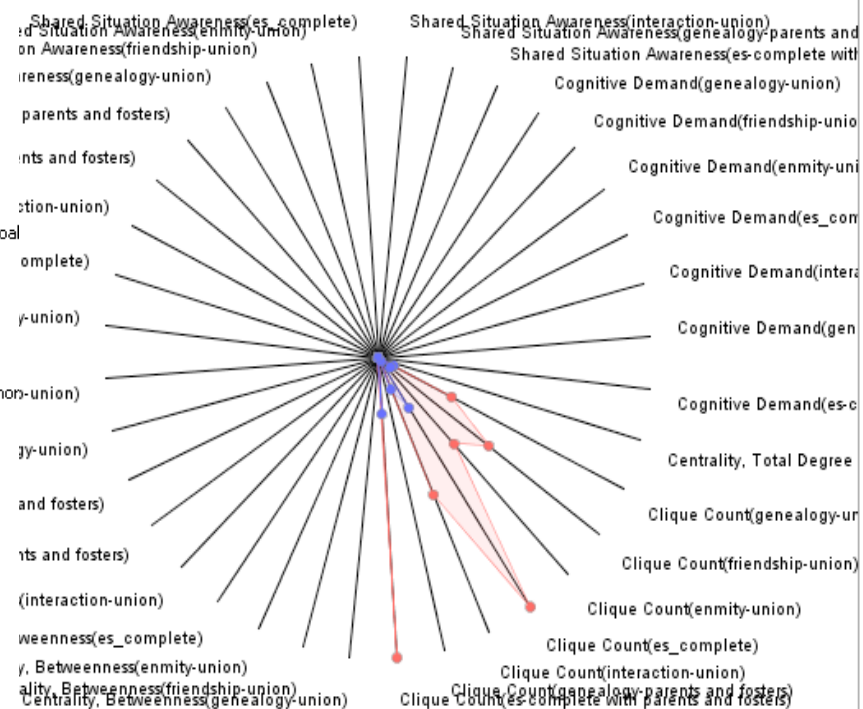
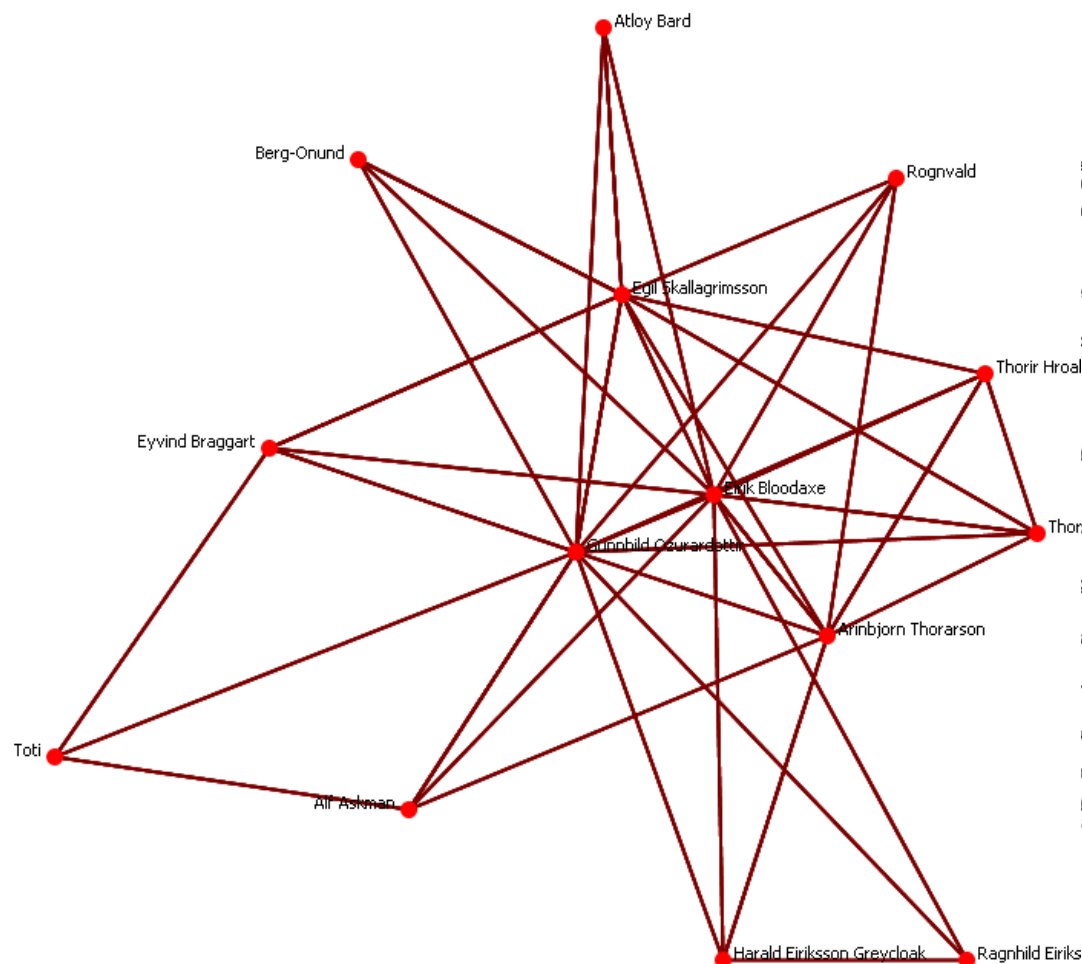


# Egonets

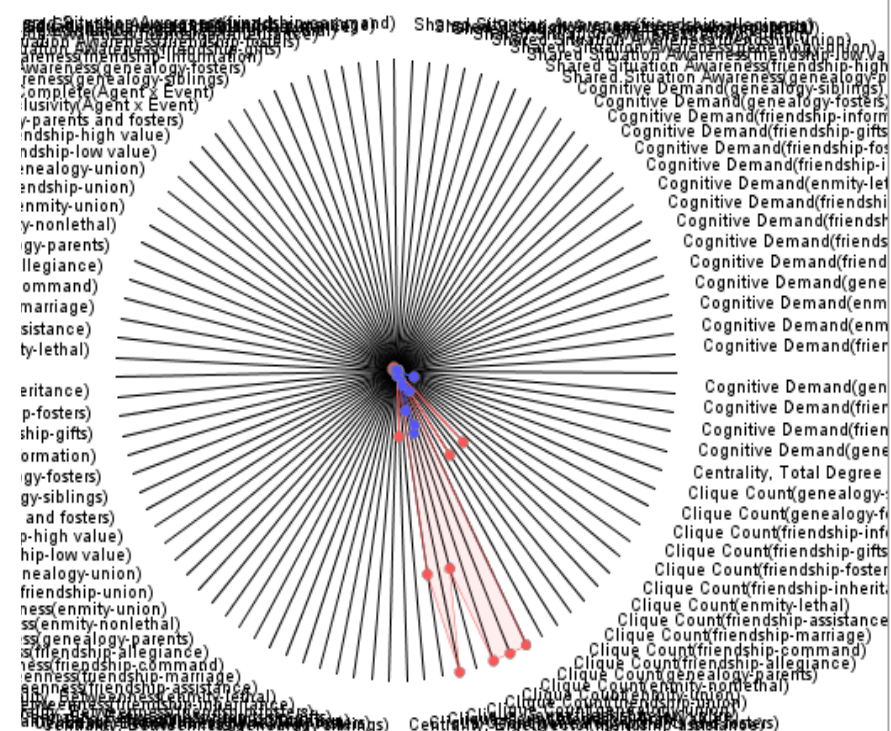
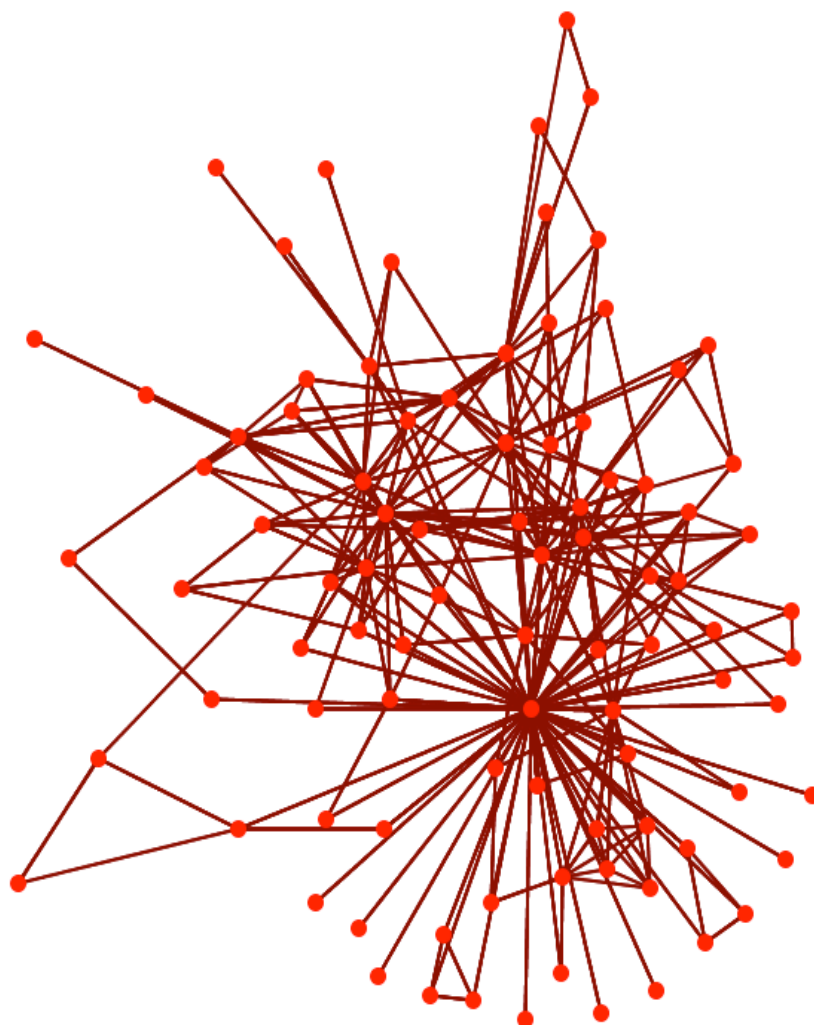
- Read the action from another character's perspective...
  - Are secondary characters really secondary?
  - A tangential character in one saga can be a major figure in another saga (eg. Egil in Egils saga and Laxdoela saga)



## Ego and Average Measure Values



### Ego and Average Measure Values



powered by DRA, CASOS Center @ CMU

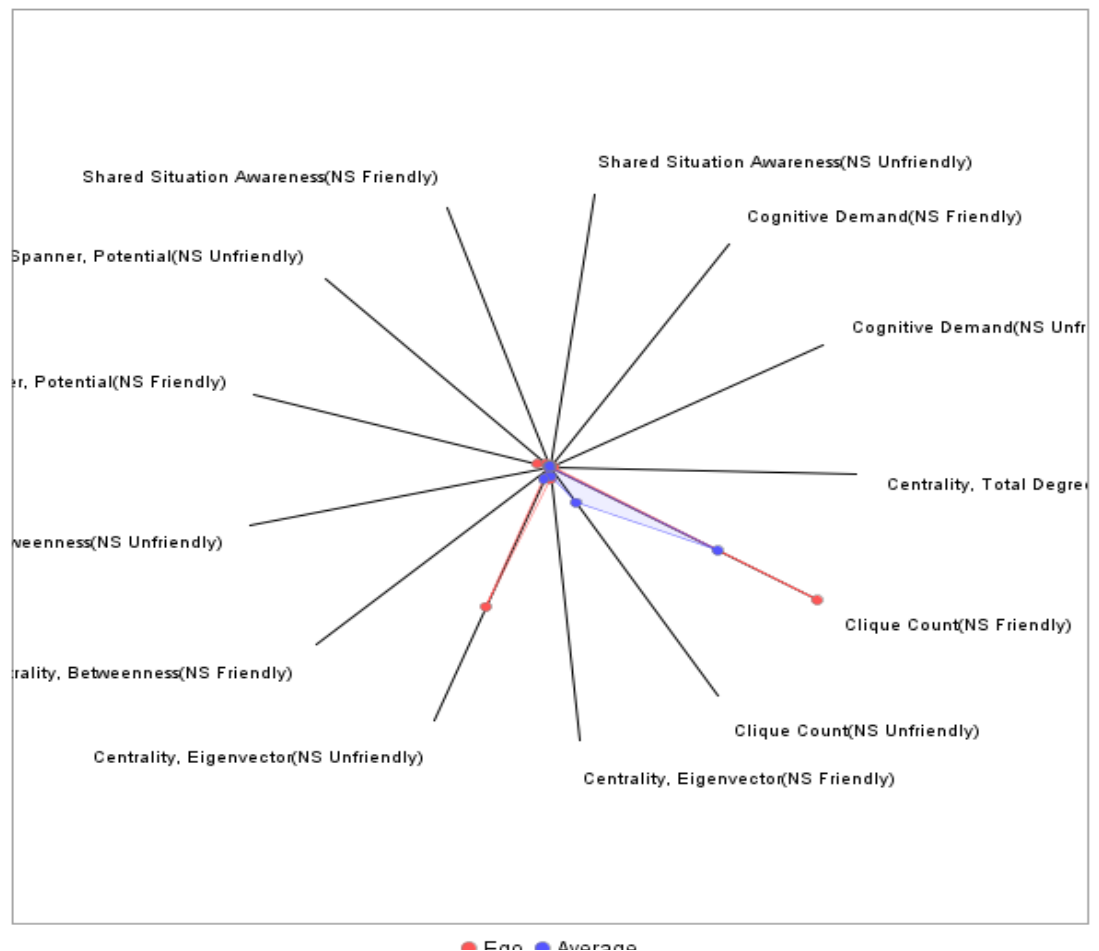
● Ego ● Average

# NS: Hallgerðr Höskuldsdóttir

Hallgerðr: unfriendly radius 1



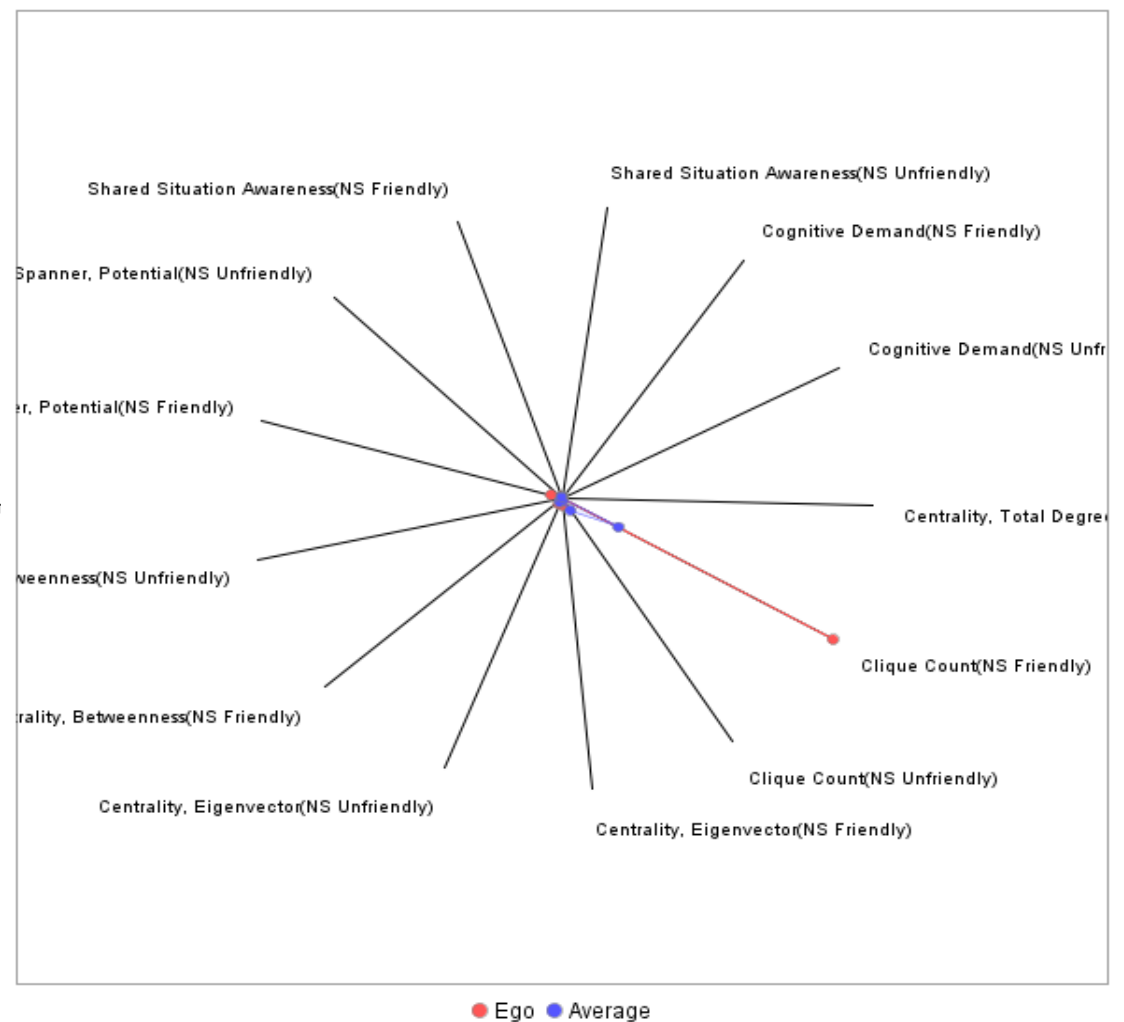
Ego and Average Measure Values



## Bergþóra: unfriendly radius 1

Hallgerd Hoskuldsdottir Long-hose — Bergþóra Skarphedinsdottir — Flosi Th

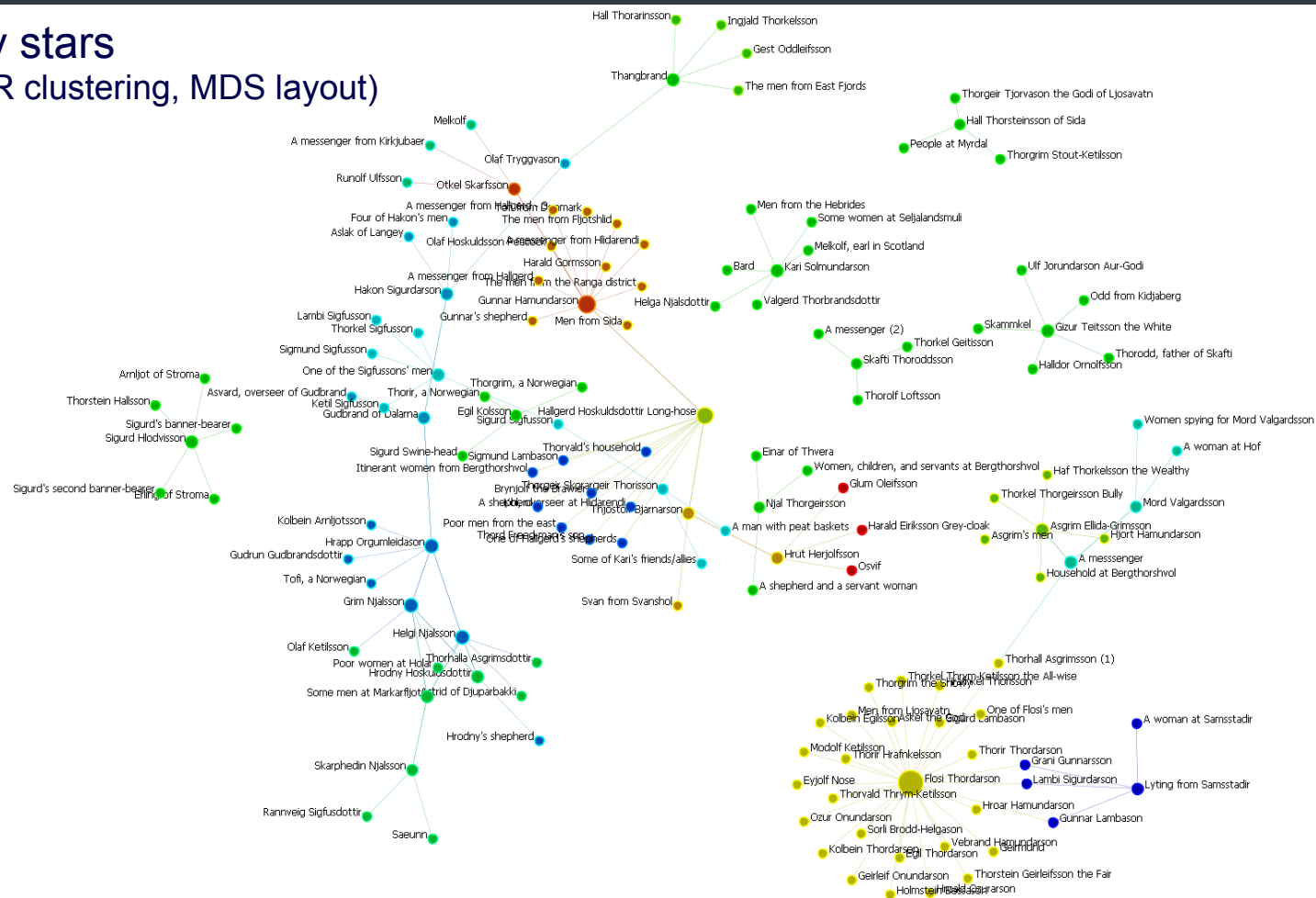
Ego and Average Measure Values



# Cliques and Stars: NS Friendly

## Friendly stars

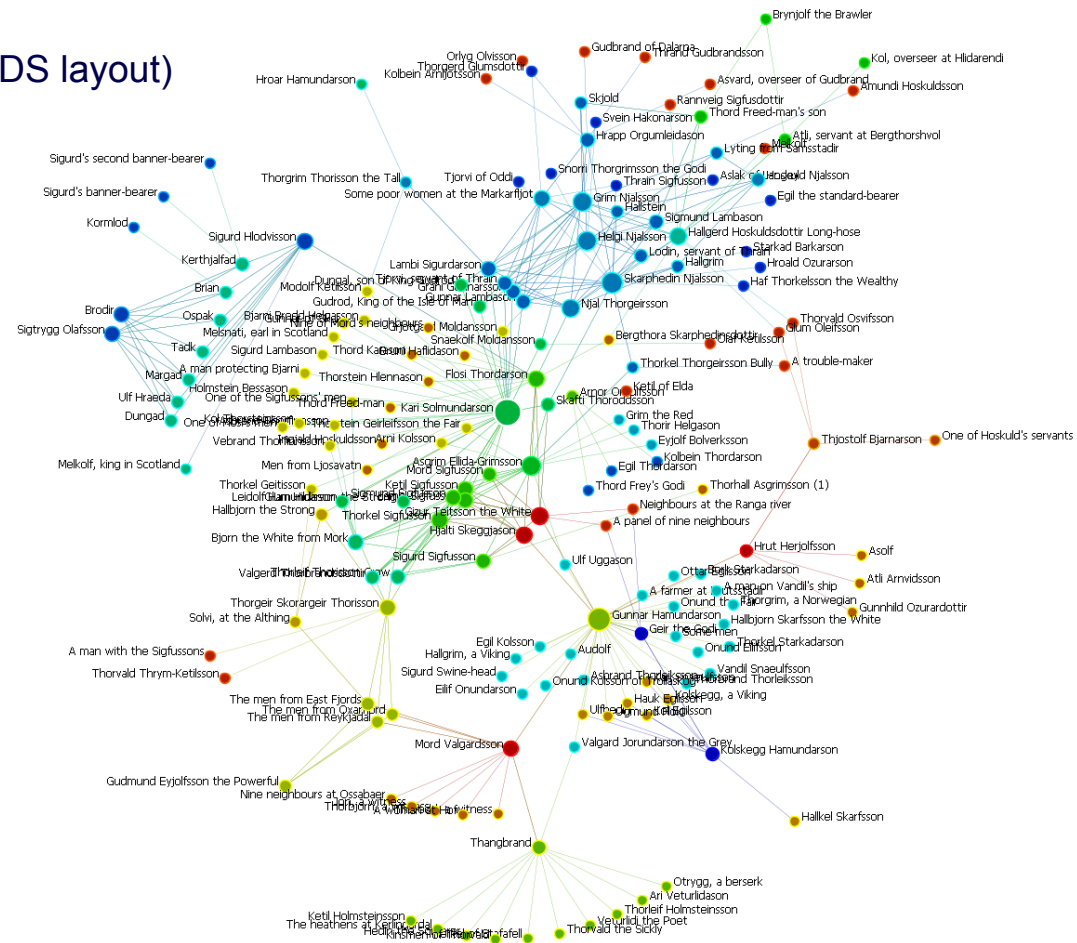
(CONCOR clustering, MDS layout)





# Cliques and Stars: NS Unfriendly

## Unfriendly stars (CONCOR clustering, MDS layout)



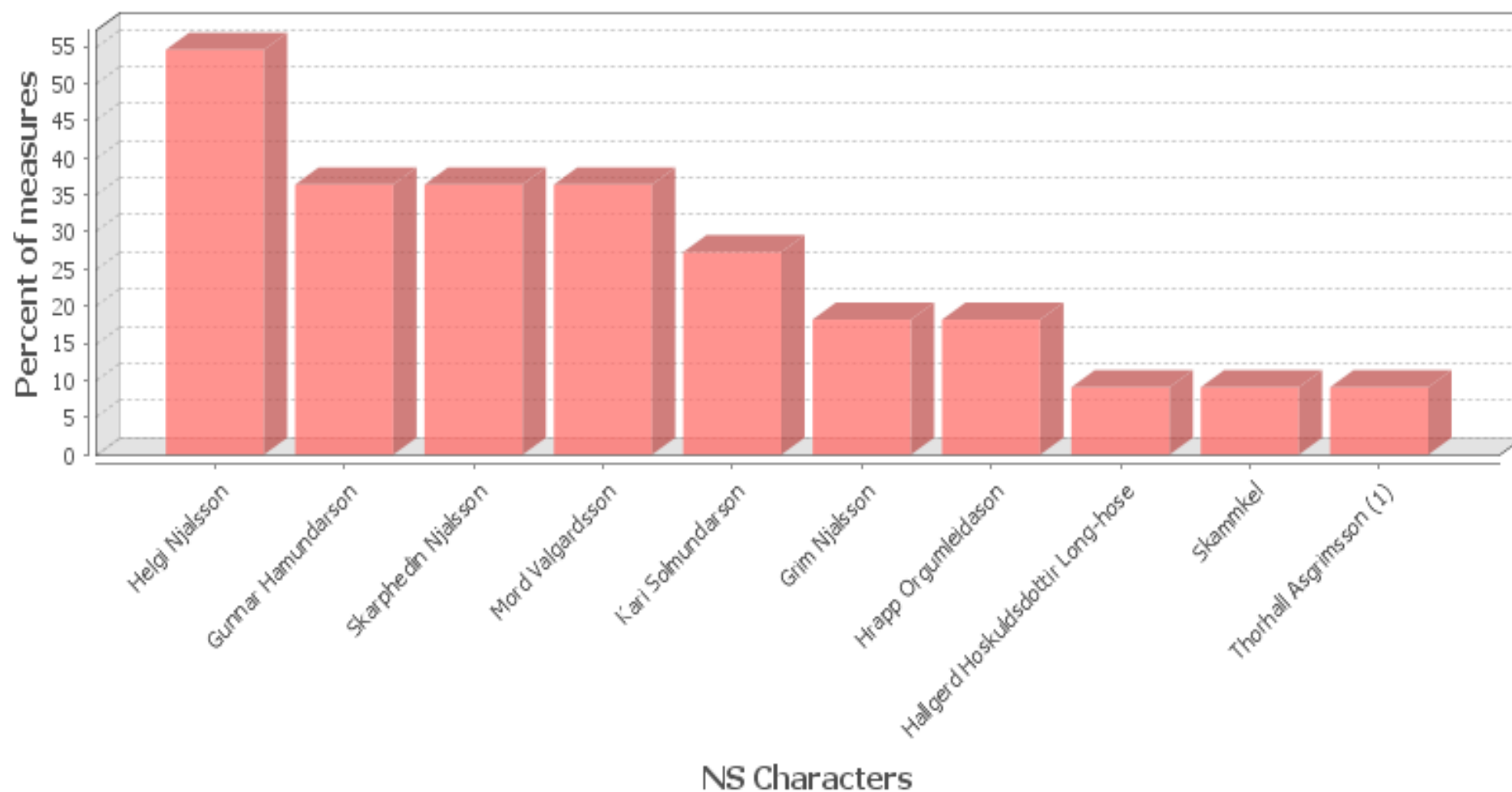


# Key Entities in Njáls saga

- SNA allows us to identify key entities along the network
- Need to make the analysis more dynamic (currently running it on the “endstate” at the end of the saga)

# Key Friendly and Unfriendly Entities

Recurring Top Ranked NS Characters - Unfriendly Interaction



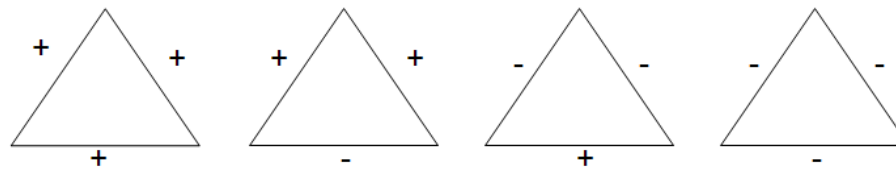


# Network analysis and visualization

- R, Matlab
- Sci2
  - Provides excellent data interchange
  - Excellent network analysis options
  - Port to Cytoscape and Gephi
- Cytoscape
  - Open source
- Gephi
  - Go to for network graph visualization
- Casos / ORA
  - Very easy to use and clearly explained network analysis
  - Only runs on Windows machines



# Structural Balance theory



T3

T2

T1

T0

Davis 1967, 185-186; Bott 1957; Rapoport 1963



# Egil's visit to Athelstan

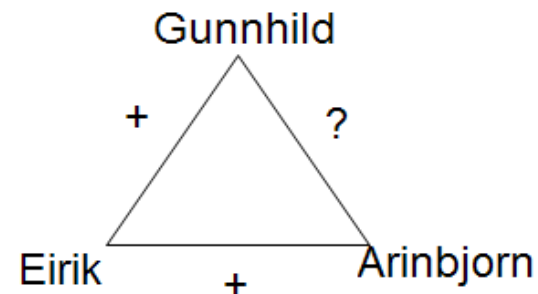
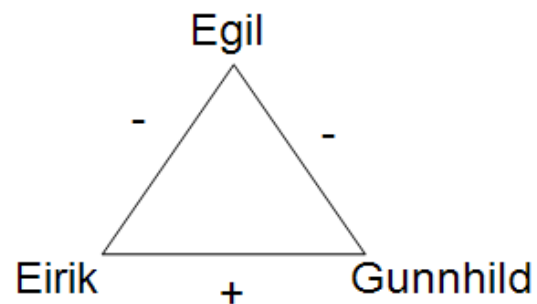
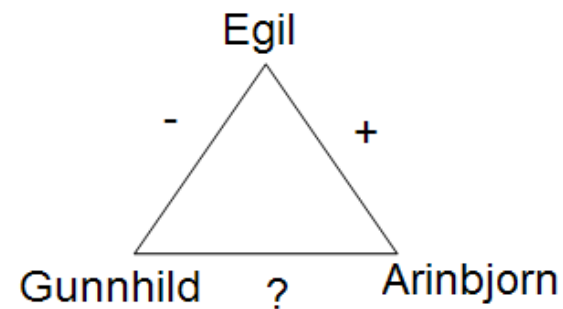
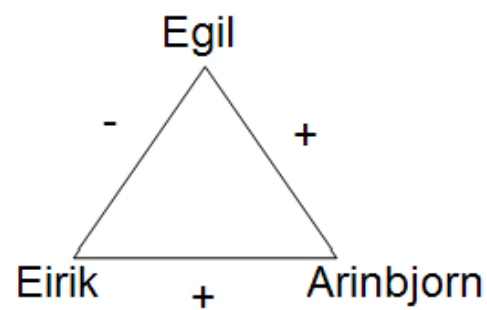


fig. 3: Status of friendship / enmity triangles at the beginning of Chapter 60

# After the Höfuðlausn...

o

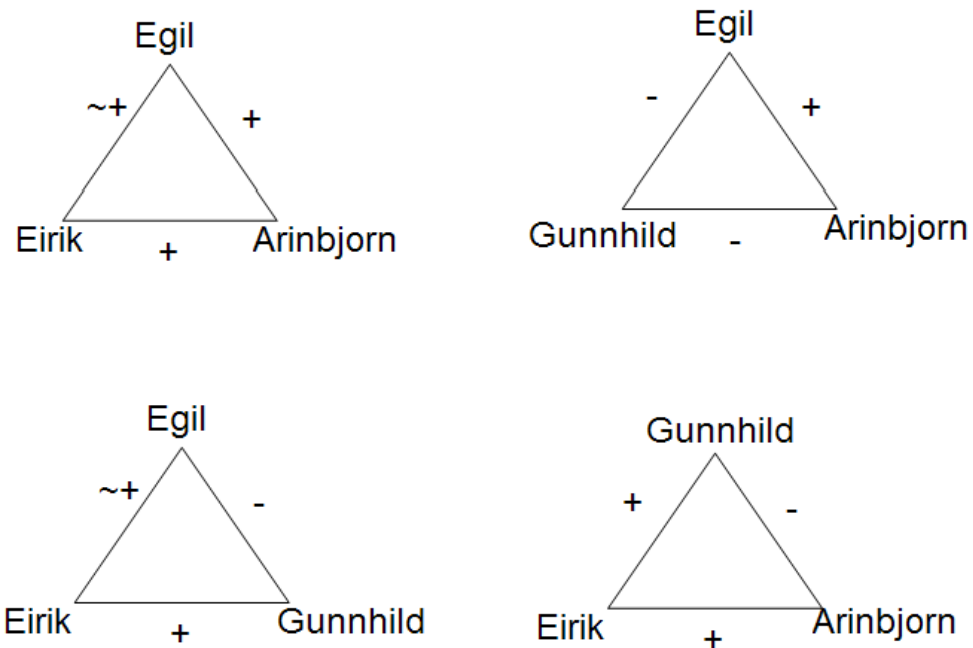
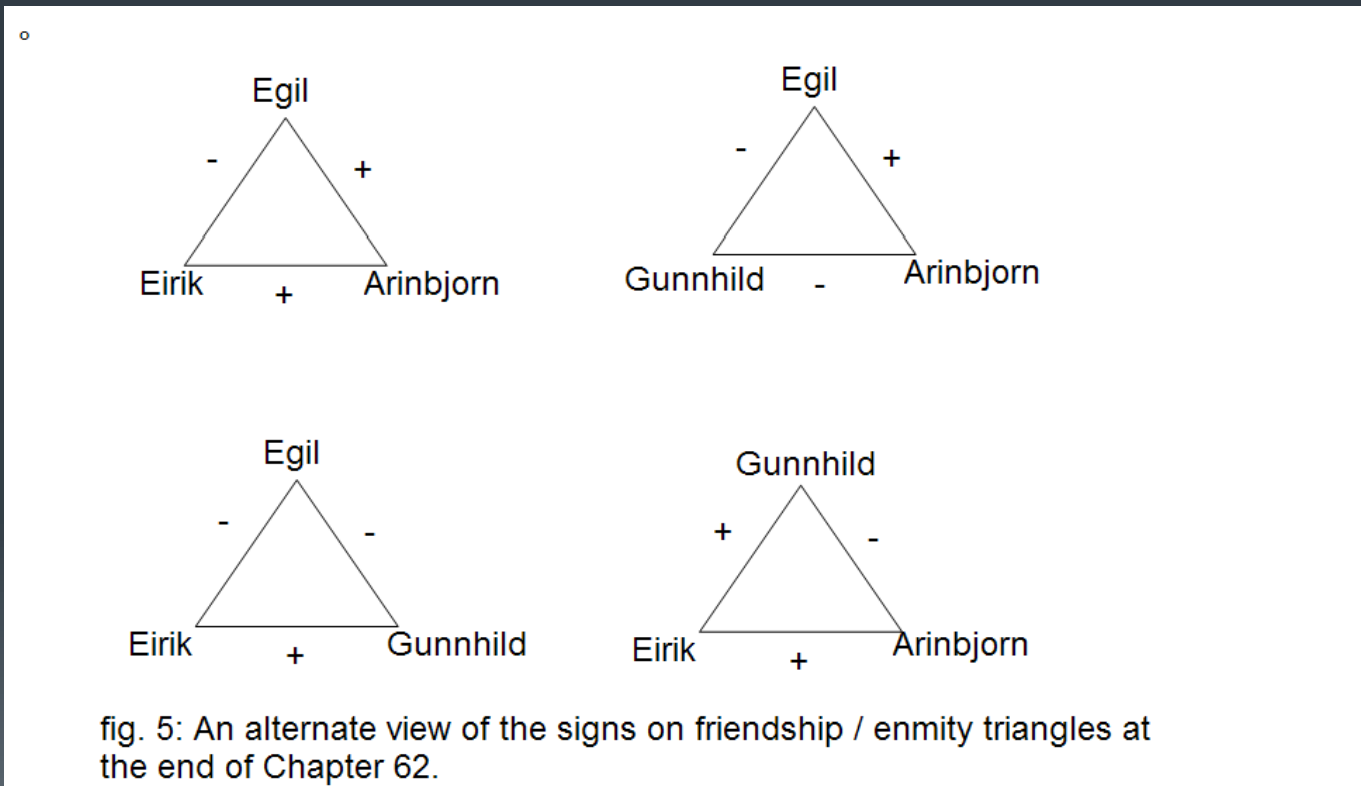


fig. 4: Status of friendship / enmity triangles at the end of Chapter 62.  
~+ indicates a temporary positively signed relationship.

# Another possible solution



# Status theory

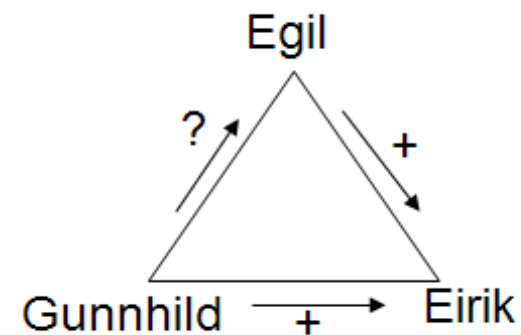
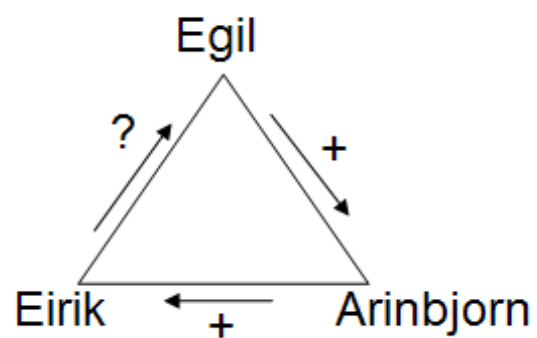


fig. 6: Triads signed according to status at the start of chapter 60.

# Keeping one's head

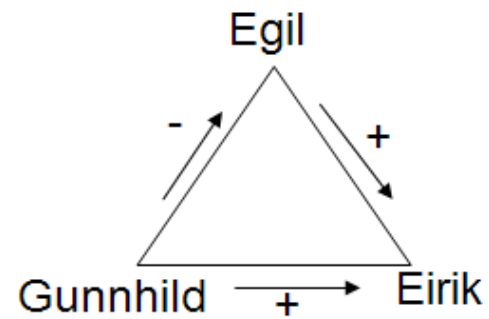
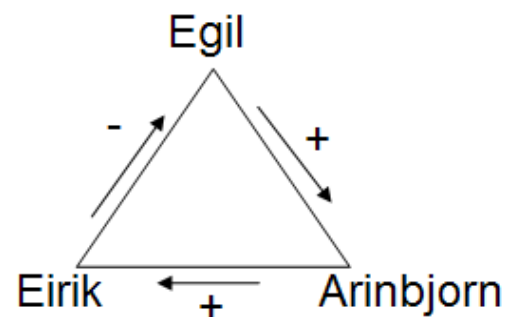


fig. 7: Triads signed according to status at the start of chapter 62.