

Mike Dow: For the last couple years, I've been responsible for Health Catalyst Text Analytics Initiatives. During this time, I've gotten to know Dr. Chapman, her work in her department of the University of Utah. We're excited to have Dr. Chapman join us to share her perspective on NLP in Healthcare today. I'll start with her official bio, and give you a few additional thoughts before we jump in. Dr. Wendy Chapman earned her Bachelor's Degree in Linguistics and her PhD in Medical Informatics from the University of Utah. Between 2000 and 2013, Dr. Chapman held roles at the University of Pittsburgh, and then the University of California San Diego. In 2013, she returned to the University of Utah to become the Chair of the Biomedical Informatics Department. Her research focuses on developing and disseminating resources for modeling and understanding information described in narrative clinical notes. That's the official bio.

My thoughts are these. Dr. Chapman is a national leader in clinical NLP, making major contributions in this area over the last couple decades, both in research and how we use NLP in industry. In the last couple of years as I mentioned, I've become familiar with the Department of Biomedical Informatics that she chairs at the U. The quality of the curriculum and the students from her department is fantastic. They have a focus on the emerging fields of machine learning and natural language processing, in addition to a more traditional curriculum. Today's webinar is titled Tapping into the Potential of Natural Language Processing in Healthcare. In my opinion, this will effectively be a master's class in NLP in Healthcare Today covering the potential of NLP in healthcare, the challenges, and Dr. Chapman's perspective on what we'll see out of NLP in the future. Dr. Chapman, thanks for joining us today, and I'll turn it over to you.

Wendy Chapman: Thank you for that kind introduction, Mike. It's a really exciting time to be in healthcare right now. I think there's so many possibilities, and natural language processing is a big part of that. Most of the people on the webinar probably have seen a slide like this where we see that the United States is spending a lot of money on healthcare, and their outcomes are kind of in the middle across the world. Because of that, our healthcare system is in crisis, and that means that the US economy is in crisis because the healthcare system is such a large part of our economy.

Well, now that I'm the Chair of a department, I've been taking a lot of leadership classes, and I've learned that crisis is really just another word for opportunity. There's a ton of data available right now, and we really have the opportunity to create a learning health system where we're learning from the way that we treat patients so that we can continue to improve.

When you think about the data that we have, the words big data, machine learning, AI, those are words that everyone knows now. They're not esoteric anymore, and articles are even being published in journals like "The New England Journal of Medicine." Everybody has an electronic health record now since the HITECH Act has gone into place, and so there's a lot of data available, and there are a lot of opportunities to leverage the electronic health record to help us deliver better care.

This is a video, which we'll send out the link after the presentation, but I'm just going to describe it about the EHR of the future, and as you listen to me and you watch this video later, think about how close is this to your EHR? This EHR is called "IRIS," and the physician is sitting with a patient, and he's talking about ... she's asking him when's the last time you got a colonoscopy, and IRIS is listening in because it's got speech recognition. It's got natural language processing. It understands what's being said in that conversation, and it pulls up past records of the colonoscopy and says, "You actually had a colonoscopy about six years ago, and here's what the findings were, and it was recommended that you schedule five years from then." And then she says, "Do you have any family history of colon cancer?" And he says, "I think I might've had an aunt," and IRIS is listening on again, and she pulls his family history chart of cancer.

And this is the type of possibility that we can have using the data that we have and the EHRs, the technology, and the computing power that we have nowadays. And we see articles every day in the newspaper about all the great things that are coming our way and that industry's working on and researchers. But let's step back to reality a little bit that the EHR that's in this video is not at all like the EHR that we have now, and in fact, EHRs are a big source of dissatisfaction and even worse than that, inflicting enormous pain on doctors, and doctors spend about 50% of their time on the EHR, which leaves less time for patients and for thinking about the higher level reasoning type of questions that they need to do.

Maybe it's not that the possibilities are endless. Maybe the pain is endless. Well, it's my view that NLP can be a big part of that solution. Let's talk about how NLP can come to the rescue and really help us in this space, but I want to have a caveat that this is the hype part of the talk, and we will also, like we did with the EHR, get to the reality part of the talk about NLP. But first of all, let's talk about the cool things that can be done with NLP.

Knowing somebody's social history can be really helpful in predicting if they're going to be readmitted and in helping them after a surgery better take care of their wounds, things like that. And so, you might think that in the EHR, there are these dropdown lists, and everyone fills in this type of crucial information that's needed for taking care of patients. But in reality, that information like other information is really embedded and hidden in these textual reports, and it requires natural language processing to extract it out so that it can be used.

The high-level picture of what natural language processing is it takes as input some text. The text might come from the EMR. It might come from the literature, social media, and what it does is it processed that text and outputs some kind of structured data that then is machine interpretable. The output might be used to classify patients into certain categories. You might be extracting particular information that you're looking for, or it might summarize information.

I first wanted to know about how many of you are currently using NLP in your work?

Sarah Stokes: Great. Thanks, Wendy. We've come to our first poll question, and as Wendy said, we'd like to know, are you currently using NLP in your work, and your options are yes or no. The votes are pouring in. Wow. I don't think I've ever seen those numbers go up that fast.

Wendy Chapman: This is an exciting topic, Sarah.

Sarah Stokes: It is. Okay. We're going to go ahead and close the poll there and show the results. Sixty-seven percent said no so that's your majority, but 33% said yes. Does that surprise you or is that kind of what you would expect?

Wendy Chapman: I think that's pretty close to what I would expect. I think a lot of people want to use NLP, but knowing how to use it is a bit of a mystery, and then there are people working on it so that's really interesting.

Sarah Stokes: Okay. You'll just click there. There you go.

Wendy Chapman: All right. Okay. Let's talk about some of the exciting ways that NLP is being used. And I'm going to talk about four different areas. The first is the usability of the EHR. This is a screenshot from New York Presbyterian Hospital, a typical EMR interface. The information is arranged by the encounter, the encounter that the patient had, and it's very difficult to find what you're looking for when you're trying to care for a patient. Noemie Elhadad, who's at Columbia University, she worked on a system that she called Harvest, and her goal was really to present information from across the EHR to the clinicians in a way that would help them find what they're looking for faster.

This is the interface that they have embedded into their EMR, and if you look in the middle where it says Salient Problems, all the problems that were mentioned in that patient's record are displayed here, and the larger problems are the ones that are mentioned more frequently. The one in purple is the one that the user has clicked on, "Dyspnea." All of the mentions of dyspnea are shown on a timeline at the top, and those could be diagnoses, ICD codes. They could be mentioned in notes. Anywhere in the EHR that they're mentioned, it would show up there. And now, the user could click on those areas and drill down and see where it's mentioned. All the notes about dyspnea are listed in the bottom left, and so every note that mentioned it is listed. If you click on one of those notes, then it shows you the note and unfortunately, it looks like my picture got cut off and I didn't notice. But, a little bit down, it says "shortness of breath," and it's highlighted, and so you can see why this note was selected as representing dyspnea.

She did an evaluation on this to see how clinicians like it, and it was very positive, that they felt like they were being able to find things that were

typically buried in information and could pick up on things that they normally wouldn't be able to find because it would take too much time. That's an area where quite simple natural language processing actually can go and annotate the patient record, and it can be used in a way that really helps users.

The second area where NLP is being used a lot is predictive analytics, and this is an example from using social media to help identify patients that you're worried about committing suicide. And so, this is a study and we know suicide's a big problem, and it's been increasing in the last decade, and this system was able to predict 70% of suicide attempts by monitoring each person's social media posts, and it only had a 10% false positive rate. And what they found was that the indicators of suicide were emojis, for one thing. When they started using fewer emojis, and they narrowed the emojis that they used to certain groups, mainly like you see here about hearts and broken hearts and sad hearts, and then there was an increase in angry and sad tweets before a suicide attempt. And so, using that information with machine learning on the text was able to predict that someone was going to commit suicide. Social media can be really powerful in looking at trends and attitudes.

The next area where NLP is being used consistently is in phenotyping, and a phenotype is the observable physical or biochemical expression of a trait, and so it might be your physical appearance. It might be a symptom that you have, a behavior that you have. It's what you can see and measure. There's a lot of phenotyping going on because you're trying to say, "What does this patient look like. Let me find a cohort of patients that has these traits," but most of the information that we use for phenotyping is from structured data because that's the easiest thing that we can capture. And if you look at this diagram, the green areas are the structured data, but underneath all of that structured data is follow-up appointments, vitals, charges, orders, encounters, symptoms, all kinds of information that's really contained in the EHR but is in text. And so, that information could bring a lot more to bear so that we could create better phenotypes of patient groups.

An example of a project that I did in the phenotyping area is with the collaborator at UCSF named Salomeh Keyhani, and she wanted to compare, does medication or surgery work better for patients who have severe carotid stenosis? And this is a very typical problem. You want to compare two things, you need to create these two cohorts of patients. What are the patients that have medication? What are the patients that have surgery?

First, you have to identify patients with severe carotid stenosis, and that information is in radiology reports, and it's not so simple because it's not always stated as, "Patient has severe carotid stenosis." There's a lot of different ways that it's said, and we care about the severity of it. We care about where it occurs, and so we used natural language processing to filter these patients because there are hundreds of thousands of them, and it's just not feasible to do manual review to create these cohorts. And so, the natural language processing would exclude patients, and then it would include patients that had

severe carotid stenosis, and those were manually reviewed so that she could perform the study. And we were able to decrease the amount of manual review that had to be done by about 70%.

While looking at, does a patient have this, yes or no, that's a typical type of phenotyping that we do, but there are a lot more questions that we have and a lot more types of patient cohorts that we want to create that are much richer, and that information is in the text, and so this is describing a project that was done by some collaborators at Pittsburgh and Harvard. If you look at a pathology report, there's all kinds of information in there about the condition the patient has, the location, the stage of the cancer, the procedure they had, the medications that they're on, and even their genetic status. If you can extract that information using NLP, and we can extract a lot of it, then you can answer much more detailed questions than just find me the patients who had breast cancer and the patients that didn't, and that allow us to be able to do a much deeper, richer research studies and better decision support.

The fourth area where NLP is often used is quality improvement in the hospital. One area of quality measurement that's required by the government is in colonoscopy exams, and they want us to measure how often do we find adenomas in people's colons? And the adenoma detection rate is a very simple calculation. Of all the colonoscopies that a physician did, how many patients did they find at least one adenoma in? It's a really simple measure but it's very important.

To do this, what hospitals typically do is they pay a lot of humans to take a sample of these patients that receive colonoscopies, and read through their pathology and colonoscopy reports and then calculate that rate. But it's a very small sample, and it's not in real time. It's used retrospectively. And so, Andrew Gawron is a researcher at the University of Utah who is working on creating a report card for physicians, actually, of how well they do on adenoma detection rate.

It's a really important problem because if you can increase a physician's adenoma detection rate by 1%, then it's been shown that you can decrease their patients' mortality by 3%, and so taking someone that gets 8% adenoma detection rate when the average is 12, and getting them up to nine or 10, is going to save a lot of lives. And if you can give them feedback about what their adenoma detection rate, then it's been shown that they change their behavior. That's a really powerful place for natural language processing.

Okay. I think I've gotten you all excited, and you're thinking let's do it, let's do some natural language processing, but let's take a peek under the hood and look at the reality. Why are only a third of you doing NLP and not all of you? This is my favorite car. Do you know what car this is, Mike?

Mike Dow:

I know it's a Ferrari.

Wendy Chapman:

Yes. This is the 458 Italia. I just think it's a beautiful car, and this is like natural language processing. It's just amazing, it's shiny, it's fast, but let's look inside the engine and see what's really going on. All right. If we want to answer the question, why doesn't everyone have an NLP system running at their institution, there are a lot of challenges to making NLP work, and I'm going to talk about four of them.

The first of them is the old adage, "Garbage in, garbage out." You can only extract information that's actually there and clear and easy to identify, and one of the things that's causing NLP a lot of trouble is that people are typing their notes now, especially at the VA, but also at other places, and because they're typing, they're doing a lot of shortcuts. They're creating templates, and they have the power to create all kinds of templates, any kind they want, and this is a big challenge for NLP because if you're trying to ... we're looking for sentences and we're training our system on certain things and handling these templates can be a real challenge. This is part of the issue of we're only as good as the data that we're given.

A researcher found that there are 58 different classes of this type of information that's pasted into notes including templates, and tables, and check boxes.

Another area of where it's really hard to deal with the data that you're given, is that because people are taking shortcuts, they're copying and pasting. And we've heard about note bloat, and you're putting information in the note that's outdated or inaccurate. You don't know who the author is, and this false information gets propagated across the record. And so, clinical notes are becoming less useful, and NLP can't solve that problem.

The second area is, NLP runs on text, and text is just a bunch of words, but what's the meaning behind those words? And so, you can build a really simple NLP system, and it can be quite powerful, but how do you really model the meaning behind that? And as you start to get more sophisticated in your natural language processing capabilities, you have to have a way to model that information.

And this is something that, you know, when I started my career in NLP as a graduate student, I never intended on doing anything about information modeling. That was not my interest, but it has been the core of my work because if you're not modeling it well, you just can't scale it out.

An example is you want to find patients with a cough for decision support, let's say. And so, you have define cough, and we're going to define cough as at least two episodes of a severe cough in one year, alright? That seems really simple to a human. It seems like one idea. But really, there's so much information that has to be modeled under that.

First of all, you have to start with the report itself, and the information in the report, and the mentions. You have to be able to find words that indicate that the patient had a cough. And then you have to know that it's severe, and that they had the cough, not that they didn't have the cough. But then you have to look over the whole report and you have to look at the episode and then you have to look at the patient level. And so there are all these layers of information that you need to model.

And you might say, "Well, let's just throw our terminologies out there. We'll just map to terminology and that's all I have to do to model the knowledge." And so the unified medical language system is a thesaurus of thesauri and if you look up cough in there, there are 25 different variations of cough that you can map to.

So, that's a start and you do have to decide which types of cough are you looking at. Do you care about bovine cough, acute cough? There's all these different types. But then, even more important is the information that occurs around the word cough.

Do you care about how long the cough lasted? What if it changed over time? Do you care if it responded to treatment? How severe of cough are you looking for? And so, all of this kind of information is part of your phenotype and it can be found in the text, in the words surrounding the word cough. But you have to be able to represent that somehow.

The third area is sublanguage. A sublanguage is a subset of natural language. And medical language is a sublanguage. It has a subset of vocabulary, it has different grammatical rules, kind of a subset of those, and it has these regularities peculiarities that you can rely on.

And so understanding ... being able to model the language that you're using is very important to be able to extract the meaning of it. We all know that social media is a sublanguage. It has new vocabulary in it, it has emoticons, it has abbreviations. And I know ... I tried to ... I'm not very good at using emoticons and I put one in a text once when my daughter said something sad, and it was a little emoji that was crying. And I thought I'm going to express my sadness to her, and she said, "That emoji that you put in there, means laughing so hard I'm crying."

So, I obviously don't understand the sublanguage of social media and so I shouldn't be using it. But your NLP ... you cannot take an NLP system trained on newspaper text and run it on social media and expect it to extract the meaning because it doesn't understand the sublanguage.

And so, medical language has a sublanguage and they're different. Medical blogs have one particular type of sublanguage and clinical notes have another type, and you can differentiate them by their use of verbs, by the types of

measurements they have, the types of ... the semantic types that they're using and there's also a lot of grammatical and syntactic information.

And so, your NLP system cannot be just bought or downloaded off the shelf from a company that runs on newspapers or on ... and you really have to be able to tailor it to understand the language that you're working on and that takes time.

The final area, is just the variation that you see in language. There's so much linguistic variation. There's so many ways to say the same thing with different words. And in English, one way that we get this variation is through derivation. And so you can say the same thing in an adjective form or in a noun form, and they really have the same meaning, they just have the different part of speech.

Also, we have two types of inflection. Something can be plural or singular, or we also have tense. Where it's in the present tense or the past tense or the future tense or all the other tenses that we have. And English is a lot easier than other languages in terms of inflection. But your NLP system has to be able to understand that cough and coughed really are the same concept.

But one of the main challenges to NLP with this variation, is the synonyms. There are so many synonyms for the same concept, and you'll never run out of new things that you do not have in your dictionary. And you think you may have a really complete dictionary and you'll just be surprised at how many ways people can say the same thing.

On the other side of the coin, is polysemy. Polysemy is when you have one word but that one word has multiple meanings. So you may have to differentiate between the different worse senses of a word like discharge. Are they talking about discharge from the hospital, or are they talking about stuff that comes out of your body?

And again, and a big challenge to NLP is the acronyms and abbreviations. And when you see APC in a text, what does it mean? Does it mean activated protein C? I mean, it just goes on, the number of meanings that, that acronym could have. And so, that's one of the big challenges in getting an NLP system to work the way that we as humans understand.

Okay, so now, I got you all excited and then I depressed you. So now, I'm going to conclude with you can still do it. There are a lot of challenges to NLP. Oh and I have a typo on there. So, see? You can do it.

Mike Dow: Yes. And that's another challenge of NLP is-

Wendy Chapman: That's right.

Mike Dow: Dealing with typos and other issues like that.

Wendy Chapman: Exactly. Alright, so let's talk through how do you really make this work without having a giant research NLP team working with you? Okay, so I wanted to divide the world into the types of fruit. There's high hanging fruit and there's low hanging fruit. And there's a lot of low hanging fruit that can be really useful and not require too much work for NLP to succeed.

Low hanging fruit looks like this. It has explicit mentions. When you're looking for chest pain, it says chest pain in the text. You understand the expected classes and the relation. You know that you're looking for diseases and anatomic locations. And the vocabulary's unambiguous. If you see the word stenosis, it always means stenosis. And if you see pneumonia, it means pneumonia.

So these are things that you can build simple systems, simple key-word base systems to extract that information. And I'm going to give three areas where we've applied natural language processing to this low hanging fruit and seeing some success.

But first, in looking at these areas, I want to ask the audience, if you are using NLP, what are you using it for? And I've given you four categories, but you could also have another category if I didn't include them all.

Sarah Stokes: Yeah, thanks Wendy. So our second poll question here, as Wendy said, we want to know what you're using NLP for. So that 33% who said you were using it, we want to know what you're using it for.

Your first option is quality, your second is research, your third option is decision support, fourth is predictive analytics and our last option is other. And we'll give you just a few moments to get your votes in there. That one takes a little more thinking than the last one.

And this is a good opportunity if you've already voted in the polls, to submit questions. If there's anything that's on top of mind, if you want to ask Wendy or Mike at the conclusion of today's presentation, now's a great time to submit that.

Okay, we'll give you just one more moment.

Okay, we're going to go ahead and close that poll and share the results. So 16% of you said that you are using NLP for quality, 23% said research, 18% said decision support, 27% said predictive analytics and 16% said other.

So no major winner there. Looks kind of even spread across the board. Is there one that you thought would've been more primarily or did you kind of expect a spread?

Wendy Chapman: Well, I'm a bit surprised at the amount of research and probably I don't understand the complete makeup of the audience, so that's more research than

I would've expected. Predictive analytics is the area that health systems are really trying to get into and so that's not surprising that, that's up at the top.

I think quality's a big area of opportunity for NLP, but there are a lot of different ... sometimes approaching quality, you want to do something that's broad based and can address all of your quality needs and you have to start small and maybe one project and that's probably what people are doing.

And decision support's a bit harder because you're usually talking about the point of care and you have higher expectations of the output. And so, the mistakes that NLP makes are often not tolerated in a decision support model. I'd be really interested in the other categories. And so I hope maybe there's a way to-

Mike Dow: Maybe we'll get some questions-

Wendy Chapman: Maybe we'll get some questions about that.

Sarah Stokes: Yup. Comments.

Mike Dow: Yeah and I just want to add a little bit, so the research and predictive analytics coming up on top made sense to me when it first came up. We've seen a lot more recently where organizations for that chart abstraction facility and chart abstraction research. That's a huge investment in time that takes place in research and just basic key word search and some of the negation can really help with automating or semi-automating that chart abstraction. So that makes sense.

And predictive analytics, obviously a lot of that information that's needed for predictions isn't necessarily in structured data-

Sarah Stokes: Yeah.

Mike Dow: So, makes sense.

Wendy Chapman: Yeah and I think what you brought up of ... researchers are a great audience to bring NLP to because they've had no help in the past and anything you give them helps them and they're not as picky about mistakes because it's not as critical that it's perfect.

Sarah Stokes: Okay. Just have to click back in there.

Wendy Chapman: Alright. I just keep forgetting what button. Okay. Okay. So, let me talk about some of the projects that we've done in these categories. The first is in quality. So at the University of Utah Health System, we generate ... we use NLP, and we process all of the relevant radiology reports every week. And then we're looking for pulmonary embolism and venous thrombosis in these reports.

And we generate ... we run over our data warehouse and generate a report using the business object system that we have and it lists all of the PEs and VTEs that have occurred over that week. And the chief of the division that looks at this for quality, says it's about 1.5 times more accurate than ICD codes, but we also give a lot more information.

What's the certainty, what side did it occur on? And it makes it really much more informative to understand their PE and VTE detection rate.

So that's one of the only things that I've worked on that's actually running in the hospital and being used outside of a research area.

This is one that's exciting with Andrew Gawron who I introduced earlier, that he's developed this report card for colonoscopy physicians, colonoscopists. And this is a mock-up of the report card that he's about to deploy across the United States through the VA system. And it shows each ... it shows the physician their adenoma detection rate over time and each different physician, so they can compare against each other.

And he is going to do a randomized clinical trial to see if this improves the adenoma detection rate for physicians to give them the speed back.

The secondary, and this is probably the area I've spent most of my time on is in assisting researchers. And so we talked about the project with Salome and being able to decrease chart review by 70%. We also worked with group health in the past doing this for pneumonia and decreased the chart review by 90%.

And it was really powerful because to do that research at Group Health, they hired two people full-time for two years to do the chart review and that's just not feasible. There's so many research studies that go undone or that are underpowered. And if you can really use NLP to filter the data in the beginning and then highlight information and make the chart review faster for the ones that need to be reviewed by a human, it can be a great boom.

The third area is decision support. And we are just getting into this. At Intermountain Healthcare they do a lot of NLP for decision support. In my world, we're just starting this. So, this is a project led by Ken Kawamoto in our department at the University of Utah. Where they are trying to flag patients from the EMR that have a family history of breast or colorectal cancer. And the idea is that there are a lot of patients.

We know the ages that are recommended but if you have a family history, you should get screened earlier. So, we're mining the EMR, finding those patients who have a first or second degree relative with breast cancer or colon cancer that started before they were 45, the onset, and then when we find those patients, we go into the patient portal and we email them and we tell them you

have a higher likelihood of getting breast cancer because you have this aunt and your mother had breast cancer when they were young.

And so we would recommend that you talk to a genetic counselor. And then we make it easy for them to make that appointment and so that they can have that. I think we have about eight people. We just launched this about six weeks ago, and we've had eight people in with the genetic counselors so far.

How well does the NLP work? You know it's not perfect. If we're looking at the age of onset for the relatives, it's pretty good. It's got high recall and high precision in the low to mid 90s. If we were looking at when they were deceased, the age and the range, then it's a little bit lower.

But you can really put these things in place even with imperfect performance, depending on how you're going to use them. I mean if we were going to alert a physician at the bedside, it might require higher performance, but the genetic counselors are going to look over this before we send the email to the patients. And so, it's this combination of human and NLP system working together that really works I think.

And so, with this we've been able to show that there is structured data about family history and we did kind of a socio-technical study to look ... we went to all of the University of Utah clinics across the state and observed and interviewed people to see where are they putting this data, how is it being captured?

So we capture the structured, but we capture about 50% more information from the text.

Okay, so there's some low hanging fruit ideas where you really can make a difference and do something exciting and meaningful with simple techniques.

There's higher hanging fruit that we're working on and these are things that are much more complicated and what higher hanging fruit looks like is it requires inference. If you're looking for social support, the low hanging fruit instance would be ... it says, "Patient has social support." But you just don't see that very often. In fact, we only see that about 5% of the time.

The other times, it says things like, "The brother's at the bedside." And you have to make an inference so they have social support. It has ambiguous vocabulary. So, brother and bedside in the first sentence, mean that they have social support, but in other instances, those words don't have anything to do with social support. And so you have to be able ... you'll get a lot of false positives and have to filter that out.

And the semantic roles that people are playing in the sentences matters. Who's the subject and who's the object? Because a key word system that's just looking

across a sentence, is not going to be able to differentiate between these two sentences, which have completely opposite meanings.

And so we've been looking at social risk factors to help predict readmission because they have been shown to increase the predictive capability, and they also give you a potential opportunity to intervene in some way for some of them. And so, these are some of the types of information that we've been building an NLP system called Moonstone, to extract. And we get performance between 75 and 98%, depending on the concept.

Alright, so what does this all mean? Well by the way, I think many of you probably know that Amazon just released a clinical NLP tool. You can go on to their website if you're a customer, and you can play with the tool and see what it does. And that will really kind of help you understand what do we mean when we say clinical NLP? It's pretty fun to input text and see what it can do and what it can't do.

There are a lot of tools out there that are available, that you can download for free, that are opensource to do classification, to find name identities like chess pain, and to look for the contextual information like the negation and the family history in them. And so there are tools out there that you can piece together and start with.

And I think really, the identification of the information from the text, is really just the first step. To really make NLP useful in healthcare, it has to be part of an application. Whether it's the decision support application, or the predictive analytics, and integrating it into that application, it takes knowledge beyond the NLP. And the hardest part is really how do you take that application now and put it into the workflow so that it can be used because technology is just technology and we've seen a lot of technology fail because we haven't really thought how can the user make use of this in the best way?

And so it's really NLP is this tiny part of this bigger picture to be able to think about how we can apply NLP in our health systems and our labs. Find a partner. There's a lot of people that do NLP. We're happy, excited research partners. Health Catalyst is doing natural language processing and there are a lot of people out there that you can work with.

And let's remember that challenges and opportunity and if we don't get onto that mountain and try it out and fall a few times we're not going to get where we want to go. I will conclude with that, and let's Sarah, ask the final question.

Sarah Stokes:

Yeah. Thank you Wendy. That was wonderful. Before we dive into our Q&A session here we just have this final poll question for you. While today's topic was an educational webinar focused on the challenges and potential of NLP in healthcare some of you may want to learn more about Health Catalyst products or professional services or the work we're doing in this space. If you'd like

someone from Health Catalyst to follow up with you please answer this poll question.

And once you've voted, yet again, we encourage you to get those questions in. We've had a lot of great ones trickling in throughout the presentation. And I'm actually going to turn the time over to Mike who's going to manage the Q&A here with Wendy.

Mike Dow: Great. Thanks Sarah. Yeah, Dr. Chapman that was great. That was the first time I heard the presentation in full in its allotted amount of time. We did a brief overview last week and a lot of that really resonated with me. Just to recap, the main things that you talked about were the potential of NLP and the hype of NLP and then some of the challenges that are really out there, and then capped us off with what can we do? What is achievable today?

So going into that what we can do today, the discussion around focusing on that low hanging fruit I think is ... I just want to reemphasize that because a lot of conversations that I hear around NLP, whether it be conversations with health systems about what our capabilities are as an analytics vendor or what's going on in research, there's a lot of focus at the really high end of NLP, the really high effort, higher yield that you're discussing around in that challenges section.

But that's maybe not where the most value is for your time spent. I just want to encourage our listeners to really think about not what can NLP do, and can it solve these really complex use cases, but what can keyword search and maybe negation do to help researchers. Another part of it that you mentioned I thought was really salient and I've had a number of conversations recently about this is it's...well, let me put it this way. So, one of the terms I've heard recently is the acronym AI shouldn't stand for artificial intelligence, excuse me. So everyone who's on the call I have a cold so it's a little bit difficult for me to speak today, and thank you for all the people in the room with us today bearing with me and my cold.

So, instead of artificial intelligence we should be thinking about AI as augmented intelligence, meaning that human in the loop so to speak, where NLP can provide an indicator of what might be happening, and human can review it. For example, when that researcher's case reducing chart review by 70%. That's a huge win. We might think well, why don't you reduce it to zero, but if you can reduce it by 70% that's a really huge win.

Wendy Chapman: Yeah.

Mike Dow: In the challenges section I wanted to maybe add a couple challenges that I've observed over the last few years and get your thoughts on it. So, one of the big challenges that I've observed, really over the last 18 months, is the overall infrastructure for NLP. There are different technologies than most data ecosystems and integrating that into a data ecosystem is a technology challenge

that has nothing to do with NLP. Deploy an elastic search cluster and optimizing that has nothing to do with NLP or solar or whatever your search index is. But that complicates the use of NLP. Is that something, it's a little off topic of what you're focused on, but have you observed that as well?

Wendy Chapman: Yeah, that's why I feel like NLP is such a small part of the whole cycle of what you're doing. You have to get the data, you have to...I mean this is the whole idea about data science. Just getting the data and cleaning the data and getting it ready to even run the sophisticated tools over is a huge burden and difficult. And then how you mark up the text. You might be able, if you use Solar or things like that to markup key words and map to vocabularies, that can be a really helpful first step. You don't necessarily have to do that but there are so many ways to approach it.

Mike Dow: Yeah, absolutely, and that's what you mentioned about the machine learning ecosystem and that's something that our president of technology talks about a lot is in machine learning, it's not just machine learning. That's a small part of the overall thing and he references a Google paper where they have a great illustration of the machine learning ecosystem and one small box in the middle is the actual machine learning model.

Wendy Chapman: Yeah.

Mike Dow: Excuse me. All right, so one of the things that I wanted to mention and go back to was Amazon Comprehend Medical. You mentioned that a couple minutes ago. Announce last week, it promises, according to the website, to extract information from unstructured medical text accurately and quickly with no machine learning experience required. So, I want to get your take on two questions. So, what would you say is the most exciting aspect of it and what should we be skeptical of?

Wendy Chapman: This is kind of like the helping clinicians work at the top of their license.

Mike Dow: Mm-hmm (affirmative)-hmm.

Wendy Chapman: There's been so much work on name identity recognition in marking up terms, negation, we don't need to reinvent those things. And so, because you're going to spend so much more time thinking about how do I really integrate this into the application and time on the front end about getting the data and getting it ready for use. If the more tools that you have available to really leverage without having to reinvent the wheel the more success you're going to have. And so I think it's really exciting. It does make me a little nervous, but I went like, 'Am I going to have a job in 10 years?' (laughter) But I really think the key to NLP in healthcare for me, marking up the words, marking up the negation, that's just...it's very low level.

It's really mapping that information in the text to your phenotype; that is a real challenge. And then scaling it out, making inferences from it across, you've got the mentions like I said, and then you've got the document and then you've got the encounter. So you have to be able to make inferences there. There's so many steps beyond that that I think it's great that we can do that, that there are more tools becoming available for us.

Mike Dow: Great.

Wendy Chapman: And I would be skeptical of a few things. If you play with it it seems to be really good at negation. It doesn't map the terminologies; it'll mark dyspareunia, it'll mark shortness of breath, but you won't know they're the same thing, and that can cause some noise when you do machine learning on top of it. And this is all just from my experience on it. And it doesn't do family history or temporality; there's so much left to do, and so I think there's a ton of work in NLP and the more people we have working on it the better.

Mike Dow: Yeah. And one other thing I'd say about the Amazon announcement, there have been a couple blogs I've seen in the last few days really that people have gone into it and provide examples of their use cases of it, so you can get reactions to it that are a little bit unbiased outside of Amazon of what it's capable of and what it works well for.

Wendy Chapman: Yeah.

Mike Dow: All right. So, the other area before we jump into the official Q&A that I want to ask about is you emphasize the importance of embedding NLP into applications and those apps into existing workflows. One of the things that I try to think about when we're looking at NLP use cases is not to focus on what we're trying to find, like we're looking for this concept and this type of note, but rather ask the question of what are we trying to accomplish. For example, we're trying to reduce the amount of chart abstraction for researchers or we're trying to increase the detection rate for certain type of cancer to enable a lower mortality rate. Does that resonate with you in terms of...I mean, have you seen that where people focus on the use case or focus on the technical problem without really understanding how that's going to fit in?

Wendy Chapman: Yeah, that really resonates with me and so collaborating with the people who are going to use it and thinking...You know, when we did the research studies of decreasing the chart review there's just a lot of questions to ask about when do you trust the NLP system and you can just say, 'don't even look at this report' and when do you need to look at it. That takes a lot of input from the people who'd be using it to trust it. And I think the trust part is huge in medicine and so, yeah, I completely agree with that.

Mike Dow: Great. Well, we have a number of questions that have rolled in so after scanning these over the last couple minutes one of the ones that just kind of pops out to

me, I've had this question asked a number of times over the last couple of years, is how do you deal with the de-identification of notes when it comes to privacy?

Wendy Chapman: Yeah, that's a really good questions. Now some institutions don't require you to de-identify. We don't need to at the University of Utah. We get IRB approval, we don't have to de-identify and other institutions require it. There are some open source systems that are available for that. There's one by Midar. I can't remember what it's called. There's a new system that's built with deep learning that was put out by researchers at MIT that's really high performing. Now, you have to train the data; you have to train these systems, and there are commercial systems too that are available. So, there are ways to do that, but you might first ask, do you really need to do it?

Mike Dow: Yeah, absolutely, that was one of the things in the back of my head is that there are use cases where you absolutely need to based on data usage agreements, but a lot of times you can actually use the raw text.

Wendy Chapman: Yep.

Mike Dow: All right, great. So, one of the other kinds of general themes in the questions asked about the skill sets that are needed in order to do NLP. That's probably a pretty broad question but can you kind of boil it down to some general things that people should be focused on in terms of skill set?

Wendy Chapman: Yeah, I've been thinking about that because now in our training program we have a data science track and NLP is part of that focus. And so I think you have to have the data science skills to be able to read in the data and manipulate it first of all. You have to be willing to really dive into the data and understand it. It can't be like the black box kind of idea, you're just going to run machine learning over it. That just doesn't work very well in general.

You have to understand the sections in the text and what they mean, the nuances of what's going on, and so you see this as sometimes as industry hops into this area and they bring a lot of computer science expertise, but they haven't thought about family history and who's experienced it and all that. And so, it's a slower ramp up, so understanding the sub language is really important. And then I think there are rule based methods, there are statistical methods and really the most successful are usually hybrids, and so I do think that you have to be able to find that balance of where the rule base is important and where machine learning helps.

Using the features...sometimes you can generate features from a rule based system and those can now be used in a machine learning system and you can show that that performs better than machine learning on text alone, or the rule based system by itself, so that balance, being able to do some machine learning also, just general ability to evaluate and knowing about good study design and how you need to have a blind test set....if you just read a lot of papers you'll see

what you really need. I don't think, you don't have to have a degree in linguistics or a degree in math. My background is linguistics and I think that's been helpful although I don't do anything really strongly linguistic in my work and the tools that I've developed have been very simple that anyone can apply.

I don't do a lot of statistical and machine learning with it either, but the skills you bring to it will be useful and you probably will get partners that have complementary skills.

Mike Dow: Yeah. The way I heard that was it's a combination of skills and so I think of Max Taggart who we hired after graduating from the University of Utah. He has some clinical training, he has machine learning training, he has a lot of different things kind of in the NLP world, but also a lot of programming and some data experience. And so, it's like a number of different things all feed into the ability to contribute in that area.

Wendy Chapman: Right, and Max is rare. He's one person that has all those skills.

Mike Dow: Yep, I know (laughter).

Wendy Chapman: Yeah, you might have to have a team.

Mike Dow: Sure, that's fair. All right, so I'm just skimming through the questions and so, one of the other themes in some of these questions are what types of technology, so what types of statistical tools and packages do you use, what are some of the open source tools that you'd recommend for NLP. That might be something we follow up with afterwards, but just quickly, what comes to mind there?

Wendy Chapman: Okay, so I'll list a few that you could play with. I've developed a tool called Context and there's a Python version and a Java version available and it's very simple. It's the rule based but it does negation. It looks for targets and the modifiers and it's simple, but it doesn't mean it's simple to use because it's all in Python or in Java and you have to go in and do things to it. You have to be able to program. But that's a simple tool that can be used.

There's a tool called Clamp and this was developed by Washu, who's at UT Houston. This is a really nice tool set that's just come out that is a drag and drop kind of system and you can leverage machine learning and lexicons in it and so I'd highly recommend that for people that don't have as much programming expertise.

Sea Takes, and a lot of people have asked about Apache Sea Takes, that's a very powerful tool that has a lot of different modules in it. So Sea Takes is awesome, it's a steep learning curve because it's in Uema and you have to really learn how to use Uema to be able to leverage Sea Takes. But it's powerful and a lot of people have it. As far as statistical, a lot of the tool kits that are out there just

Psych It Learn, Wecca, those kind of things, you have to feed features and you build those features through NLP or text processing.

Somebody also asked about ontologies and mapping your text to ontologies or to terminologies can be important for decreasing the sparse feature space and for decreasing that noise. The representation that I showed you, we represent that as an ontology, it's a different kind of ontology. I think when you get to scalability ontologies are important. If you're doing the low hanging fruit, a project here and a project there, it tends to feel like overkill. So those are my thoughts about tools.

Mike Dow: Great. I think we probably have time for just one more question. You started the presentation talking about the challenges of healthcare, the cost of healthcare and the challenges of EMR usability. So one of the questions that came up is which EMRs are using NLP and, I might add a little bit to that in terms of how are they using it and what techniques are they using?

Wendy Chapman: Yeah. I don't have a good answer to that, but I do think that all the EMRs are starting to embed NLP capabilities and some of the areas would be for search, to help use search better, but I think it's new in the EMR space and I haven't had a lot of experience with it.

Mike Dow: Yeah, yeah. In my experience with a handful of EMRs it's a relatively new feature in the last three to five years depending on the major EMR vendor, and the search is just a really good first step. I think all major EMR vendors have that now and had it for a couple years. Key word search doesn't sound like a really sophisticated thing, but like we talked about before, some really simple techniques like key word search get you a pretty long way. Doesn't get you hundred percent accuracy but it doesn't tackle synonyms, it doesn't tackle a lot of things, but it gets you maybe 50 percent of the way there and mitigates that harder stuff that will take longer and take more effort.

Wendy Chapman: Yeah, and I think an exciting area is being able to build applications on your own and then integrate them into your EMR using Fire standards and so I think in the future we don't necessarily have to rely on EMR companies. If we're doing something cutting edge we are able to integrate that.

Mike Dow: Absolutely.

Sarah Stokes: Okay, sounds like that's going to do it for today's Q&A. I'd like to thank Wendy and Mike for taking the time to share this presentation with us today and I want to thank everyone in the audience for joining us.