

so hello everyone um i think i'm on the other sort of edge of the spectrum of the the familiarity with network analysis because i'm just a beginner and this talk is mostly about exploration of network analysis as a tool that is useful and perhaps feasible perhaps not for the study of medieval annotations so it's a sort of experiment i didn't get very far still work in progress as you will see and i have lots of room from for your comments and recommendations i'll be very glad to hear what you think about it and what are your advice how to go further at the beginning let me thank to my colleague peter boat who is also participating uh from halcon's institute uh who helped me tremendously with encoding all of the material you're going to see into xml and he was also instrumental in getting it out through his excess service artery so all the data you're going to see in a form of autographs and again visualizations are due to his his skill so thank you for that peter um early medieval annotations um are a very specific kind of medieval text uh uh they differ from standard texts that you were used to in the in that that they have what may be considered low textual coherence uh in standard text we are used to the fact that a text has a certain structure a certain order of words and if the order is disturbed or if it's not there if the text has different order of words we recognize as a different text but it is not entirely true for glosses and annotations

annotations in general tend to be behave  
as  
self self  
coherent entities each annotation on its  
own even though they are transmitted in  
layers or  
assembled into larger corpora of  
hundreds or thousands of such  
annotations  
and especially some types of annotations  
those that will be referring  
here to here as glosses are sort of very  
low on that spectrum and that they are  
really behaving  
as separate lego blocks where the early  
medieval annotators  
basically built each layer of classes to  
their liking so  
it's very difficult to see how relate  
to manuscripts containing various  
annotations how related they are  
and whether perhaps they contain exactly  
the same or very similar material  
because it's for us very difficult to  
analyze that by some of our modern by  
our standard methods  
by our naked eye and this was exactly  
why i  
ran out and tried to use it to  
disentangle this problem and  
find some alternative ways how to  
approach this material uh and how to  
work with this  
or work with it um in my particular case  
in my project we're talking about  
glosses to the etymologia of  
will this text is the most important  
latin encyclopedia of the middle ages  
it was extremely important particularly  
in the earlier period of the middle ages  
when it became an important receptacle  
and vehicle of innovations it  
underwent the process which can be  
likened to wickedization and one of  
those innovations that concerned this  
text was a

annotation uh these texts began to be annotated already very early after its uh disseminations already in the seventh century in several different ways uh but i will be talking here specifically about one type of annotations what i will be calling grammar glosses that is the kind of annotations are so very low on that spectrum of textual coherence they are not very coherent in other words grammar glosses are a phenomenon very particularly specific context and that is the carolingian period they are in fact a product of a very specific innovation from carolingian environment and that is a separation of a one part of the encyclopedia the part dedicated to grammar and its separate circulation uh as a school text and this is something that happened in carolingian period and we see that the glosses come into being and begin to circulate and be transmitted in these in the manuscripts connected with this school version of the encyclopedia and they're very closely related in terms of geographical scope chronology and materiality so for example this school version was particularly popular in northern france and this is also where we see most of these grammar glosses to appear unlike in the case of other types of annotations from the early middle ages these grammar glosses for which we do have substantial evidence for their diffusion uh it's very clear that they're circulated it's very clearly ever transmitted but it's very difficult

to see how exactly it happened and it is not possible in any way to treat them in sort of a stematic fashion to reconstruct the estimate because that's not really what's happening and we're likely going to see that a particular gloss recurs in multiple manuscripts and another class occurs in another group of manuscripts it's very difficult to see how these two or how other glosses connect together um this may have to do with the fact that schools were an important channel of transmission for this material but there were probably also other kinds of channels and networks that were involved probably they were also copied in scriptoria and one of the questions i would like to try to answer to use of network analysis is to see whether we can see some tangible evidence through the network analysis of which kinds of channels were used for the transmission uh for this purpose it's important to distinguish two different types of manuscripts that will feature also in the uh the graphs and the analysis that will come i will color code them pink and blue so pink will be always grammatical handbooks and blue will be always encyclopedic copies so these are two forms two types of manuscript early medieval manuscripts in which we find these grammar glosses to be etymologia the grammatical handbooks can be connected to the schools and the instruction in the classroom there are the manuscripts which contain the oldest layers of glosses also two of the most important manuscripts are

grammatical handbooks with most classes  
but but but at the same time there's as  
many manuscripts  
of the grammatical handbook type as of  
encyclopedic copy type that contain  
glasses so that's quite  
interesting and the question is what  
does it mean for the transmission it  
we can be reasonably sure that the  
glasses emerge through and because of  
the instruction  
in classroom they're sort of a tool for  
the instruction that's how they begin  
but their life becomes much more  
complicated and their transmission is  
not only governed by  
schools and the needs of schools so let  
me tell you a little bit about my corpus  
first about the manuscripts we're  
talking about these will be used as  
notes in the network as you  
as you can guess perhaps we're talking  
here about 61 manuscripts so this is  
quite smaller set as you will guess also  
that as to say  
as the data says we have seen before  
maybe that's also a bit of a problem but  
we are talking here about  
61 manuscripts  
they can be distinguished into two  
categories on one hand we have a core  
consisting of 16 manuscripts that  
contain at least 40 classes  
up to 770. you can see them here in this  
chart uh with roughly the number of  
glasses they contain  
uh and then there are 45 additional  
manuscripts which contain less than 40  
glasses  
these glasses come from the same pool of  
material but they are much  
less represented it is clear they are  
somehow related to the core  
but it's not entirely how uh some of  
them are also in destroy the ones that  
contain the most uh

most glosses and they are actually represented by this kind of uh striped color uh as you can see from the color coding uh there are some grammatical handbooks and there are also enthalpy copies and they appear in both groups the manuscripts come from 8th to 10th century but they were annotated in the 9th to 11th century so slightly later after when they come into being and most of them come from northern france but there are also some manuscripts from england and brittany northern italy perhaps and one manuscript from germany as for the glosses the entire corpus of the grammatical gloss grammar glosses to the etymology consists of for almost 4 000 glosses to the entire work the work has 44 chapters of together about 13 000 words so that you can get an idea how heavily or not heavy gloss it is i will call this a maximum corpus this is the case when we count every single instance of glossing separately we can also look at the what is minimum corpus if we count only the unique instances of glosses so no repetitions and this corpus is in about 2 500 unique gloss pairs in this what follows i will be using a smaller set because i have not completed my analysis yet so i will be using only a maximum corpus of around three which corresponds to the first 37 chapters because i have not completed encoding yet so it will be smaller about three quarters of the material are there so this is about

slightly smaller than what it will be in the end  
these glosses uh can be distinguished it's the two types of glasses on one hand isolated glasses glasses that appear only in a single manuscript and on the other hand what are we calling shared glasses that these glasses that appear at least in two manuscripts and more manuscripts uh there are 1216 glasses in the maximum corpus or it could be 1511 in the minimum corpus so about a one fifth of the minimum corpus is our shared glasses and this is what the material will be uh i will be working with um oops okay oh yes okay here's one example of how it looks when you see a shared gloss so you would see the same gloss appearing across multiple manuscripts one big question is what does it mean if you see a glass is exactly the same glass appearing in multiple manuscripts is the source of relationship and what is the nature of the thing with the glasses and this is very particular to this material is that they can emerge for the emerge spontaneously and when they do uh it is possible that they will look similarly or alike because if you're talking about situation in a classroom and you have one particular text and the student asks what does this word mean uh then the teacher there's a likelihood that the teacher will come up independently with a similar very similar answer and for this reason i engage in sort of what i call non-triviality ranking so i decided to rank the glasses give them a rank from one to four based

on how trivial or non-trivial they are  
uh courses with high rankings are the  
ones that cannot be explained  
independently  
and therefore point to some form of  
transmission circulation  
uh glasses with low ranking one and two  
are the ones which are likely to  
or may occur independently and  
especially glasses with ranking one may  
be independent on the contrary glasses  
with number four cannot be explained in  
this way  
and therefore it is very obvious that  
they are transmitted and they are heavy  
there are these uh these ranks will be  
used as weights on the glasses  
so they'll be adding weights to the  
material so what can we do or what i'm  
going to do is  
very similar to what other people have  
already done we can start by  
looking at the relation you can see  
which classes appear in which manuscript  
and use it as a source of relationship  
building them both as known as a sort of  
bipartite network uh because it will be  
quite a large network and not very  
comprehensive we can immediately  
simplify it instead of individual  
glasses we can look at gloss clusters  
a cluster is a group of glasses that  
always travels together or always  
appears in manuscripts together  
and we'll see there are fewer clusters  
and in the end we can progress one step  
further and instead of  
doing a bipartite network we can go we  
can just use these clusters as edges  
and give them weight to see the  
relationship between manuscripts so  
what you're going to see is a network  
representation of  
relationships between manuscripts it is  
very important we're not going to talk  
about or



not directly see the transmission pattern we don't know it's very difficult to reconstruct and one reason for that is that we should assume for this type of material that 95 of the manuscripts containing glasses are gone we're looking into into a very sort of very fragmented ecosystem uh where the question is how can we get from this fragmentation into sort of a picture nevertheless uh it needs to be also added many of the sort of manuscripts in which with transmitting losses were transient manuscripts because they were classroom material uh they were there to this they were they disappeared within one or two generations so this is this is very clear that there is lots of fragmentation so let's start at the beginning the very first step if we plot a top 33 manuscripts on one side and all 511 unique shared glasses on the other side with their labels this is the kind of bipartite representation you can get we can immediately simplify it because this is quite difficult so we will instead we can look at which classes always travel together and or appear together and say this is going to be a cropping group thing of some sort and instead of 511 unique shirt glasses we can get 275 groupings of losses this is roughly how they are distributed when it comes to how many glasses appear in such groupings you can see there are one or two very interesting groupings quite a few interesting ones and there are many many many ones that are

actually very unimportant  
in fact there are 115 of these 175  
groupings  
consist of only a single glass and  
moreover many of them have a very low  
nontriviality ranking so we can  
immediately forget about them  
these are not real clusters these are  
sort of a noise in the background which  
is likely to emerge in such  
in this sort of situation there are 20  
more one more groupings which consist of  
two or three glasses with very low  
rankings and we can also forget  
about them and in this way we end up  
with a smaller set of  
39 groupings which can be considered  
clusters  
they have a weight of six or more weight  
in this case meaning  
uh the rank of the glass is non-trivial  
non-triviality  
rank combined together and then you can  
see  
how many glasses they consist of on one  
side and also what is the combined  
weight of these glasses here  
there are 13 let's say important  
clusters which can  
which have a weight of at least 20 and  
there are two  
top clusters which uh stand out from  
among the rest as being extremely  
clustered that occurs in two manuscripts  
harley3941 and early on 296  
uh consists of 76 glasses with a total  
weight of 199  
another one is a a cluster that occurs  
in manuscripts only on 296  
and vlo41 which has consists of 29  
glasses and has a  
weight of 55 so slightly lower  
i will return back to the clusters and  
sort of because there's a lot more  
interesting to say about them but let me  
first draw try to do the third step and

draw you a nice sort of network graph  
where we use these  
clusters as edges and when the nodes are  
the manuscripts  
the nodes are colored according to  
schema introduced so pink  
is for grammatical handbooks blue is for  
insecurity copies the edges are also  
called a colored  
along the same principle so a pink edge  
is an edge between two grammatical  
handbooks  
a blue edge is an edge between two  
encyclopedic copies  
and sort of a purple edge goes from one  
type to another  
this network the representation was laid  
out  
according to the rough geographical or  
point of origin of the individual  
manuscripts in many cases this is also  
where they were annotated  
so what you see here this is northern  
france uh with  
vlo41 down there that is flurry an  
important monastery with a school  
especially active in  
10th century beginning of the 11th  
century rhymes four to six  
that's rhymes actually a seed of the  
archbishop and also see  
a locals of a very important cathedral  
school and monasteries as well  
and this sort of um cluster here  
these are manuscripts which are  
connected to paris because they were  
produced there  
or annotated there and some of them were  
both produced and annotated there  
especially in particular early on 296  
a manuscript which has most glosses and  
which has also two very strong  
connections  
relations to other manuscripts um here  
up here we see england  
in the manuscript paris 7585 which was

produced and annotated in canterbury  
this is brittany and this is northern  
italy probably  
in case of these masks we cannot be sure  
these are definitely not an  
italian so you already see some very  
interesting things here for example you  
can  
hopefully see that these manuscripts  
from paris are they do not have  
uh strong strong connections between  
each other or as strong connection  
between each other  
as they have to manuscripts from other  
places in case of northern france or  
other regions actually  
that's quite interesting and you can  
also notice that uh  
the grammatical handbooks come from this  
area and here we have  
encyclopedic copies but let me get but  
get back i hope i still have time  
to some of the facts about the clusters  
interesting things  
so i said there are 30 to nine clusters  
that can be identified and we can  
reasonably assume are actually real  
clusters not just a noise or just  
some aggregation of glasses and what is  
interesting about them is most of these  
clusters  
represent relationship between it is  
unique to two manuscripts  
so we see a group of glasses that  
appears recurs in two manuscripts only  
and not not less but not more sorry  
there are some examples of  
uh clusters that occur across three  
manuscripts but  
very fewer way fewer in this group and  
in this the top 32  
nine and they're only one there's only  
one cluster  
of only two glasses moreover but this is  
shared between five manuscripts and only  
one cluster

that shows relationship between six manuscripts uh  
i don't know exactly why is that this may have have to do with the level of fragmentation of the whole sort of original ecosystem that we are looking actually at uh  
small bits and pieces of what would be sort of uh clusters which would be sort of  
of  
shared more broadly perhaps but we do not see that because many of the manuscripts are gone  
but it also may suggest sort of a low transmission rate for the the grammar glasses  
and it is even possible in fact that these clusters are very transient in nature that actually a cluster comes into being  
because of transmission needs so there is a circle  
or a school or an institution that is sending material to another school an institute or institution  
that creates a cluster for this purpose the cluster is transmitted  
embedded into the different manuscript on the other side of manuscripts on the other side  
and then it sort of dissipates and if that second institution or circle passes material on  
it creates a rather a new cluster than passing on the material that was already available  
available to it another interesting thing about the clusters is that actually they come in two different varieties and that becomes quite obvious after sort of a bit of analysis  
we have what may be called visible and invisible clusters visible clusters are of the kind that sort of good philologists would also notice and they can be seen from collation of the

manuscript  
and the main reason for this is that is  
that they concern  
uh smaller segments of text they appear  
close to each other  
they are not interspersed by glasses  
from other clusters or isolated glasses  
and they usually contain a smaller  
section such as one or two chapters  
but it's also interesting about them  
these are these clusters tend to have  
relatively high average ranking of about  
2.5  
and many of them contain a high number  
of non-trivial glasses with ranks  
three and four here you see the averages  
and here you see  
the six uh visible clusters i was able  
to identify  
you can also notice that they all uh you  
perhaps could not notice this but these  
are  
many of them are manuscripts of the  
encyclopedic type  
on the other hand we have what may be  
considered invisible clusters and these  
are these are ones i  
did not see before i performed this sort  
of analysis  
this quantitative analysis  
because they came out of this exercise  
so to say it is because the glasses that  
constitute these clusters do not find do  
not appear next to each other  
they are already distributed across the  
antarctics or most of it  
not close to each other interspersed by  
other glasses whether shared or isolated  
and also it is interesting they tend to  
have relatively low average rank  
of under two or two or less and some of  
them are  
made only of glosses with ranks of one  
or two for example this cluster  
which appears which was the second most  
important cluster in the whole corpus

actually is constituted only with  
glasses of rank one  
and two and also you can notice  
or maybe you don't but many of the minus  
in most of the manuscripts here  
are of the grammatical type of the  
grammatical handbook type  
if you plot it here you can see that two  
the yellows are here are visible  
clusters and green are invisible  
clusters and you can also see that  
that the yellow ones appear between  
manuscripts of the encyclopedic copies  
and the green are either between  
grammatical manuscripts or my  
or in one case rhymes 46 is also quite  
interesting because it contains many of  
these invisible clusters as well  
even though it is an in an encyclopedic  
copy  
and there is a good reason for it a  
small detail here i hope at least those  
who can read latin can see it here it is  
because this rhymes manuscript contains  
uh the oldest layer of glossing there  
are stylist glosses that these glosses  
they were not put in by ink  
but by uh lead stylus uh this is  
the stylus and this is the instrument  
with which it is used that is back  
tablet  
so they're probably carried over from  
avax tablet this is extremely unique  
this is the  
in my corpus and i have not seen it  
elsewhere yet it is in latin context  
box tablets were used in schools and  
it's possible that what we are looking  
at here is that this enthalpy copy was  
annotated directly from material that  
comes  
to this transient medium from school  
environments  
i think i should end here so i just want  
to reiterate uh  
these visible and invisible clusters

that was the most surprising thing i've seen  
it's very interesting and one are more aligned with encyclopedic copies and the clustering the pattern of clustering is more consistent with copying rather than with let's say instruction in classroom  
and therefore may be indicative of the work of scriptoria on one hand while the invisible clusters  
uh which are more common in grammatical handbooks and also show other patterns for example low ranks of losses may have to do with oral and mnemonic transmission  
maybe more more like more connected with schools and the transmission via school  
uh this is where i want to end for now so this is what where the analysis so far brought me as you can see i didn't get very far and i would like to get further  
uh i it was it helped me to identify several interesting clusters of glosses including such they would not be able to see otherwise it helped me to see differences between two types of clusters those that i call visible  
and invisible and also have a sense many maybe  
looking at this visible and invisible cluster it's looking at two different channels of transmission copying in scriptorium on one hand and transmission is called environment in on the other hand yes this is what i will stop for now because i think i'm already speaking for too long and thank you for your attention  
thank you avina i think you're perfectly in time we  
even we would have even still have some some more time so we have plenty of time for questions and discussions um



it was a very interesting presentation  
and i'm actually impressed  
by the amount of them have  
so i don't think that it's a small  
corpus to be honest  
so um other questions remarks  
sorry  
can i ask the question okay yeah yeah  
then uh yeah  
it's uh me i have my camera turned off  
um hello uh i just wanted to ask  
about something at the very beginning  
where you were talking about the  
types of glosses and you said there are  
three types  
scholarly the insular and the  
grammatical glosses and i was just  
wondering  
what  
boss