

[COVID Information Commons \(CIC\) Research Lightning Talk](#)

Transcript of a Presentation by Jayavanth Shenoy (Onai) January 30, 2024



Title: [Scaled Medical Records Analysis](#)

[Guha Jayachandran CIC Database Profile](#)

NSF Award #: [2028008](#)

[YouTube Recording with Slides](#)

[Winter 2024 CIC Webinar Information](#)

Transcript Editor: Shikhar Johri

Transcript

स्लाइड 1

मेरा नाम जयवंत है और मैं ओनाई में इंजीनियर हूँ। आज, मैं एनएसएफ से वित्त पोषण के साथ मेडिकल रिकॉर्ड के स्केल किए गए विश्लेषण पर हमारी परियोजना के बारे में बात करूंगा।

स्लाइड 2

समस्या यह है कि मेडिकल रिकॉर्ड बहुत महत्वपूर्ण जानकारी हैं, लेकिन हमारे पास बड़े पैमाने पर उनका अध्ययन करने का एक अच्छा तरीका नहीं है। परंपरागत रूप से, उस अस्पताल के भीतर रिकॉर्ड पर अलग-अलग अस्पतालों में अध्ययन हुआ। वहां भी, यह श्रमसाध्य है, निवासियों और छात्रों को मैन्युअल रूप से स्प्रेडशीट में डेटा ट्रांसक्रिप्ट करना है। यह प्रक्रिया भी वास्तव में धीमी है। COVID-19 जैसी स्थितियों में, यह अस्वीकार्य है - हम तब तक इंतजार नहीं करना चाहते जब तक कि एक अस्पताल के पास सांख्यिकीय रूप से महत्वपूर्ण अध्ययन होने के लिए पर्याप्त डेटा न हो। यही घाव हो रहा है - मेरा मतलब है, यही आमतौर पर होता है। लेकिन हम वास्तव में प्रसार के शुरुआती चरणों में परिणाम प्राप्त करना चाहते हैं। इसके लिए, हमें कुल मिलाकर कई अस्पतालों में अध्ययन करने में सक्षम होना चाहिए। वास्तव में बुनियादी प्रश्न हैं जो आपको लगता है कि इस दिन और उम्र में उत्तर देना आसान होगा, लेकिन वे COVID के लिए इतने आसान नहीं थे और आज भी नहीं हैं। उदाहरण के लिए, क्या रक्तचाप की दवा COVID में सुधार या बिगड़ती है? लाखों अमेरिकी इन रक्तचाप की दवाओं को ले रहे हैं, इसलिए आपको लगता है कि रक्तचाप की दवा लेने वाले और COVID के साथ अस्पताल में घायल होने वाले रोगियों के परिणामों के आधार पर इसका उत्तर देना एक आसान प्रश्न होगा। यह Google खोज क्वेरी जितना आसान होना चाहिए, लेकिन ऐसा नहीं है।

स्लाइड 3

इसलिए हमारा लक्ष्य कई स्वास्थ्य संस्थानों में एक कुशल तरीके से और गोपनीयता का त्याग किए बिना मेडिकल रिकॉर्ड पर तेजी से खोज को सक्षम करना था। हमारे उपकरणों के साथ, स्वास्थ्य सेवा संस्थान और दवा कंपनियां सांख्यिकीय रूप से महत्वपूर्ण आबादी तक बहुत तेजी से पहुंच सकती हैं और अनुसंधान तुरंत किया जा सकता है। यह अधिक सुरक्षित भी है। तो हम दोनों दुनिया के सर्वश्रेष्ठ प्राप्त कर सकते हैं? क्या हम अनुसंधान के लिए डेटा का एक बड़ा पूल प्राप्त कर सकते हैं और साथ ही, सुरक्षा को अधिकतम कर सकते हैं?

स्लाइड 4

हाँ, हम इस क्रिप्टोग्राफिक तकनीक के साथ कर सकते हैं। इस लक्ष्य को प्राप्त करने की तकनीक एक जादुई तकनीक है जिसे बहुदलीय गणना कहा जाता है। यह पार्टियों के एक समूह को एक दूसरे या किसी और को अपने इनपुट को प्रकट किए बिना अपने इनपुट पर संयुक्त रूप से कंप्यूटर परिणाम देने की अनुमति देता है। इसका एक उदाहरण याओ की करोड़पति समस्या कहा जाता है। एंड्रयू याओ ने इसे 1982 में पेश किया और यह हमारे दो करोड़पति ऐलिस और बॉब हैं। वे अपनी वास्तविक संपत्ति का खुलासा किए बिना जानना चाहते हैं कि उनमें से कौन अधिक अमीर है। तो आपके पास दो लोग हैं, ऐलिस और बॉब। वे अपने स्वयं के धन को जानते हैं, लेकिन दूसरे व्यक्ति को बताना नहीं चाहते हैं। और मान लें कि वे किसी तीसरे पक्ष को शामिल नहीं करना चाहते हैं, इसलिए कोई आईआरएस या कुछ भी नहीं। क्या आप उनके लिए यह पता लगाने का कोई तरीका सोच सकते हैं कि उनमें से कौन अधिक अमीर है? खैर, मानो या न मानो, यह पता चला है कि ऐसा करने का एक तरीका है। एक अन्य उदाहरण के रूप में, मान लें कि हम इस वेबिनार में सभी की औसत सामाजिक सुरक्षा संख्या की गणना करना चाहते हैं, लेकिन हम अपने सामाजिक सुरक्षा नंबर को एक-दूसरे के सामने प्रकट नहीं करना चाहते हैं। क्या कोई तरीका है जिससे हम इसे कर सकते हैं? हां, सुरक्षित एमपीसी के साथ है। हमने स्वास्थ्य रिकॉर्ड के लिए एक ही अवधारणा लागू की है और एक क्वेरी का परिणाम कई अस्पतालों में गणना की जाती है, बिना अस्पतालों को किसी भी रिकॉर्ड को प्रसारित करने या प्रकट करने की आवश्यकता होती है। हमने अपनी परियोजना में कुछ अलग-अलग साइटों पर इसका प्रदर्शन किया।

स्लाइड 5

यहां उन प्रश्नों के कुछ उदाहरण दिए गए हैं जिन्हें पूछने में रुचि हो सकती है, जैसे कि 70 से अधिक रोगियों के लिए अस्पताल में रहने की औसत लंबाई, या वेंटिलेटर पर बहुत कम उम्र के रोगियों का कितना प्रतिशत घायल हो गया।

स्लाइड 6

अधिक क्रिप्टोग्राफिक रूप से समझदार दर्शकों के लिए, हमने एनपीसी तकनीक के लिए एक शुभंकर का उपयोग किया। मैं अभी विवरण में नहीं जाऊंगा, लेकिन यदि आप रुचि रखते हैं तो आप मुझसे बाद में इस बारे में पूछ सकते हैं। केवल एक चीज जिसका मैं यहां उल्लेख करूंगा वह यह है कि यह प्रोटोकॉल बेईमान बहुमत के खिलाफ सुरक्षित है, जिसका अर्थ है कि यह प्रोटोकॉल सुरक्षित है, भले ही अधिकांश प्रतिभागी दुर्भावनापूर्ण हों।

स्लाइड 7

मैं दोहराना चाहता हूँ कि हमारे दृष्टिकोण के साथ, कोई भी डेटा कभी भी अस्पताल नहीं छोड़ता है। प्रत्येक अस्पताल का डेटा उसके विभाग में पूरी तरह से साइलो है। आईटी विभागों को सुरक्षा के बारे में चिंता करने की ज़रूरत नहीं है, इतना ही नहीं कुछ भी रिवर्स इंजीनियर नहीं किया जा सकता है। कम्प्यूटेशन का समन्वय ब्लॉकचेन या वितरक खाता बही के माध्यम से होता है।

स्लाइड 8

यह ढांचा अधिक न्यायसंगत नैदानिक परीक्षणों को भी सक्षम बनाता है। परंपरागत रूप से, रोगी आबादी और नैदानिक परीक्षण बहुत विविध नहीं रहे हैं। एफडीए और अन्य की प्राथमिकताओं में से एक यह है कि और इस तकनीक के साथ, रुचि के रोगियों वाली साइटों को स्थित किया जा सकता है, भले ही वे ऐसे संस्थान न हों जो एक कंपनी आमतौर पर काम करती है।

स्लाइड 9

तब से, हमने इस दृष्टिकोण को क्वेरी करने से लेकर अलग-अलग साइटों पर AI मॉडल को गोपनीयता-संरक्षण तरीके से प्रशिक्षित करने तक बढ़ा दिया है। हम गोपनीयता-संरक्षण दवा की खोज के लिए एनआईएच के साथ भी काम कर रहे हैं। यहां हमारी अंतिम दृष्टि बायोमेडिसिन में गोपनीयता-संरक्षण खुफिया नेटवर्क है।

स्लाइड 10

सुनने के लिए धन्यवाद और मैं इस अवसर के लिए सीआईसी और एनएसएफ को हमारी परियोजना, देश भर में हमारे सहयोगियों और ओनाई में मेरे सहयोगियों को वित्त पोषित करने के लिए धन्यवाद देना चाहता हूँ। कृपया मुझे बताएं कि सत्र के अंत में आपके कोई प्रश्न हैं या नहीं। फिर से धन्यवाद।