

great so um hi um
thank you uh evina and gustavo for
putting together a wonderful conference
we're very pleased that we've already
learned a lot
um so i'm going to start by telling you
a little bit
about a manuscript uh currently held
in uh cambridge university library
under the shelfmark additional
manuscript 3479
and this manuscript is quite an
unassuming
little codex a little over 200 leaves
measuring 28.5 by 22.5 centimeters
uh there are several scribes writing in
a caroline miniscule script
no decorations apart from some plain red
initials
and in the catalogue it is described as
a compilation of patristic
texts and that description is exactly
right
um we find among its contents some
letters
a number of lectionas and then
predominantly
sermons written by or ascribed to the
church fathers
the exact origin of the manuscript is
uncertain
but is likely to be eastern france
because it contains a collection of
sermons
with limited dissemination that is known
to have been present in lyon
in the beginning of the 9th century when
this manuscript was also created
now confronted with this particular
opening to a presentation
a rather reticent description
of a relatively anonymous manuscript
you might expect that the rest of the
presentation would focus
on a heretofore unnoticed but
spectacular feature of this manuscript

perhaps an unidentified original text or
a famous medieval user but
that is decidedly not our point today
the simple fact is that manuscript
libraries across the world
are full of unassuming manuscripts
not especially remarkable in terms of
subjects decorations or history
though potentially relevant for use in
critical editions of the text they
contain
apart from the odd editor such
manuscripts are unlikely to receive
sustained attention from scholars
this is unfortunate because every
manuscript where the plain or lavish
old or young meticulously copied or
haphazardly thrown together
represents an act of reception a very
small cog in a very large medieval
machine
that toiled tirelessly to disperse
knowledge
and that has unalterably changed our
collective literary and intellectual
heritage
the problem is of course that the
reconstruction of this medieval machine
of textual transmission
is extremely challenging and that
getting an accurate
picture of each individual manuscript
as a motor of reception is to a certain
extent
contingent upon the overarching
reconstruction of that entire machine
so that means that in practice we are
still focusing to a large extent
on manuscripts that have been identified
as significant
a frustratingly vague term by previous
generations of philologists
and that we are leaving large parts of
the corpus of medieval manuscripts
untouched still much progress has been
made

since the digital turn in manuscript studies
and in this presentation we want to present part of an ongoing project that hopes to offer a small contribution to this larger effort before we delve into the specifics we want to contextualize it a bit by quickly presenting the project in general so patristic sermons in the middle ages or passing is a project that started in january 2019 and will run for five years it aims at charting and analyzing the interrelations between the medieval manuscripts that transmit latin patristic sermons that means the sermons preached by the fathers of the early church augustine gregory leo and their contemporaries these texts had a dynamic and complex medieval afterlife throughout the middle ages they circulated usually as part of larger collections of patristic preaching both the collections and the sermons themselves were heavily manipulated as part of their medieval reception additionally they circulated alongside and became intertwined with an enormous corpus of pseudo-epigraphic sermons so texts attributed to one of the authoritative fathers but of which the origins whether late antique or medieval are often uncertain now at the heart of our project is a digital research tool that consists of a database of metadata on manuscripts that transmit patristic sermons

and an interface designed to allow for sophisticated querying of this database in due course this digital tool intends to grant access to a greater number of relevant manuscripts than has yet to fall been studied it is also meant to create a semblance of order in the very extensive and messy scholarly tradition of identifying futuristic sermons and their medieval incarnations and in doing so the project strives to open up new research avenues for the study of the reception of latin futuristic preaching but simultaneously and perhaps more relevant for this conference pasim also intends to contribute to the methodological framework and digital toolbox necessary to further embed manuscripts in the study of medieval history so in this contribution we present one of the building blocks of the passing research tool that we are just starting to develop namely network visualizations of textual overlap between sermons so what we will describe is experimental and not yet fully worked out but that means that there is a lot of room for input and improvement patristic sermons are a very malleable genre they invited myriad adaptations abbreviations expansions cut and paste jobs this characteristic of sermons is particularly interesting from the point of view of reception studies these deliberate adaptations are generally intended to make the salmon fit in a particular context or adhere to a certain pattern however the abundance of different

versions
and textual overlap can be intensely
frustrating
when trying to rigidly categorize and
identify the sermons as you find them in
the manuscripts
passim is trying to solve this problem
by creating
an up-to-date set of authority records
which reflect the current comunis
on each sermon's author and critical
text
each manifestation of a sermon as it is
found in a manuscript with all its
scribal errors and different author
attributions
is linked to the corresponding authority
record
if a particular sermon circulated in
multiple distinct versions
we also create multiple authority
records for it
these authority records are then linked
to each other using one of a set of
different link
types uh in a sense this is our version
of even
triviality ranking that she presented
yesterday a way
to add more information to the links
the simplest link type is the partially
equals link which means that there is
some textual overlap
between two authority records
depending on the amount of information
offered by the reference
works we consult the link type can be
further specified in two ways
if we know more about the nature of the
overlap
we can specify that using for example
this sermon is a different
version of that sermon or this sermon
is made up of a combination of this
sermon and another one
if we have access to stematological

information
the link can also be given a direction
it can become an
arrow and the type can be refined to
this sermon is a model for that sermon
or this sermon was modeled after this
salmon
by connecting the authority records in
this manner we are
effectively creating a network of
patristic
and pseudo-epigraphic futuristic sermons
that reveals
how these sermons changed over the
course of their medieval transmission
now the next step is to make this
network
visually accessible to users of the
passim
research tool and to explore the
challenges
inherent in creating a serviceable
network
of a sermon's textual relations we will
use as a
text a test case sorry the
pseudo-augustinian
sermon 121 which is a christmas sermon
attributed to augustine but not written
by him
on the topic of the version birth
now this sermon is an apt example
as you will come to see in a minute of
the complex
interrelations that can exist between
texts
labeled as patristic sermons
previous research has shown that this
sermon is in fact
a combination of two separate texts
121a and 121b which
had vastly different manuscript
traditions
so starting with 121a this one is found
in the homiliary collections associated
with

fleury as well as two 10th and 11th century manuscripts currently in wolfenbuttel and monte cassino as well as the unassuming cambridge manuscript with which we started this presentation this 121a appears in these manuscripts as part of a cluster of christmas sermons it has been connected to various pseudo-ambrosian sources most notably the fragment in natali domini in addition to that it was itself used as a source for various other compilations it is notable finally that 121a was probably not a part of the old roman homiliary which formed the foundation of a large number of early medieval liturgical collections moving on to 121b this text is connected to the important roman humiliary it is also found in the 19th century manuscript montpellier 59 which incidentally also transmits its sibling 121a a different slightly different version of 121b with an alternative ending associated in the reference works with numerous other authors including pseudo-peter cressologus is found in the roman homiliary and other widespread homiliaries dependent upon this collection finally both this version of 121b and 121a probably rely on a shared principle source an unedited sermon found in four french manuscripts labeled as pseudo-augustine cermo caillou 10 a now this is a quick but representative overview of the current consensus on this particular salmon and if you are

at this point starting to get a little
bit lost frankly we applaud you for
sticking with us this long
already um let us immediately
hone in on the two main problems as we
see them first
what we have just described is the
textual tradition of a sermon
that has been the subject already of
several scholarly articles
the manuscript tradition is littered
literally littered with texts
that have not yet enjoyed this privilege
but that may be part
of networks just as complex as
pseudo-augustine 121
second problem this type of philological
research
research should be but hardly ever is
taken to the next level
and to the left next level in two ways
first the analysis should go deeper
and ask how all of these alterations and
combinations
actually affect the content the literary
qualities and the theological
position of the texts in question
second the analysis should also broaden
in scope
and look at the implications of this
particular case study
for our knowledge of for instance early
medieval
traditions of preaching transmission and
authority
and in the remainder of this paper we
will see whether a network visualization
can contribute
to solving these two problems and for
that i give the
floor to iris and we are quickly
switching chairs so she can control also
the power points
all right hello that was my cue um
first up before looking at the
visualizations themselves

we should explain how we made them in this experimental stage of the project all network visualizations were made using the program node excel which is a very simple program which is regularly used for social media network analysis eventually a visualization generator will be integrated in the passing research tool but for now this will do the layout of these visualizations is governed by the so-called cool and fast multiscale algorithm which is a type of algorithm designed to make visualizations aesthetically pleasing for instance by keeping all edges around the same length the networks should be understood as photos each node represents an authority record which is a standard text as shari has just explained in our database each of these authority records will have its own passing code and be linked to all manuscript manifestations of this text in the database every edge in turn represents textual overlap any text will overlap between two sermons our knowledge of this overlap is based on both existing publications and our own research we start with a simple network of the text mentioned in the previous section these connections are all quite easy to glean from the reference works we might call this a first degree network or a simple network you can find basically all of these connections in a simple in a single paragraph on pseudo-augustine seven 121. still

as noted this particular sermon may be exceptional in this regard the quantity of research on particular particularly pseudographic sermons can vary considerably the network becomes vastly more complex and expansive if we follow it to its limits this means that we've added the additional textual links to each of the notes until there are no more textual links to add this network shows that far more edges connect sermon 121 a and b to other nodes in the network than a first degree reading of the reference works reveals if you are interested in this particular sermon it can give you a quick overview in a single glance of all the texts or authority records you should look at in order to get a full picture of the models alternative versions and offspring of this sermon however that is not the only and not the most important result to be gotten from this network let's assume you're a researcher entering this network not starting from a particular sermon but with a general question about practices of rewriting and compiling futuristic materials throughout the middle ages this network can show you the most promising text to look into you see of course the network does not guarantee that all of these options give you exactly what you're looking for for example the multiple collections going to and from the cluster around the anonymous sermon on slo1 are caused by the fact that many of them feature a very popular opening sentence which was picked up in many different

segments

however this network does also have a distinct advantage

you do not have to start from a known text to find what you're looking for in this sense the network visualization has the potential to help with the problems discussed above

it incorporates more material outside of the few sermons picked up by scholarship and it allows you to tentatively

identify those texts

that are worth getting into depending on the questions that fuel your analysis

also very interesting about this network

is the fact that it turns the

traditional perspective upside down

the eye is not drawn to the originals

the authentic sermons by the church

fathers these are mostly lingering at

the outskirts of the network as you can

see

whereas your eye is drawn to the text

within most edges

that connects most other texts either

the most important source

text or the cantonese that incorporate

many of such sources

these include the traditionally

problematic or unstable text

that symbolize the medieval attitudes to

the patristic source material best

so as an exercise in shifting and

broadening one's perspective

it is quite successful still

the network at this stage has several

deficiencies as well

in one way it lacks neurons connections

are undirected

and unspecified in that sources are not

distinguished from nagleben

and there is also notice difference

between two texts

that allude to each other and two texts

that are virtually

the same except for a few small

differences
at the same time the network may be
overly nuanced
and that it shows us all textual
connections expressed in the scholarly
literature and reference works
whether or not they are still upheld in
the case of
sermon 121 for example the contradictory
ideas of several scholars are at this
point presented without
preference for one over the other
because all
instances of textual overlap are
considered equal at this point
we are experimenting with ways to add
more detail to the network
for example by adding edge labels that
specify the link type
by adding arrows at some points to give
direction to the edges wherever we know
what that direction should be or by
playing with the transparency of the
edges to indicate
that an edge is based on uncertain or
unspecified
data all these implementations however
are only
are only possible if enough reliable
data are available
nevertheless for the final part of this
presentation we want to spend some time
exploring one type of further
development of the network
that would we hope significantly
increase its usefulness
one of the most important challenges in
general to overcome for the passing
research tool
is to bring together or to truly
integrate what we know about
textual connections with what we know
about manuscript traditions
in other words to link the abstract
intertextual relationships that we've
just

shown you to their material manuscript contexts
the clusters visible in the network illustrate the key premise that at the point in time when a new sermon was created all of its sources must have been physically present in some form or other
the more information we have about direction degree and manuscript addition the more certain such a physical connection may become
this ostensibly simple principle can be a gateway into bigger topics such as patterns of transmission
the reconstruction of lost collections or better insight into the strategies of selection and dissemination all of these are especially hard to determine for the early medieval period when as we have already heard yesterday manuscript evidence is of course limited
in this version of the network the size of the nodes corresponds to the number of manuscripts
in which the sermon is found the bigger the nodes the more
accident manuscripts each color in turn corresponds to a known collection and by collection we mean here
a combination of sermons which is either transmitted in multiple manuscripts or which can be traced back to antiquity
this development of our initial second degree network
has the potential to help the user in identifying texts that were transmitted together in the same collection for instance texts that were especially popular
texts that traveled in many different contexts
and that have for example lots of manuscripts but not all of them of the same collection
and finally collections that served as

key mediators or sources of inspiration
for the rewriting of texts
without jumping to any conclusions you
can already see in this first
experimental network
how a number of texts in the network
around sermon 121
were used as sources for compilations
found in the collection alarms of pharfa
which is the blue one
alongside the aforementioned
pseudo-peter kusalaga seven
five i don't know if i can point it to
you it's there
we also have pseudo elephants sermon
seven and
uh elephants sermon eight
that are quite prominent in the network
this way of breaching the manuscript and
textual traditions
could potentially prove fruitful ground
for new questions or starting hypotheses
however again there are a number of
important issues or caveats
one of them is simple to overcome in the
program we used we can assign only one
color per node
whereas many texts are part of multiple
collections
for the passing research tool the idea
is to convert the nodes into pie charts
so that all the connection collections
the sermon is found in can be
represented
the size of the slice would in this case
show
how many manuscripts there are per
collection
the second and third problematic issue
are
a bit harder to solve the corpus of
manuscripts in the passing database is
incomplete
and so for the purposes of this
demonstration
we have estimated the number of

manuscripts based on our reading of the reference works
however this state of incomplete list will probably persist
for quite some time and so the network will have to be used with caution
finally the network in this experimental phase also reveals another important problem
when looking at this network the fact immediately jumps at you
that the notes of the authentic texts are overall
much bigger than those of these pseudographic sermons
we do not think that we can interpret this as an accurate reflection of the manuscript edition
rather this reflects the fact that authentic sermons are much better studied
there are editions and reference lists which makes it much easier for catalogues of manuscripts to identify authentic sermons and so while it seems that authentic texts had a much larger dissemination in the middle ages
our intuition would be this is an effect of the scholarly tradition and not necessarily the historical reality
this example shows that we cannot take this network as face value without an acute awareness of the material it is based on
right as being mentioned at the start of this presentation
our purpose was to explore the potential usefulness of this type of network visualization
to get more revenue out of the data we are gathering in the passim database
in terms of potential edit value we feel the experiment was successful
we can conclude that looking at the data through a visualization

indeed succeeds in pointing the user to promising research avenues and case studies beyond the traditional boundaries of the existing scholarly tradition furthermore and this was a more unexpected result the visualization forces the user to adopt a different perspective that does away with the prioritization of the original that was prevalent in stematology instead this visualization explicitly draws the eye to the medieval reception which is exactly what passim intends moreover our effort to encapsulate both textual data and manuscripts data in a single visualization may provide interesting insights into the workings of transmission and textual adaptation the visualizations we've presented so far are relatively simple and we do intend if we can to add more options giving users the opportunity to customize the visualization depending on their interests more importantly still we want to make sure that the visualization does not form an end point but rather a bridge which means allowing the user to click on each node see the details of that authority record see the manuscript connected to that authority record and continue to explore the database from there this this very flexibility of the network visualization that will allow it we hope to be adapted to new insights on particular sermons and to evolve with new research within the field however it's not all good news we have also been confronted with some pretty

challenging obstacles
because we are just starting to build
this feature of the passing research
tool
we want to take full advantage of the
expertise gathered at this conference
so in conclusion we want to briefly
mention these challenges again
in the hope that helpful suggestions or
experiences may emerge from the
discussion
it all comes down to responsible use of
the network
which firstly will remain for a long
time incomplete
which is secondly generated by an
algorithm that
may invite a user to attribute meaning
to random effects
and which is finally subject to
distortions produced by its abundance
on previous traditions of scholarship we
look forward to your input
thank you
thank you very much wonderful paper
sorry logistical arrangements
yes i see that emo has a question
sorry yeah here i am thanks a lot that
was that was very
very very interesting and i have two
questions one in more context and one
more in detail
um both very superficial
questions i suppose uh first one is the
the manuscript context of those sermons
i i suspect
that it's not any not all of those
manuscripts
include only sermons so they will
potentially be bound within
a certain context whether that context
is and plays an important role in the
overall
transmission second one is for the text
themselves
how formulaic are they so if you dive

into those texts
can you create blocks so and and then
play around with those blocks where you
can see okay there are certain ideas for
those sermons
um that move around and you have we call
that in in
what i'm trying to do tomorrow objects
and you can then
try to create networks through objects
so i'm wondering how
how much you want to deal with the text
and if they can be divided up
in significant blocks and whether that
isn't
okay yes thank you uh wonderful
questions both of them
uh with regard to the manuscript context
so yes
sermons also travel a lot with other
types of
texts other genres um we've decided to
focus primarily on um
manuscripts that contain let's say
collections of sermons
rather than uh manuscripts where there
are one or two sermons
interspersed with lots of other
materials
that was mainly a feasibility choice but
also because
um we wanted to really explore the
context
where sermons were still sermons and not
theological treatises
or like adapted because as a genre
sermons are also not that easy to
delineate
they tend to shift a little bit over the
course of their transmission
but if you focus first on on collections
of sermons
then you keep that sort of preaching
context
a little bit intact and that's our first
starting point

for for this project knowing full well
that we will have to eventually expand
and also incorporate
other types of manuscript compilations
um second question about these objects
as you call them
well i would say that at this point
our first intuition is that no they are
not
especially formulaic but
i have to immediately add that
especially the corpus of
pseudo-epigraphic sermons is so little
studied
when it comes to its comp actual content
of these texts
that we cannot know that for sure yet so
that is actually one of the things that
we are hoping to
um explore further and discover because
you have
you so you start with the the text of
the church father some of which are
quite
like literary very beautiful pieces and
certainly not formulaic in any way
and then you come to a new tradition
where these texts are broken apart
or where new texts are created that sort
of
allude to them or try to incorporate
their style or ideas but
are completely different in terms of the
context in which they were generated and
the purposes they served
so it is it is that new um early
medieval context that we are very
interested
in exploring further because there's
very little know
known about that at this point
anyone else i think this
comes back to the question that even
asked
in the previous paper what is similarity
and what's the difference and how do you

rate it
and can you somehow visualize it too
in the model um
but these are all very difficult
difficult uh
questions to solve um i see that sarah
corrigan has a question
hi sorry um it well it follows on from
that same issue i suppose and you you've
partially answered it
in your answer to immo but
i was immediately um intrigued by
how so what's the minimum
unit of a sermon and which i suppose is
a bit limited if you're focusing on
collections of sermons then you likely
aren't dealing with very many
fragmentary texts or very
small bits and pieces but i was
wondering like at what point
do you say this is no longer part of the
sermon that we're tracing and and how
have you defined that even practically
for the point of the project
yeah that's a very uh complex and
difficult issue and
we are actually i showed you a bit of
the link types that we are putting in
place
um to show to like further define the
connection between two texts
and we do see uh quite often that some
like
first and second paragraph of a sermon
are combined with third and fourth
paragraph of another one that type of
thing we see quite often
but we are struggling ourselves with
creating
a representative spectrum of link types
that covers everything while keeping
the visualization or the network still
you know correct accurate um and that is
something that we're still quite
working on but you know we're kind of
giving ourselves

a way out of that by saying that you
always have to use this network
as a starting point to point you to the
places where you have to go and look
next um and then you know
that might turn your hypothesis upside
down or it might just turn out to be a
dead end
but because this corpus is so huge like
for pseudo-epigraphic sermons attributed
to augustine
there are over 3000 known sermons
of that type so we cannot investigate
all of them
but we're trying to find like the
minimum amount of information
we need in order to point out
the good places to look further that
that go beyond what
19th and 20th century philologists have
already explored
thank you i see that dominique also has
a question
yes thank you a small question um
you have said for the hooray uh research
project that we are
identifying texts so with text reuse
identification methods
have you applied things like tracer
from marco bischler to your characters
to the edited text
no not yet and we're actually i was
really liking
i'm going to email and ask you more
technical details about that but
no right now we are using some very good
um reference works uh and catalogues
um also because quite a lot we are
making a very clear distinction in the
database between the text
as it occurs in the manuscript and the
text as it is
identified like sort of the authority
record or the standard version
but um not all of them are edited
not all of them are properly identified

so it's really an ongoing process of yeah correcting previous scholarship and adding to it so yeah it might be that using machine-based recognition could really help with that thank you mateo uh yeah thank you thank you for the presentation i just have a small question maybe i missed it and but do you have the dates of these documents so or how precise is the range of that you can apply by dating these documents um yeah that's a a good question and quite an issue for our manuscript data um for these texts it's it's quite impossible because like the the for the pseudo-ethographic sermons we have sort of a range you have the first manuscript they occur in and then you have the source text and somewhere in between of those two this text came into being but usually the label is just like this may be late antique or this is probably somewhere early medieval and so there's not a lot the idea of creating a beautiful like timeline that shows you um how these texts changed would be nice but will probably only be possible for specifically delineated case studies and not for the entire corpus okay i think okay maybe uh just one uh a shorter a short suggestion uh in spite of this even if you keep your network undirected you could nevertheless try uh to build communities to identify communities uh by applying so community detection uh techniques which are quite standards using your overlaps data and it might be might be a different way

to explore your network to use your network for explorative reasons as you said
so just as a suggestion thank you and then there's a question from uh nicole vollmering who doesn't have a microphone but she asks on the same note will it eventually be possible to identify identical segments recurring in various sermons and recessions so that goes back to the question we had before of sarah yeah basically at this point we are using things like incubate explicit um to identify a sermon that that's not the whole story because that does you know uh offer it needs to be more nuanced than that but we are not doing full text comparisons at this point um but we are trying to create authority records for for example excerpts that had a separate life in the manuscript tradition so that's a first start for that type of um that type of identification but you know the database is perfectly capable of doing this we just don't have time at this point to do all of the full text comparisons yes and then there's a question from eskate573 i can't hear you sorry you you haven't unmuted oh you are still muted i'm sorry yes um it was very brilliant thank you very much um the first first remark i i don't know why you did begin your paper with the manuscripts of cambridge but uh it's it's a manuscripts um it's a manuscript from samuel actually i mean i

think it's
it's it has a next laborious so it
confirmed the um eastern
probably the eastern france province of
the manuscript
the second thing i wanted to to ask you
is um do you work
primarily with the catalogue of
manuscripts or from the literature
are you working um at the beginning with
freda and mahilson and all these
repertories
or are you working with catalogues of
manuscripts because the the result would
be very different if you are taking old
manuscripts from a military collection
you will probably have a lot of pieces
which are um unedited um
and ascribed and and and all that stuff
are you um oh are you dealing with this
text
are you cycling yourself this text or
whether with a new cycle
yes so yeah um
[Music]
the answer is a little bit of both so
well a little bit actually hopefully all
of both but
um so our way of importing
data is to use different um avenues
different sources for import so my
health
and grizzle are very important um
because they supply what is
the common uh identification codes for
the texts that are known have been and
have been identified
um but then once we identify a text in a
manuscript
the idea is to then describe the entire
manuscript
so that um and what we also do is we
have a few
sort of key very important collections
for which we decide we we try to
get all of the manuscripts described in

in the database
so we try to because we know we can't do
everything like if we start with a
library catalog
a we will never get probably not even to
c
let alone to the end so um
we wanted to have sort of a target using
the knowledge we already have but we
also want to include
for example the
and all of the wonderful small case
studies that have been done there we
want to integrate those to see what they
mean for the network as a whole
uh so that is a little bit our our
strategy but it's also constantly being
adapted
when new partnerships become available
or
when when a mass mass
semi-automated import becomes an ops
option so we are constantly trying to
get more clean data in the database as
as fast as possible
but for for the ameliaris i mean from
from
bavaria or um or from oser
are you taking the the number the entry
number from barry or
other uh other repertories as a sidle
for your database or not
um um we have them recorded in the in
the database but we want to give
eventually each authority record um a
passing code with which will be
um uh which will you will be able to
cite but it is
uh generated randomly so it doesn't have
the problem that the
analog reference systems have that you
have to continue to add
like a's and b's and everything to to
the
um identifiable codes but so basically
every reference

label that is out there is present in
the database but we're trying to
um sort of clean up that entire
tradition of referencing these texts
okay thank you very much
i think looking at the clock that we
should conclude here
we have again a short break uh and then
we'll reconvene
at uh ten past uh three
is it ten past three even yeah for the
next
uh paper thank you again cheri
and irish for a wonderful presentation