# CHoRUS Bridge2AI Standards Office Hours-20240926_175429-Meeting Recording

September 26, 2024, 4:59PM

55m 2s

Talapova, Polina 1:04
Hello everyone.

Schmidt, Heidi E. 1:08
Hello.



Alvarez, Marta 2:53
Tell me now.
Sorry about that.
We're running a little over on the last meeting.
Looks like everybody on the call is has been here before so.
Paulina, would you like to start take things off.

Talapova, Polina 6:13
Yeah. Thank you, Mark. Tim.
Let's start.
So good day.
Today we'd like to introduce the Senate operating protocol for cataloging unmod terms.
And our research consortium uses data from 15 sites to support advanced machine learning and artificial intelligence. As you know, and collecting and unifying these diverse data is challenging because it includes various types.
Like structured information like structured HR, images and waveforms.
To manage this complexity, we use the observational Medical Outcome partnership common data model OM OP CDM and this model helps standardize our data sets, making integration and analysis more straightforward.
We've also aimed to develop specified specific concepts within the OMOP to handle multimodal files effectively, and these diverse data sets.
So we created reference mappings through a structured Delphi process where experts identified and prioritized important flowsheet descriptions for critical care research.
These mappings were validated by 14 clinical experts and they assist us this mapping.
Samine assist us in cataloging terms across our data sets and some terms are not covered by the existing mappings.
Leading to the unmarked terms and our standard operating protocol that we will share today addresses these issues by providing a clear process for identifying and cataloging decent MapReduce.
So we will offer guideline for documenting them and ensuring they are included in the future work and analysis.
Despite the fact that this SOP is currently marked as pending or awaiting approval, it has already undergone review during which changes and improvements were made and if no suggestions and comments arise after today's meeting, THESSF will be approved finally.
So let me share my screen and show you the page with SOP and also share the page with.
This current place of this SOP.
The first of all link.
In the chart and now I need to share my screen.
OK.
So you should be able to see my screen and.
My background so I already.
Told you a story of what we hear about in in general, and the purpose of these SOP.
It's it's a centrally managed and mob source. Terms within the chorus bridge to a project and this SOP provides detailed steps for integrating existing mappings, identifying unmap terms and documenting them for further action.
This is so P applicable to all project team members involved in data mapping and mapping creation and now procedures.
There are five steps and the first step is to integrate reference mappings into your homomorphic instance. The first sub step here is to verify homosplym deployment and accessibility.
So you need to ensure that your homophobia instance is correctly deployed and accessible.
And particular attention we need to put to this omop center. That's vocabulary schema that all tables.
Presented in our omop CDM instance, the next sub step is to confirm that all required standardized vocabularies have been had been downloaded from Odyssey Athena installed and they are from the latest official Odyssey vocabulary release.
The second sub step in the first subsection is import mappings into the vocabulary tables.
In this case you need first access the mappings called Bridge two AI ontology available in the chorus mapping stage.
Ontology CSV files.
Just let me show you.
This place.
Chorus mapping stage ontology and there are CSV files with all delta tables for each omop CDM table.
Let's come back to this place.
So if you don't have access to these directory, please request it from Jared Hotelling.
And there is his e-mail address.
After this access, if access is granted, lower the mappings from chorus mapping stage into the corresponding tables within the OMOP CDM vocabulary schema. So we populate concept ancestor using concept ancestor delta. We populate concept class table using concept class delta. We also populate concept table using concept DEL.
Concept relationship using concept.
Relationship delta.
Concept synonym using concept synony table delta sorry source to concept map table using source to concept map. Just source to concept map and vocabulary table using vocabulary delta. We also offer mapping metadata table with DDL and possibility to populate this table as well with information related to Ma.
Precision mapping, origin and maintenance.
And now we can go to the second step.
Second step is to prepare concepts to be mapped.
In this place will compile a set of source terms requiring requiring mapping.
We need to prioritize these terms, usually based on their frequency, and also it's very important clinical relevance.
So then we categorize terms based on their respective domain within the OMOP CDM and these domains can be the most frequent measurement, observation, measurement values, condition, procedure, drug and device.
Step #3 is to develop matching algorithm.
So once we develop the concepts to be mapped, we can join this list of concepts with their bridge 2AI ontology.
User usually we use join query and we can implement the following matching techniques.
First one it's a case normalization, so we ignore lower-upper case and use one of them.
Then prefix removal some prefixes acronyms without important.
Sort of clinical sense can be present in the source data, and they also can be trimmed.
Then terms splitting.
In some cases it's relevant to split the source storm to retrieve more relevant source description.
And also removing redundant information.
Also, there is an option such as a full name mention.
Concept synonym name matching using the concept synonym table.
Also another option for matching technique is removal of digit spaces and special characters that can prevent the matching of these terms.
Also, fuzzy matching is applicable, such as Leven, Stein distance or Jared Winkler distance and also bag of words comparison that was demonstrated by Jared previously.
So then we also consider combining multiple matching techniques to optimize accuracy.
And after this we perform mapping and validation.
So literally we execute the join query to identify potential mappings between our source terms and target terms.
Then we usually validate the generated mappings for accuracy and completeness, and then create a delta tables containing source descriptions that remain unmapped.
You may notice there is no such a stop like a manual mapping or using Usagi for mapping.
This information can be covered as a separate SOP or additional guidelines.
It means that it's possible to enrich your mappings using manual mappings and Osaki tool, but it's not obligatory.
It means that if there is, if there is any difficulty to perform this, it means that you can skip these steps and only perform joining technique to.
Define the unmapped terms and then we go to the Step 5.
So it's documenting and sharing unmapped terms.
Here we record the delta table from map source terms, ensuring all relevant metadata is captured and organized for clarity and ease of reference.
What I mean by metadata?
So first of all, we need a field indicating the origin of the term.
It can be source table name if its proprietary. For instance it can be replaced by some identifier like flow sheets or procedures or conditions.
So origin of the source data.
Then the field count the number of times the source terms appears in the source data. So it's a frequency of occurrence of term in the data data set.
Usually it's numeric.
The next field is a source code.
So unique source code associated with the source term or concept, it can be.
Numeric can be alphanumeric.
It also can be just.
The same as a source description. It also can happen, but anyway we need this field.
The next field is source description.
It's a factual textual description of label corresponding to the source code.
Then we have optional fields that are applicable for particular domains. For instance marks, source value and marks source unit.
So marks source value.
It's the most frequently absorbed value for a specific source term across the data set, and it's relevant for observation and measurement domains.
And it means that.
Particular value can help can help us to map unmappable term and better understand its semantics.
The same this Max source unit field is the unit of measurement associated with the most frequently observed value and also it's applicable to measurement domain.
So it means that for conditions or procedures we do not use mark source value or Mark source unit and comments.
This field captures notes or any sorts like reasons for an unsuccessful mapping or challenges encountered, or who did the attempt for to do this mapping.
So any kind of dis information can be stored here and also important to note that for observation and measurements, usually we need to extract units separately.
And they can be extracted from a separate column or they can be embedded within the source description and they should be extracted from the source description.
Then we map these units to standard units in the Amorph vocabulary. If a unit cannot be marked with document.
In the delta table with appropriate comments in the same way, in this case source code and source description will be description of the unit.
And units also can be submitted as unmappable terms, so any kind of domain can be submitted.
And another important point is we need to recognize that measurement values also can be a concept.
It, for instance, qualitative results like positive or negative. They can be mapped separately to standard Omo concepts.
So we recommend attempt to map these values and also document any unmapped values in the delta table.
Created separately for values mapping and the last sub step here is to distribute the delta table to the standard team using a preferred method.
It can be e-mail, so I kindly ask Marty to help with this.
So send directly to Marty Alvarez or use GitHub.
It's possible to upload the file with unmapped terms to the specific unmapped terms.
Folder in chorus mapping stage repository. I will show you.
So here is unmapped folder.
And example of table with fields to be populated.
And coming back to the end of this SOP.
So name and convention.
We also ask you to apply specific formatting as shown here.
Through the upper camel case.
Through the camel case we have unmapped category which specifies the semantic category of unmapped terms. It can be flowsheet items or flowsheet values or conditions or procedures.
Then site name is the institution and then the date.
A digit format.
And also important to mention that we would ask you to choose the method for sharing the delta table of a maps or storms that adheres to your organization data policy and addresses the proprietary concerns.
So ensure that the selected approach complies with your security protocols and everything should be OK.
According to additional consideration, they are very brief.
So regularly review and update the mapping algorithm to improve matching accuracy.
If this map and algorithm is good, it means that you can reuse it again and again and customize it according to each mapping and also establish clear guidelines for handling ambiguous or multiple matches or any kind of logic that you applied for selecting the best mapping.
So that's it.
According to this SOP, if you have any questions, please let me know.
I would be happy to answer.



Williams, Andrew E 21:54
Thank you so much, Polina.
Just want to emphasize a couple things you said so.
Working hard to do as much automatically as we can to take that burden off you. As Polina said, it's possible if you have the resources and the interest to try and do some manual mapping to do things that extend what can be done automatically using the techniques that.
You covered, but we're really not expecting that.
Not only because it's, you know, it's a little bit time consuming because the expertise is really a bit hard and I and I.
Not necessarily intuitive as to why it's hard. I want to.
Get on a soapbox for a little bit, or maybe be usefully didactic.
That's it.
Not a soapbox? It'd be usefully didactic.
Why is this hard to do?
You might think a rational smart person would think.
The name for a concept is descriptive.
It's going to kind of tell you what that is, and that's only really very partly true.
This whole way of working with concepts that are standardized.
Means that the meaning.
Isn't inherent in that description.
Most of the meaning is in the relationships to other concepts.
You know what something means by knowing it's a parent of this thing is a child of this thing.
It has these other attributes.
It has these other synonyms.
It's not all wrapped up in the description.
It has to be represented.
In axioms that use formal description logic that allow the whole thing to be computable in a particular way that has specific rules associated with it so.
It's both supposed to have some transparent meaning. When you take into account all of the relationships and attributes and synonyms and so on, and you know its domain and you know the other, you know things about it. Not all of that being encapsulated kind of in a in.
A description.
So that's why it's challenging like that means.
To understand those hierarchical relationships and in relationships other attributes that might be associated with things, you have to have.
Been kind of living in the space where these vocabularies are and how they're used and how they're generated to kind of get a feel for what things really mean. This specific task and the guidance that Polina has provided on, say, take the label that you see in your.
Flow sheet and say does it match this thing or not after she's drafted? It is a very useful, important thing and it's can be slightly misleading. Is like that's how you understand the meaning of all these concepts in a lot of ways. You don't right we need.
This whole structure to really define how those things.
You so when you're trying to go about this task, if you decided to like take that on locally, it's kind of you need both the content expertise, somebody who understands where these flow sheet data come from and this pretty.
Unusual set of skills about how how ontologies work, how all of that description logic plays out when you're actually trying to use it in algorithms.
How it what the domains mean? So on and so forth and you've all gotten, you know, some really useful steps on that through this, if you didn't have it already.
But it's still a lot to take on.
And so that's, I just wanted to kind of amplify kind of why it's, you know what the.
What we mean, we say the sort of ontology specific expertise for doing this the right way.
What kind of things that Paulina and others are bringing to the table and do that.
It's it's, it's that stuff. And having been immersed in it for a long time and seeing how it plays out, that really is required to to do the job in the way that we want to. you know, make sure it gets done so that all the algorithms work.
As they're supposed to, and people interpret the results and so on so forth.
So that's I don't think we've ever covered that specific thing about how ontologies work and what definitions mean.
Descriptions because it's a lot to do.
I I I'm not really trying to make sure you walk away from this little.
Blurb feeling like you understood exactly everything I said, because each of those things.
Like what we mean by axioms?
What formal description?
Logic means what?
You know how you how you actually do that?
That's that's the stuff that's kind of, you know, it's the it's a set of skills and so on that we don't expect people to have or to have the interest to really acquire a lot of. So I just.
Wanted to elaborate on that. In addition to thanking you pleina for driving this and all the work you've done in leading it and.
That's a little bit of a cautionary tale, though, about like taking on a lot of mapping locally, but it might seem like it's more straightforward than it is. Probably does seem more straightforward than it is. So that's why we're trying to automate it and we're going to take.
All the things that you contribute through this and you know, work on them centrally and collate them and develop, you know something that combines.
Ontology expertise and clinical expertise and algorithms to try and.
Take on this whole lot large bolus of things that haven't been mapped already and and develop good contributions to the omop vocabulary that that handle all of that and make it standardizable and computable.
So it's a little long winded.
Hopefully that's useful context though.
Anything to add to that Paulina?
About what?
What I was saying about why it's why it's so great to have a polina on your team.
Just mostly.

Talapova, Polina 27:17
No, I have nothing to.

Williams, Andrew E 27:18
Mostly that's what that was like. What?
You know, it's it's not as easy as it might seem, so.
I think we had some other things on there.
Are there other questions about this process and how it might play out at your site so.
I'm expecting, I guess, not much in a way.
Of suggested modifications of this. So I think it was very carefully put together.
It's really nice if there is some. We're certainly welcome to it, but we're expecting it to be approved.
That's why it's been shown. But there are other comments about how you might actually use the SOP rather than modifying it.
But like what you think of this process and how it's going to play out at your sites as you try and do it?



Balakrishnan, Kasturi 28:11
Hey, Andrew and polina. I just had one question.

So during the validation process, which is where we take the mappings provided and review it with what we have, if we notice anything that's that seems like it needs attention.
Do we also submit it in the same format for? Because I think I saw something for the unmapped and I'm sorry if I missed this, but if we have questions about the things that are mapped, but we think that our values look different.



**Talapova, Polina** 28:33
Good question.

**Balakrishnan, Kasturi** 28:40
And we want to bring it to your attention.
What do we do about that?



**Talapova, Polina** 28:45
Yeah, this is a great question. And this is a topic for separatists.
Sophie, what to do with a questionable mappings? But in fact, yes, you could share them with us as well in the same way, and it would be great.



**Balakrishnan, Kasturi** 28:59
I have not sorry at Eddie's looking at some of these mappings and we notice a couple of things together and I don't know if it's gonna be a huge number or not, but I just wanted to, you know, raise that if there was a process for it.



**Williams, Andrew E** 28:59
So let's go ahead.
Sorry, no. you go ahead.



**Balakrishnan, Kasturi** 29:13
What else I could just set up something with you guys and talk through it.



**Williams, Andrew E** 29:18
Set up a convention right now.
That distinguishes questions about a mapping versus can't map and how the same mechanism that Paulina just described can be used for both and well probably differentiate it later.
Pauline is saying but like for right now so that you can take the stuff you're finding with Eddie and and do something with it.
Let's just say use exactly the same mechanism that.
Apalina described.
Name the whatever it is the file.
Questionable. Put questionable in the title.
Other thing that you're submitting and will use that as the indicator for the time being, but we'll come back with a here's what you do with questionable versus. Here's what you do with campmap and make that sort of a less Kluge solution. That's not a terrible solution, but.
Does that seem like it's going to work for the time being?



**Balakrishnan, Kasturi** 30:08
Yes, it will. Thank you.



**Williams, Andrew E** 30:09
Erfect.
Anything else related to this SOP?
Violent and uproarious applause for all that went into it.
Just astonishment.
Yes, I think so. I'm with you.
Yey, thank you so much.
So we've got a growing set.
We heard you all loud and clear.
We're right with you. When we knew the size and complexity of your whole bridge to AI project needed some strategy for getting single sources of truth about things.
Tierra and other people worked a lot on building this out.
The strategy for doing that?
Jared and Polina and and Marty and others have done that and since that time we've taken, various people have produced what you just saw.
Sop that gives sort of the definitive process for doing some part of your work.
And there is a landing page we've shown you before, but we want to show again now that it's starting to be populated more.
We really want to kind of start to direct.
People more often to this as the place to go to get answers and Marty.
'S.
Busy calendar.
Is on full view.
Ah, there we go.
Marty, you want to take it there?



**Alvarez, Marta** 31:55
Sure. Yeah.
So obviously there's been a lot of people working on this and I think the originator of creating this was Jared.
And I have just been helping to try to link some things to this page, but we wanted to just at least start to talk a little bit about what we're doing and maybe get some feedback and.
Guidance from people about what might be useful moving forward, but the graphic right now is split up into basically 2 sections.
The one on the left is for data generating sites and the one on the right side is for. If you're working with data infrastructure in the cloud and then going back to the left side, the data, the data generating sites, there's two ways to look at Sops.
There's D1 through D6, and then there's also via data source.
Currently there's 601 through D.
So is 6 labeled sections and then of the data source. Obviously is listed simply by source.
Currently, there's three finalized.
And one provisional sop linked to this graphic.
And I can show a little bit of that.
And also we adding the catalog.
Cataloging of unmapped terms soon as well. So that'll be 4.
And there's also a new little bubble that we created here. How to find an SOP on this page which still again all of this is a work in progress, but I'll, I'll show an example. I'm just gonna click on one of these so that you can see what.
It looks like currently.
Anything that has a that looks like a link obviously links to the SOP page.
Example of the first SOP that was approved and linked, which is for clinical validation of mappings.
Sorry.
And then if you come over to the data source section.
Then that same thing is mapped here.
So there's basically two ways to look for Sops currently.
That give a little bit of a of a good intro.



**Williams, Andrew E** 34:44
It does.
And if you were doing it, you know, it was clear that there's a lot more than Sops here, right?
And as that section in the lower right shows, there's discussions, office hours, ways of monitoring progress that are also linked here.

**Alvarez, Marta** 34:50
Yeah, yeah, yeah.
Yes.



**Williams, Andrew E** 34:58
So it's a very useful page for a number of different views.
So maybe if you click on one of the same ones you showed before, you can see how you've done this great work of linking where the office hours are. You can see them.
Labeled and other kinds of resources.
Code bases, et cetera that are that are relevant to some of the tasks that are involved in standardizing data elements here.
And so it's not only a place for sops, but for some of the guidance that have been provided on office hours or some of the code that's used and so on.
And so that's the only thing I would add.
It's, I guess, to reiterate stuff we've gone over before the D1 through D2 were meant to.
Simplify the overall view of what's involved.
From a data generating site's perspective, is getting the data, standardizing it, linking it, making you know the different modalities of data, integrate with one another.
So in a, you know, unified analysis can be done.
So how the ID structure for?
Procedures is linked to the waveforms that are required as a result of those procedures, for example. So the EHR data.
Is linked to the waveform data and so on.
That is what's covered there, the deidentification.
And quality assessment of data.
And then how to submit it and the kind of feedback that one might get for D5 and D6 so that D5D6 is meant to be easy to intuitive thing? Like where am I on my progress of doing this meant to be organized that?
Way.
And then similarly and probably less relevant to most folks on the call, there's.
A similar kind of chain of activities associated with data once it's submitted to the cloud and is being assessed and eventually approved for merger with the main data set. That's in general how that's organized. So the relevant materials are all meant to be easy to understand and as.
Marty said.
We're eager to improve this if it isn't easy to understand or you don't find it.
Useful. It's really only good if it is useful to you.
So we're eager for your thoughts on it.
It's the only thing I would add, but thank you Marty and everybody, Jared and Chester and Sierra and all the people who divined sops and delivered office hour stuff for generating all that content and putting it here.
Again, I'm going to pause for the violent applause and uproarious joy at seeing all your questions answered in one convenient location because of the hard work of Marty and Chad.
It's just it is.
It's just that good.
I can't contain myself either.
Going to have to applaud publicly for it, OK.
Else, do we have in our agenda today Marty?
That's it.
We done? We gonna let those were the two topics we had?



**Alvarez, Marta** 38:14
Those are the two topics that we had.



**Williams, Andrew E** 38:17
I I have other things to bring up because we had a really productive meeting with.
Mark Wainwright, an entire crew from.
And a couple of things came up and I guess maybe we can have a little bit of a discussion about that we you know that we're involved in trying to get an initial stab at one of the things we knows is required, which is being able to you.
Know provide I DS for the type of ICU that data acquired from.
And that's a current gap in the Omo vocabulary we're gonna fill.
Gonna be reporting to you on that. Other things. I'm not sure we do have yet, but I wanna bring up here. Specially since you're on a call. Eric, you know I opined that death data are in a little bit of a special category because death as an out.
Is so important and we know there are inherent limitations to capture of death itself and cause of death when it happens in in hospital data and some.
Strategy for.
Demarcating at the end, when the chorus data set is finalized as to where we think sites have had access to out of hospital data and where they have not is probably a good idea.
So I want to run that by you.
I I advocate for that because I think it is so often an important outcome.
It's one that we want to pay special attention to and have metadata about so that people don't go in and say, oh, all these sites have the exact same kind of completeness and granularity with respect to their death data capture.
The other.
We thought is where sites don't currently have access to it or not aren't sure where they do.
We're probably going to share some things inside Epic user web for where we at Tufts are aware that those data are stored.
So just because they it's possible to do an epic doesn't mean it is being done at your own epic site or your own server site.
I think we want to share that so that you can at least check 'cause. Again, I think it's.
Extremely important outcome and we wanna kinda underline it in a way that we might not underline all the other kinds of data that we get but really help understand both how to maximize access to it and.
Capture appropriate metadata about the extent to which we're we're confident in its completeness and so forth.
So Erica, you have thoughts on any of that?



**Rosenthal, Eric S.,MD** 40:37
For death data.
Why? How would we not capture it? In other words?
Or I know it's sort of a negative question but.
Where? Where would it be captured?
So we we, you know, there are death flags in the electronic health record in some, I know Cerner very well. But in Epic and the challenge about these is that if you're looking at a patient from a year ago, they could have died after the discharge, right?
So are you asking about death during the discharge?
I would think that would be a discharge disposition.
And that's where the data would come from.
In other words, it'd be to cease to discharge.
Are you asking about the hospital in admission or are you asking about follow up outcomes of if and when the patient died? Because then you're dealing with, I think a couple other sources public, you know, data sources as well as.
Flags on the record, which you may not have the exact time for.



**Williams, Andrew E** 41:42
Was talking about things like National Death Index or state death registries that would supplement what you see in Epic where that's captured.

**Rosenthal, Eric S.,MD** 41:45
Yeah. OK.



**Williams, Andrew E** 41:49
So I think sites may very well vary whether or not they use the capability of Epic. To do that, I think you clearly see it at MGH.
It's not a given that all sites will do that and do it in the same way, so I guess some special attention to it because of its importance as an outcome is the main thing. It's.
Where we think we seeing it.
Information about death and or cause of death. So it's not necessarily just the occurrence of death where you've got something in an in hospital record about it, but where the additional adjudication about cause of death might happen subsequent to that as part of whatever's happening in the.
National or state registry.
Where that information is, we've got a little bit more understanding about cause of death.
I think that's the kind of thing I wanna make sure we're just paying extra special attention to both in terms of helping sites get that data.
And then being transparent about where we think it's come from coming from the end of the day, yeah.



**Rosenthal, Eric S.,MD** 42:48
Do you think we'd have to pay at each site for that information?



**Williams, Andrew E** 42:54
Don't know. I think at Tufts I could say we get it from the Massachusetts Department of Vital Statistics tends to be at this point about six months out of date.
They spun up.
They speed up the release cycle for it during COVID because there was a need for more timely releases of data, they're slightly less confident about some of the results of adjudication because it's faster than it used to be.
But they give that for free.
It's easy to just ask them for it, and they'll they'll give it.
You can link it.
I don't know.
How that applies to all your sites in chorus. So I think it's probably variable and but just in general paying extra attention to it I guess is one of the main things I'm floating here is something we we probably want to do because of how important death is.
It's now come.
That makes sense to you.



**Rosenthal, Eric S.,MD** 43:47
Yeah it does.
Wonder.
Maybe we should be doing a little survey of rpis, whether they are aware of source, that they could use.
The national indices you have to pay for them unless and.
You're passing the unique identifier of the patient, so we couldn't do that centrally, so.



**Williams, Andrew E** 44:19
You could use a private preserving record linkage strategy like we do in some other context to to do that in a way that is OK at least has been approved. But.

**Rosenthal, Eric S.,MD** 44:29
Sure you could. but.
Yeah, it would.
It would take. I was sort of thinking about within our existing RV.
So that would be a a new like a new amendment to our.
To our IRB. I guess this would be anyway because it's a new source of data to.



**Williams, Andrew E** 44:47
Mm hmm.

**Rosenthal, Eric S.,MD** 44:48
You know to search.

**Williams, Andrew E** 44:51
Think it's not absolutely, categorically new in kind, as I think we had already planned to have some external publicly accessible data sets related to social terms of health and other kinds of things, be a part of what we're dealing with.

**Rosenthal, Eric S.,MD** 44:52
It's worth it.

**Williams, Andrew E** 45:06
So it's not.
It isn't the case. I don't think that we never said anything about non EHR.
Or healthcare generated data.
I think we talked about.
Publicly available data sets without specifying and then event we could we could look into it, I agree.
Bears scrutiny as far as what was approved and what IRB approval might require.
But think that's useful.
There are, I think, open source pprl, privacy, preserving record linkage approaches that don't necessarily cost money to implement that we could investigate, but mostly I guess I'm interested in not adding a huge undue burden for new things, but just making sure we don't pass by one of the.
Main outcomes without paying appropriate attention.
To it, when we're talking about acquisition and standardization of it, so.
Guess more to come on that Mark also brought up excellent points in relation to that about withdrawal of life sustaining therapy.
How it's captured and whether that's something we want to have here.
So I guess it relates to the cause of death information.
I don't know, Mark. I I mentioned your work and.
My privilege of helping to collaborate in some minor way with it.
But Mark, I don't know if you have additional thoughts on on that.
Have it to Eric.

**Wainwright, Mark** 46:26
Yeah.
Alexandra, I just put it in the chat now.
I mean, it's not.
It's an issue likely well aware of for a long time, Eric distinguishing between what's withdrawal of care and what's actually despite escalation of medical care. Even that would be an advance, I think, for for this field to it's.

**Rosenthal, Eric S.,MD** 46:54
So.
The way I've typically seen this done is by looking at so-called early limitation and withdrawal of licenting therapy.
And different people define that differently. In the first, you know Fortyweight or 72 hours, but.
And part of the reason I think is because in essence, almost, you know, every patient eventually in The Who dies in the hospital has withdrawal of of care.
You know, whether it's that the code is ended or you know well, some people maybe found it, but you understand so.
So sometimes setting a time limit gives some information about us to whether that that was the case.
The other way we could look at it is to develop some type of therapeutic intensity level measure. And if you look at like track TBI for neuro, they the the therapeutic intensity level is pretty hard to.
Compute.
Because it has a lot of things about for the intent of X.
In other words, not just that you're getting sedation, but it was in sedation for the intent to do something and.
Outside of text, it's different.
Difficult to often intuit in the intentionality, but I don't know some of those things are would be of interest and Andrew or others. If you think there's a way to compute some measure, it won't be necessarily always the same for each disease condition. But if there is some.
Way.
To understand. you know that that would be of interest for.
Some types of life sync therapy, you know, of course it's intensive until it's withdrawn.
But.
Yeah. So I I'm not sure how to get around that, that aspect of it, but.
But I I do agree.

**Williams, Andrew E** 48:54
Might do something relatively efficiently. Go ahead.

**Rosenthal, Eric S.,MD** 48:55
No, no, I do agree. I think the the minimum that and it came up in the gap analysis chart where we were creating these or annotation or whatever we wanna call it that chart of looking at data elements we were missing was orders you know for limitation and.
Withdrawal by sync therapy.
But Mark, I sort of hear you wanting to go a little further, which is sort of looking at the intent of what people are trying to do.

**Wainwright, Mark** 49:21
Ideally, but I agree. I'm well aware of the limitations trying to abstract intent from the electronic medical record. I think that would be likely to fail and not worth the effort.

**Rosenthal, Eric S.,MD** 49:25
Yeah.

**Williams, Andrew E** 49:35
I guess you know in the cancer domain their intent is described for palliative care as opposed to therapeutic.
Intervention. I think that's often codes for it that are distinct because it's a part of a formal plan that's, you know, got a a distinct class of therapeutic intent associated with specific codes that it can be attached to orders and I think.
It probably isn't the case in medical care, but I think.
Investigating to make sure that's true is one thing we could do. I think also if we're defining targets for the NLP I would be interested in exploring that. Again, it seems like it's an important thing that's of interest to researchers and perhaps of interest to people doing algorithm develop.
If we had some confidence that we could through notes and some combination of a measure of intensity or the kinds of therapies that are being developed, if they seem mostly.
Palliative at some point and don't involve, you know, some broad class of things that are.
I don't want to put the words to it, but you all know the feel better than me. Whatever it is, that one might chunk kinds of therapeutic inventions.
And into understanding their intent into that, we could investigate that relatively easily without committing to saying we're going to do this and do it well enough to be, you know, confident in it and ask everybody to implement it. But some initial investigation seems at least within scope, so.
I don't know. That's my.
Initial thoughts?

**Rosenthal, Eric S.,MD** 51:12
So is it just in terms of the orders aspect though?
Is that a procedure in the omop lingo?
In other words, if someone writes an order for 'DNR DNI.
Does it get entered as a procedure or is it an observation of the patient's wishes?
What is it?

**Williams, Andrew E** 51:37
Good question.
Like we'd have to. you know, I I wouldn't want to answer the top of my head on that.
I would assume it is a procedure, but I could be wrong, so I think that's a a thing to to be clear about.
I don't know if Polina or Jared or others know the answer to that more definitively than I do.

**Talapova, Polina** 51:59
There is.
Node domain clinical observation concept class that addresses quitate.

**Williams, Andrew E** 52:10
Concet class do not resuscitate as a kind of note that.

**Talapova, Polina** 52:12
Is it as a? As a note. yeah, kind of note.

**Williams, Andrew E** 52:15
A kind of note.
So it would just be without even seeing what's in the note. just some indication that a note that follows into that class has.
Been created during a particular.
Point in the stay is that the the gist?
I.
I guess I haven't seen that before, so that's interesting.

**Talapova, Polina** 52:37
And also there is observation with clinical find.
Status with supporting documentation.
From later than this way.

**Williams, Andrew E** 52:48
Well, that might be low hanging fruit to just investigate whether, let's say, at a couple of sites we can see.
And systematically relate the creation of those notes with a particular point in care where they happen.
That's at least an initial indicator.
Might be straightforward if we see that being widely implemented, so it's great.

**Rosenthal, Eric S.,MD** 53:10
And death is is that in a person table or is it?
Is it a? It's a death table.

**Williams, Andrew E** 53:14
Its own table.
The death table.

**Rosenthal, Eric S.,MD** 53:19
OK.
I don't want that table.

**Williams, Andrew E** 53:22
Yeah, try and avoid that one.
Right. Well, that was very useful. I think we have some to do's following those discussions. Just to reiterate, I think we were gonna share in at least the epic user web, perhaps some additional investigation of what we did when we were Cerner shop to get that data.
About death and cause of death in hospital, Cerner. And come back and say where it's been shared in epius or web or where it's been shared elsewhere for the Cerner. Just for places for people to go look and find their death data.
And probably some initial analysis of what it looks like to win these.
DNRD NI notes are created and see if we can just reliably capture them and the the points in a visit where they're cat, where they are created and come back with that and not necessarily commit to doing anything with it, but just as an invest.
Of it.
So that's really useful. Great.
Alright, sorry Marty was let it.
Marty was trying to give you all some free time and I foiled her well-intentioned plan.
I'm sorry, but that was useful and thanks again, Polina and Marty for presenting.
That's it.
5 minutes.
Back. See you all next week.

**Houghtaling, Jared** 54:51
I think so. See you.

**Talapova, Polina** 54:51
Thank you everyone. Goodbye.

**Alvarez, Marta** stopped transcription