

WEBVTT

NOTE duration: "01:28:33.877"

NOTE Confidence: 0.5274791

00:00:01.120 --> 00:00:01.620 So

NOTE Confidence: 0.98047274

00:00:02.053 --> 00:00:02.553 in

NOTE Confidence: 0.9973992

00:00:02.986 --> 00:00:03.486 the

NOTE Confidence: 0.9925713

00:00:03.919 --> 00:00:04.960 previous session, I think,

NOTE Confidence: 0.94184804

00:00:06.559 --> 00:00:08.160 we mainly focus on to

NOTE Confidence: 0.94184804

00:00:08.160 --> 00:00:09.679 show you how to access

NOTE Confidence: 0.94184804

00:00:09.679 --> 00:00:11.119 the data, how to request

NOTE Confidence: 0.94184804

00:00:11.119 --> 00:00:12.900 computing resource. Right?

NOTE Confidence: 0.9852571

00:00:13.265 --> 00:00:14.545 In the next session, what

NOTE Confidence: 0.9852571

00:00:14.545 --> 00:00:15.745 we will do is more

NOTE Confidence: 0.9852571

00:00:15.745 --> 00:00:16.245 about

NOTE Confidence: 0.9181464

00:00:16.864 --> 00:00:18.085 technical details,

NOTE Confidence: 0.95738536

00:00:18.945 --> 00:00:21.125 how you can call existing

NOTE Confidence: 0.95738536

00:00:21.185 --> 00:00:22.165 large engine models

NOTE Confidence: 0.909582

00:00:23.185 --> 00:00:25.364 in the CHP safe environment.
NOTE Confidence: 0.98965263

00:00:26.070 --> 00:00:27.590 Or if you want to
NOTE Confidence: 0.98965263

00:00:27.590 --> 00:00:28.810 build your own
NOTE Confidence: 0.63741624

00:00:29.430 --> 00:00:31.850 customized large engine model. Right?
NOTE Confidence: 0.9151659

00:00:33.670 --> 00:00:35.350 By leveraging is it existing
NOTE Confidence: 0.9151659

00:00:35.350 --> 00:00:36.550 model by using your own
NOTE Confidence: 0.9151659

00:00:36.550 --> 00:00:38.070 data, then how can you
NOTE Confidence: 0.9151659

00:00:38.070 --> 00:00:39.510 train the model using the
NOTE Confidence: 0.9151659

00:00:39.510 --> 00:00:41.795 environment? So that's what will
NOTE Confidence: 0.9151659

00:00:41.795 --> 00:00:43.255 happen in the next session.
NOTE Confidence: 0.9597303

00:00:43.715 --> 00:00:46.115 And I will provide very
NOTE Confidence: 0.9597303

00:00:46.115 --> 00:00:46.615 brief
NOTE Confidence: 0.8095837

00:00:46.995 --> 00:00:49.075 overview about large engine model,
NOTE Confidence: 0.8095837

00:00:49.075 --> 00:00:49.575 then,
NOTE Confidence: 0.98716235

00:00:50.035 --> 00:00:51.235 a number of speaker will
NOTE Confidence: 0.98716235

00:00:51.235 --> 00:00:52.515 actually jump in to show

NOTE Confidence: 0.98716235
00:00:52.515 --> 00:00:53.655 the different tools
NOTE Confidence: 0.96264327
00:00:54.510 --> 00:00:55.950 that are available on the
NOTE Confidence: 0.96264327
00:00:55.950 --> 00:00:57.790 CHP safe environment, which you
NOTE Confidence: 0.96264327
00:00:57.790 --> 00:00:59.390 can use. Okay? For example,
NOTE Confidence: 0.96264327
00:00:59.390 --> 00:01:00.510 if you want to annotate
NOTE Confidence: 0.96264327
00:01:00.510 --> 00:01:01.890 data, we have a tool.
NOTE Confidence: 0.96264327
00:01:02.110 --> 00:01:03.150 If you want to train
NOTE Confidence: 0.96264327
00:01:03.150 --> 00:01:03.870 the model,
NOTE Confidence: 0.8540691
00:01:04.350 --> 00:01:06.110 call existing API like a
NOTE Confidence: 0.8540691
00:01:06.110 --> 00:01:07.230 keyway tool, how to call
NOTE Confidence: 0.8540691
00:01:07.230 --> 00:01:09.385 it, and how to stuff
NOTE Confidence: 0.8540691
00:01:09.385 --> 00:01:10.905 on using your own data
NOTE Confidence: 0.8540691
00:01:10.905 --> 00:01:11.944 to train them out. So
NOTE Confidence: 0.8540691
00:01:11.944 --> 00:01:13.385 that's what will happen next.
NOTE Confidence: 0.8540691
00:01:13.385 --> 00:01:14.185 Let me just do a
NOTE Confidence: 0.8540691

00:01:14.185 --> 00:01:14.685 quick,
NOTE Confidence: 0.7731308
00:01:15.065 --> 00:01:15.565 introduction
NOTE Confidence: 0.9284867
00:01:16.584 --> 00:01:17.865 of the large engine model.
NOTE Confidence: 0.9284867
00:01:17.865 --> 00:01:19.305 I think because of running
NOTE Confidence: 0.9284867
00:01:19.305 --> 00:01:20.345 of time, so I'm trying
NOTE Confidence: 0.9284867
00:01:20.345 --> 00:01:21.545 to shrink,
NOTE Confidence: 0.9730878
00:01:21.944 --> 00:01:23.245 reduce my session.
NOTE Confidence: 0.97617203
00:01:32.170 --> 00:01:33.289 I think I'm going to
NOTE Confidence: 0.97617203
00:01:33.289 --> 00:01:34.670 skip this. I will
NOTE Confidence: 0.9609124
00:01:35.049 --> 00:01:36.329 was planning to give a
NOTE Confidence: 0.9609124
00:01:36.329 --> 00:01:38.775 short history about AI. Basically,
NOTE Confidence: 0.9609124
00:01:38.775 --> 00:01:39.575 what I'm trying to say
NOTE Confidence: 0.9609124
00:01:39.575 --> 00:01:41.334 is not just starting now.
NOTE Confidence: 0.9609124
00:01:41.334 --> 00:01:42.854 We have been doing this
NOTE Confidence: 0.9609124
00:01:42.854 --> 00:01:44.215 for many times, but this
NOTE Confidence: 0.9609124
00:01:44.215 --> 00:01:45.735 wave of generative AI is

NOTE Confidence: 0.9609124
00:01:45.735 --> 00:01:47.575 slightly different from the previous
NOTE Confidence: 0.9609124
00:01:47.575 --> 00:01:49.174 one in a few ways,
NOTE Confidence: 0.9609124
00:01:49.174 --> 00:01:50.875 like, how the model was,
NOTE Confidence: 0.9222265
00:01:52.545 --> 00:01:54.100 trained. It's much bigger than
NOTE Confidence: 0.9222265
00:01:54.100 --> 00:01:55.940 previous model and focus on
NOTE Confidence: 0.9222265
00:01:55.940 --> 00:01:57.700 generation task rather than trying
NOTE Confidence: 0.9222265
00:01:57.700 --> 00:01:59.620 to do, prediction or analyze
NOTE Confidence: 0.9222265
00:01:59.620 --> 00:02:01.860 the data and also heavy
NOTE Confidence: 0.9222265
00:02:01.860 --> 00:02:03.000 rely on the GPUs.
NOTE Confidence: 0.9226687
00:02:03.780 --> 00:02:04.900 And one thing I really
NOTE Confidence: 0.9226687
00:02:04.900 --> 00:02:05.940 want to mention because we
NOTE Confidence: 0.9226687
00:02:05.940 --> 00:02:06.980 talk about this is a
NOTE Confidence: 0.9226687
00:02:06.980 --> 00:02:08.755 large engine model session, and
NOTE Confidence: 0.9226687
00:02:08.755 --> 00:02:10.275 I'll just give a brief
NOTE Confidence: 0.9226687
00:02:10.275 --> 00:02:11.175 history about
NOTE Confidence: 0.9350986

00:02:11.475 --> 00:02:13.555 language models. Language model has
NOTE Confidence: 0.9350986

00:02:13.555 --> 00:02:14.514 been there for a while.
NOTE Confidence: 0.9350986

00:02:14.514 --> 00:02:15.735 Start nineteen sixty.
NOTE Confidence: 0.8961434

00:02:16.275 --> 00:02:18.115 Really a promising model trying
NOTE Confidence: 0.8961434

00:02:18.115 --> 00:02:19.235 to give a sequence of
NOTE Confidence: 0.8961434

00:02:19.235 --> 00:02:20.595 words, trying to predict next
NOTE Confidence: 0.8961434

00:02:20.595 --> 00:02:21.900 word. Okay. Then
NOTE Confidence: 0.76518255

00:02:22.200 --> 00:02:23.160 later on, neural,
NOTE Confidence: 0.909754

00:02:23.880 --> 00:02:25.560 language model showed very good
NOTE Confidence: 0.909754

00:02:25.560 --> 00:02:27.240 performance, but it suffers with
NOTE Confidence: 0.909754

00:02:27.240 --> 00:02:27.720 all those,
NOTE Confidence: 0.8063671

00:02:28.200 --> 00:02:28.700 computational
NOTE Confidence: 0.65543723

00:02:29.240 --> 00:02:29.740 efficiency
NOTE Confidence: 0.9542594

00:02:30.120 --> 00:02:31.560 and all those issues. So
NOTE Confidence: 0.9542594

00:02:31.560 --> 00:02:33.415 didn't really scale up. Until
NOTE Confidence: 0.9542594

00:02:33.415 --> 00:02:35.335 later, two thousand seventeen, the

NOTE Confidence: 0.9542594
00:02:35.335 --> 00:02:36.955 transformer model was proposed
NOTE Confidence: 0.83207774
00:02:37.655 --> 00:02:39.435 together with abundant of GPUs
NOTE Confidence: 0.83207774
00:02:39.575 --> 00:02:40.075 available,
NOTE Confidence: 0.94360423
00:02:40.455 --> 00:02:41.895 make it possible actually to
NOTE Confidence: 0.94360423
00:02:41.895 --> 00:02:42.395 train,
NOTE Confidence: 0.7439441
00:02:43.334 --> 00:02:43.834 neural
NOTE Confidence: 0.93808967
00:02:44.294 --> 00:02:45.575 language model with a lot
NOTE Confidence: 0.93808967
00:02:45.575 --> 00:02:46.694 of data. Then we move
NOTE Confidence: 0.93808967
00:02:46.694 --> 00:02:49.139 to this pretrained language models.
NOTE Confidence: 0.93808967
00:02:49.360 --> 00:02:50.639 Okay? At that time, like
NOTE Confidence: 0.93808967
00:02:50.639 --> 00:02:51.919 a BERT model, you probably
NOTE Confidence: 0.93808967
00:02:51.919 --> 00:02:53.280 heard a lot too. Right?
NOTE Confidence: 0.93808967
00:02:53.280 --> 00:02:54.800 Show the good performance that
NOTE Confidence: 0.93808967
00:02:54.800 --> 00:02:55.780 you you pretrain,
NOTE Confidence: 0.917651
00:02:56.560 --> 00:02:58.740 a reasonable text. But then
NOTE Confidence: 0.917651

00:02:58.960 --> 00:03:00.960 two thousand twenty two, really,
NOTE Confidence: 0.917651

00:03:00.960 --> 00:03:01.840 we moved to what we
NOTE Confidence: 0.917651

00:03:01.840 --> 00:03:03.375 call large language model.
NOTE Confidence: 0.98295563

00:03:03.935 --> 00:03:04.435 And,
NOTE Confidence: 0.8720233

00:03:05.055 --> 00:03:06.915 basically, it's a transformal based
NOTE Confidence: 0.8720233

00:03:06.975 --> 00:03:08.895 pretrained language model, but trained
NOTE Confidence: 0.8720233

00:03:08.895 --> 00:03:10.195 out a lot of data.
NOTE Confidence: 0.9082212

00:03:10.655 --> 00:03:12.735 And the the rationale behind
NOTE Confidence: 0.9082212

00:03:12.735 --> 00:03:13.455 this is,
NOTE Confidence: 0.9885584

00:03:14.014 --> 00:03:14.995 oh, I think I
NOTE Confidence: 0.9795842

00:03:15.455 --> 00:03:16.195 is this
NOTE Confidence: 0.8939684

00:03:16.630 --> 00:03:18.550 emergent phenomenon of a large
NOTE Confidence: 0.8939684

00:03:18.550 --> 00:03:20.470 range model. Because people find
NOTE Confidence: 0.8939684

00:03:20.470 --> 00:03:21.690 that when you train on
NOTE Confidence: 0.8939684

00:03:21.830 --> 00:03:22.810 a lot of data,
NOTE Confidence: 0.9103052

00:03:23.270 --> 00:03:24.870 the model not just can

NOTE Confidence: 0.9103052
00:03:24.870 --> 00:03:26.470 do one thing in a
NOTE Confidence: 0.9103052
00:03:26.470 --> 00:03:28.230 reasonable performing. It can actually
NOTE Confidence: 0.9103052
00:03:28.230 --> 00:03:29.350 do a lot of different
NOTE Confidence: 0.9103052
00:03:29.350 --> 00:03:31.850 tasks, always a reasonable performance.
NOTE Confidence: 0.9103052
00:03:32.085 --> 00:03:33.045 That's what we call the
NOTE Confidence: 0.9103052
00:03:33.045 --> 00:03:34.805 emergent phenomenon of the large
NOTE Confidence: 0.9103052
00:03:34.805 --> 00:03:36.245 language model. Suddenly, it become
NOTE Confidence: 0.9103052
00:03:36.245 --> 00:03:37.065 very smart.
NOTE Confidence: 0.9461985
00:03:37.685 --> 00:03:39.285 And that leads to the
NOTE Confidence: 0.9461985
00:03:39.285 --> 00:03:40.245 the the the a lot
NOTE Confidence: 0.9461985
00:03:40.245 --> 00:03:41.865 of development of all those
NOTE Confidence: 0.9461985
00:03:42.005 --> 00:03:43.765 large language model, open source
NOTE Confidence: 0.9461985
00:03:43.765 --> 00:03:44.265 models
NOTE Confidence: 0.86073595
00:03:44.565 --> 00:03:46.245 like a LLAMA, DeepSeq, you
NOTE Confidence: 0.86073595
00:03:46.245 --> 00:03:47.790 probably heard, which give you
NOTE Confidence: 0.86073595

00:03:47.790 --> 00:03:48.610 all the weights.
NOTE Confidence: 0.9659783
00:03:49.070 --> 00:03:50.190 You can actually use your
NOTE Confidence: 0.9659783
00:03:50.190 --> 00:03:52.370 own data to continue pretrain
NOTE Confidence: 0.978988
00:03:52.670 --> 00:03:54.590 or fine tune. Okay? And
NOTE Confidence: 0.978988
00:03:54.590 --> 00:03:56.290 commercial model like GPT,
NOTE Confidence: 0.9394568
00:03:57.390 --> 00:03:58.830 you they used to be
NOTE Confidence: 0.9394568
00:03:58.830 --> 00:04:00.350 closed model, so you cannot
NOTE Confidence: 0.9394568
00:04:00.350 --> 00:04:01.550 really fine tune. But now
NOTE Confidence: 0.9394568
00:04:01.550 --> 00:04:03.045 GPT also have a service.
NOTE Confidence: 0.9394568
00:04:03.205 --> 00:04:04.885 You can upload data, do
NOTE Confidence: 0.9394568
00:04:04.885 --> 00:04:06.185 some kind of fine tune
NOTE Confidence: 0.9394568
00:04:06.245 --> 00:04:07.444 on their side, then post
NOTE Confidence: 0.9394568
00:04:07.444 --> 00:04:08.724 the model. The fine tune
NOTE Confidence: 0.9394568
00:04:08.724 --> 00:04:10.084 model on the GPT side.
NOTE Confidence: 0.9394568
00:04:10.084 --> 00:04:11.045 You still have to pay
NOTE Confidence: 0.9394568
00:04:11.045 --> 00:04:12.165 every time you call them.

NOTE Confidence: 0.9394568
00:04:12.165 --> 00:04:14.084 Okay? And then there's also
NOTE Confidence: 0.9394568
00:04:14.084 --> 00:04:16.425 different architect encoder versus decoder.
NOTE Confidence: 0.926009
00:04:16.729 --> 00:04:17.690 And the the trend is
NOTE Confidence: 0.926009
00:04:17.690 --> 00:04:19.370 also we're moving more towards
NOTE Confidence: 0.926009
00:04:19.370 --> 00:04:21.610 multimodal large entry model. Instead
NOTE Confidence: 0.926009
00:04:21.610 --> 00:04:22.810 of just a text based
NOTE Confidence: 0.926009
00:04:22.810 --> 00:04:24.589 model, your text plus image,
NOTE Confidence: 0.926009
00:04:24.810 --> 00:04:26.410 text plus other genomic data,
NOTE Confidence: 0.926009
00:04:26.410 --> 00:04:27.370 and all those things. A
NOTE Confidence: 0.926009
00:04:27.370 --> 00:04:28.669 lot of things going on.
NOTE Confidence: 0.82617366
00:04:29.210 --> 00:04:29.710 And,
NOTE Confidence: 0.84698564
00:04:30.933 --> 00:04:33.426 in particularly, in the NLP
NOTE Confidence: 0.84698564
00:04:33.426 --> 00:04:33.924 world,
NOTE Confidence: 0.9958482
00:04:34.423 --> 00:04:36.915 especially in the biomedical NLP
NOTE Confidence: 0.9958482
00:04:36.915 --> 00:04:37.414 world,
NOTE Confidence: 0.9589079

00:04:37.912 --> 00:04:40.404 we often focus on one
NOTE Confidence: 0.9589079
00:04:40.404 --> 00:04:42.897 NLP task called information extraction.
NOTE Confidence: 0.9589079
00:04:42.897 --> 00:04:44.891 So the idea is,
NOTE Confidence: 0.9344058
00:04:45.550 --> 00:04:46.830 in the clinical data, there's
NOTE Confidence: 0.9344058
00:04:46.830 --> 00:04:47.950 a lot of unstructure there.
NOTE Confidence: 0.9344058
00:04:47.950 --> 00:04:49.070 For example, it's a text
NOTE Confidence: 0.9344058
00:04:49.070 --> 00:04:50.190 document and a lot of
NOTE Confidence: 0.9344058
00:04:50.190 --> 00:04:50.690 details.
NOTE Confidence: 0.94430923
00:04:51.390 --> 00:04:52.850 So the task of information
NOTE Confidence: 0.94430923
00:04:52.910 --> 00:04:54.510 extraction is like, okay. Given
NOTE Confidence: 0.94430923
00:04:54.510 --> 00:04:55.250 this document,
NOTE Confidence: 0.9966955
00:04:55.710 --> 00:04:56.910 can you extract all the
NOTE Confidence: 0.9966955
00:04:56.910 --> 00:04:58.830 disease information of the patient
NOTE Confidence: 0.9966955
00:04:58.830 --> 00:04:59.970 out of this document?
NOTE Confidence: 0.89721805
00:05:00.285 --> 00:05:01.325 So I would say this
NOTE Confidence: 0.89721805
00:05:01.325 --> 00:05:03.165 is about comes to seventy,

NOTE Confidence: 0.89721805
00:05:03.165 --> 00:05:04.385 eighty percent of,
NOTE Confidence: 0.9952812
00:05:04.685 --> 00:05:06.365 the the the requirements for
NOTE Confidence: 0.9952812
00:05:06.365 --> 00:05:07.505 a lot of, like,
NOTE Confidence: 0.9675229
00:05:08.045 --> 00:05:09.505 EHR based analysis,
NOTE Confidence: 0.96573627
00:05:10.445 --> 00:05:11.185 both for,
NOTE Confidence: 0.6576495
00:05:11.965 --> 00:05:12.465 practical
NOTE Confidence: 0.969087
00:05:13.005 --> 00:05:13.505 practices
NOTE Confidence: 0.9832907
00:05:13.805 --> 00:05:15.005 as well as for clinical
NOTE Confidence: 0.9832907
00:05:15.005 --> 00:05:15.505 research.
NOTE Confidence: 0.9494245
00:05:15.910 --> 00:05:18.150 So today, most all almost
NOTE Confidence: 0.9494245
00:05:18.150 --> 00:05:19.270 all the work we show
NOTE Confidence: 0.9494245
00:05:19.270 --> 00:05:20.390 here today is on this
NOTE Confidence: 0.9494245
00:05:20.390 --> 00:05:22.010 information extraction task,
NOTE Confidence: 0.9706254
00:05:22.390 --> 00:05:24.070 and it can further divide
NOTE Confidence: 0.9706254
00:05:24.070 --> 00:05:25.290 into three subtasks.
NOTE Confidence: 0.824318

00:05:26.150 --> 00:05:27.670 The first one is called
NOTE Confidence: 0.824318
00:05:27.670 --> 00:05:28.730 named entity recognition.
NOTE Confidence: 0.9664465
00:05:29.425 --> 00:05:30.545 So the idea is that
NOTE Confidence: 0.9664465
00:05:30.545 --> 00:05:31.425 you need to give a
NOTE Confidence: 0.9664465
00:05:31.425 --> 00:05:32.725 document. You need to,
NOTE Confidence: 0.9919987
00:05:33.265 --> 00:05:34.645 the system need to recognize
NOTE Confidence: 0.9919987
00:05:34.865 --> 00:05:36.725 MRI of the abdominal
NOTE Confidence: 0.95648456
00:05:37.425 --> 00:05:38.405 is a test.
NOTE Confidence: 0.87832105
00:05:38.945 --> 00:05:39.445 Okay?
NOTE Confidence: 0.8876491
00:05:39.745 --> 00:05:40.785 You need to know both
NOTE Confidence: 0.8876491
00:05:40.785 --> 00:05:41.825 the type is the test
NOTE Confidence: 0.8876491
00:05:41.825 --> 00:05:42.725 and the boundary.
NOTE Confidence: 0.8984965
00:05:43.060 --> 00:05:44.419 And you need to know
NOTE Confidence: 0.8984965
00:05:44.419 --> 00:05:46.020 June eighteen two thousand eight
NOTE Confidence: 0.8984965
00:05:46.020 --> 00:05:47.400 is a temple expression,
NOTE Confidence: 0.9492426
00:05:47.779 --> 00:05:49.220 that type of entity what

NOTE Confidence: 0.9492426
00:05:49.220 --> 00:05:50.419 type of entity and what's
NOTE Confidence: 0.9492426
00:05:50.419 --> 00:05:51.000 the boundary.
NOTE Confidence: 0.943485
00:05:51.380 --> 00:05:53.060 That's the MER task. The
NOTE Confidence: 0.943485
00:05:53.060 --> 00:05:54.660 second one, we also call
NOTE Confidence: 0.943485
00:05:54.660 --> 00:05:56.125 it relation instruction. So what
NOTE Confidence: 0.943485
00:05:56.125 --> 00:05:57.165 do you need to know?
NOTE Confidence: 0.943485
00:05:57.165 --> 00:05:58.365 You want to know this
NOTE Confidence: 0.943485
00:05:58.365 --> 00:06:00.285 June eighteen two thousand eight
NOTE Confidence: 0.943485
00:06:00.285 --> 00:06:02.065 is a modifier of MRI.
NOTE Confidence: 0.967909
00:06:02.764 --> 00:06:04.445 Right? That's a relation between
NOTE Confidence: 0.967909
00:06:04.445 --> 00:06:06.125 those two entities. So it's
NOTE Confidence: 0.967909
00:06:06.125 --> 00:06:07.324 very important for you to
NOTE Confidence: 0.967909
00:06:07.324 --> 00:06:08.385 recognize the context
NOTE Confidence: 0.9884232
00:06:08.764 --> 00:06:10.145 of that clinical entity.
NOTE Confidence: 0.9341382
00:06:10.669 --> 00:06:11.790 And the the third one
NOTE Confidence: 0.9341382

00:06:11.790 --> 00:06:13.089 is called concept normalization.
NOTE Confidence: 0.97934395

00:06:14.029 --> 00:06:16.190 So if you read the
NOTE Confidence: 0.97934395

00:06:16.190 --> 00:06:17.790 notes, what you saw here
NOTE Confidence: 0.97934395

00:06:17.790 --> 00:06:19.330 is renal cell carcinoma.
NOTE Confidence: 0.96911204

00:06:20.029 --> 00:06:20.830 But if you want to
NOTE Confidence: 0.96911204

00:06:20.830 --> 00:06:22.510 build a clinical decision support
NOTE Confidence: 0.96911204

00:06:22.510 --> 00:06:23.870 system, you need to be
NOTE Confidence: 0.96911204

00:06:24.110 --> 00:06:25.550 this entity need to be
NOTE Confidence: 0.96911204

00:06:25.550 --> 00:06:27.565 coded to a concept in
NOTE Confidence: 0.96911204

00:06:27.565 --> 00:06:28.705 the medical terminology.
NOTE Confidence: 0.9482559

00:06:29.165 --> 00:06:30.605 Could be ICD ten. Could
NOTE Confidence: 0.9482559

00:06:30.605 --> 00:06:32.205 be SNOMED. Right? So that's
NOTE Confidence: 0.9482559

00:06:32.365 --> 00:06:34.285 you want to normalize this
NOTE Confidence: 0.9482559

00:06:34.285 --> 00:06:35.185 detect entity
NOTE Confidence: 0.9830545

00:06:35.645 --> 00:06:37.404 to a term in the
NOTE Confidence: 0.9830545

00:06:37.404 --> 00:06:39.345 vocabulary in the standard vocabulary.

NOTE Confidence: 0.9406617
00:06:39.690 --> 00:06:40.650 And as you can see,
NOTE Confidence: 0.9406617
00:06:40.650 --> 00:06:42.350 it's actually not straightforward because
NOTE Confidence: 0.9406617
00:06:42.410 --> 00:06:43.950 you see renal cell carcinoma,
NOTE Confidence: 0.9406617
00:06:44.010 --> 00:06:45.330 but, actually, the term is
NOTE Confidence: 0.9406617
00:06:45.330 --> 00:06:47.150 a, malignant neoplasma
NOTE Confidence: 0.97249115
00:06:47.690 --> 00:06:49.710 of, kidney in the terminology.
NOTE Confidence: 0.97249115
00:06:49.930 --> 00:06:50.970 So you need this kind
NOTE Confidence: 0.97249115
00:06:50.970 --> 00:06:52.910 of mapping. Okay? So
NOTE Confidence: 0.9815793
00:06:53.595 --> 00:06:54.875 today, most of our work
NOTE Confidence: 0.9815793
00:06:54.875 --> 00:06:56.075 will show how we do
NOTE Confidence: 0.9815793
00:06:56.075 --> 00:06:57.755 those three tasks and build
NOTE Confidence: 0.9815793
00:06:57.755 --> 00:06:59.295 the system to extract information
NOTE Confidence: 0.9815793
00:06:59.355 --> 00:07:00.335 out of the text.
NOTE Confidence: 0.8944059
00:07:00.714 --> 00:07:02.635 I just primarily talk about
NOTE Confidence: 0.8944059
00:07:02.635 --> 00:07:03.455 three different,
NOTE Confidence: 0.9984642

00:07:04.635 --> 00:07:05.135 approach.
NOTE Confidence: 0.96651506
00:07:05.830 --> 00:07:06.870 I'll skip those.
NOTE Confidence: 0.997553
00:07:07.590 --> 00:07:08.630 I have a few slides
NOTE Confidence: 0.997553
00:07:08.630 --> 00:07:09.370 about history.
NOTE Confidence: 0.9265872
00:07:09.830 --> 00:07:12.150 Two thousand twin two thousand
NOTE Confidence: 0.9265872
00:07:12.150 --> 00:07:13.350 ish, we mainly work on
NOTE Confidence: 0.9265872
00:07:13.350 --> 00:07:14.950 rule based system. You have
NOTE Confidence: 0.9265872
00:07:14.950 --> 00:07:16.390 a dictionary. We try to
NOTE Confidence: 0.9265872
00:07:16.390 --> 00:07:17.590 look up all the disease
NOTE Confidence: 0.9265872
00:07:17.590 --> 00:07:19.595 from the dictionary. Okay? Then
NOTE Confidence: 0.9265872
00:07:19.595 --> 00:07:21.215 two thousand ten, we annotate
NOTE Confidence: 0.9265872
00:07:21.275 --> 00:07:22.555 corpus. We start to do
NOTE Confidence: 0.9265872
00:07:22.555 --> 00:07:24.475 machine learning. So what happens
NOTE Confidence: 0.9265872
00:07:24.475 --> 00:07:25.914 is if you have a
NOTE Confidence: 0.9265872
00:07:25.914 --> 00:07:27.755 pack of red blood cell
NOTE Confidence: 0.9265872
00:07:27.755 --> 00:07:28.655 as a entity,

NOTE Confidence: 0.95943403
00:07:29.115 --> 00:07:30.715 beginning of entity will label
NOTE Confidence: 0.95943403
00:07:30.715 --> 00:07:33.010 as b, Intermediate token of
NOTE Confidence: 0.95943403
00:07:33.010 --> 00:07:34.450 the entity will label as
NOTE Confidence: 0.95943403
00:07:34.450 --> 00:07:36.130 I. Then all other outside
NOTE Confidence: 0.95943403
00:07:36.130 --> 00:07:37.570 entity will label as o.
NOTE Confidence: 0.95943403
00:07:37.570 --> 00:07:38.850 So then you convert this
NOTE Confidence: 0.95943403
00:07:38.850 --> 00:07:40.790 to a sequence labeling task.
NOTE Confidence: 0.95943403
00:07:41.090 --> 00:07:42.450 You label each word. Is
NOTE Confidence: 0.95943403
00:07:42.450 --> 00:07:43.730 it b, is it I,
NOTE Confidence: 0.95943403
00:07:43.730 --> 00:07:45.525 or o? So become machine
NOTE Confidence: 0.95943403
00:07:45.525 --> 00:07:47.385 learning task to learn. Okay?
NOTE Confidence: 0.95943403
00:07:47.525 --> 00:07:49.045 So that's what we're doing
NOTE Confidence: 0.95943403
00:07:49.045 --> 00:07:50.965 around that time. And then
NOTE Confidence: 0.95943403
00:07:50.965 --> 00:07:51.865 twenty twenty,
NOTE Confidence: 0.96681434
00:07:52.165 --> 00:07:53.225 it's more on,
NOTE Confidence: 0.9543405

00:07:53.765 --> 00:07:55.365 deep learning. I think many
NOTE Confidence: 0.9543405

00:07:55.365 --> 00:07:56.910 of you heard about. At
NOTE Confidence: 0.9543405

00:07:56.910 --> 00:07:58.030 that time, we are looking
NOTE Confidence: 0.9543405

00:07:58.030 --> 00:07:59.870 for those context embeddings like
NOTE Confidence: 0.9543405

00:07:59.870 --> 00:08:01.390 a BER model. So we
NOTE Confidence: 0.9543405

00:08:01.390 --> 00:08:03.390 actually fine tune the BER
NOTE Confidence: 0.9543405

00:08:03.390 --> 00:08:05.070 models from open domain with
NOTE Confidence: 0.9543405

00:08:05.070 --> 00:08:06.670 clinical data and show the
NOTE Confidence: 0.9543405

00:08:06.670 --> 00:08:07.170 performance.
NOTE Confidence: 0.9402269

00:08:08.110 --> 00:08:09.310 So this is just trying
NOTE Confidence: 0.9402269

00:08:09.310 --> 00:08:10.910 to summary. All you want
NOTE Confidence: 0.9402269

00:08:10.910 --> 00:08:12.435 to know is, like, moving
NOTE Confidence: 0.9402269

00:08:12.435 --> 00:08:13.955 from rule based machine learning
NOTE Confidence: 0.9402269

00:08:13.955 --> 00:08:15.414 to deep learning, the performance
NOTE Confidence: 0.9402269

00:08:15.474 --> 00:08:16.615 that you're getting better.
NOTE Confidence: 0.94375366

00:08:17.074 --> 00:08:18.275 And now we move to

NOTE Confidence: 0.94375366
00:08:18.275 --> 00:08:20.194 large language model. How you
NOTE Confidence: 0.94375366
00:08:20.194 --> 00:08:21.634 can use large man lang
NOTE Confidence: 0.94375366
00:08:21.715 --> 00:08:23.235 language model to do this
NOTE Confidence: 0.94375366
00:08:23.235 --> 00:08:23.735 information
NOTE Confidence: 0.9854333
00:08:24.290 --> 00:08:26.130 extraction task. I'll give you
NOTE Confidence: 0.9854333
00:08:26.130 --> 00:08:27.030 three examples,
NOTE Confidence: 0.95121384
00:08:27.490 --> 00:08:29.250 three different approach we have
NOTE Confidence: 0.95121384
00:08:29.250 --> 00:08:31.090 worked on. And, we actually
NOTE Confidence: 0.95121384
00:08:31.090 --> 00:08:32.070 have a a
NOTE Confidence: 0.97962344
00:08:32.450 --> 00:08:34.210 I think, potentially, if you're
NOTE Confidence: 0.97962344
00:08:34.210 --> 00:08:35.170 going to work on your
NOTE Confidence: 0.97962344
00:08:35.170 --> 00:08:36.210 own task, those are the
NOTE Confidence: 0.97962344
00:08:36.210 --> 00:08:37.730 three approach you may can
NOTE Confidence: 0.97962344
00:08:37.730 --> 00:08:38.230 take.
NOTE Confidence: 0.9703888
00:08:38.765 --> 00:08:40.525 The first one, you probably
NOTE Confidence: 0.9703888

00:08:40.525 --> 00:08:42.045 all know. You have GPT
NOTE Confidence: 0.9703888

00:08:42.045 --> 00:08:42.785 over there.
NOTE Confidence: 0.99943227

00:08:43.245 --> 00:08:44.065 All you need
NOTE Confidence: 0.9646077

00:08:44.445 --> 00:08:45.404 to do is write a
NOTE Confidence: 0.9646077

00:08:45.404 --> 00:08:47.404 prompt. Right? Say, give this
NOTE Confidence: 0.9646077

00:08:47.404 --> 00:08:49.404 document. Tell GPT what I
NOTE Confidence: 0.9646077

00:08:49.404 --> 00:08:51.005 want. So that's what we
NOTE Confidence: 0.9646077

00:08:51.005 --> 00:08:52.365 did as a first experiment
NOTE Confidence: 0.9646077

00:08:52.365 --> 00:08:53.505 here. Basically,
NOTE Confidence: 0.8790318

00:08:53.990 --> 00:08:54.649 we give
NOTE Confidence: 0.89976746

00:08:55.110 --> 00:08:56.309 GPT three point five, GPT
NOTE Confidence: 0.89976746

00:08:56.309 --> 00:08:57.829 four at that time. We
NOTE Confidence: 0.89976746

00:08:57.829 --> 00:08:58.870 we say we want to
NOTE Confidence: 0.89976746

00:08:58.870 --> 00:09:00.730 extract a medical problem treatment
NOTE Confidence: 0.89976746

00:09:00.790 --> 00:09:02.649 test out of clinical nodes.
NOTE Confidence: 0.97302645

00:09:03.509 --> 00:09:05.350 So the main exercise here

NOTE Confidence: 0.97302645
00:09:05.350 --> 00:09:07.214 is really about prompt. So
NOTE Confidence: 0.97302645
00:09:07.214 --> 00:09:08.755 we tried actually a different,
NOTE Confidence: 0.9818685
00:09:09.375 --> 00:09:10.975 strategy for prompt. You define
NOTE Confidence: 0.9818685
00:09:10.975 --> 00:09:12.514 the task, define the output.
NOTE Confidence: 0.9818685
00:09:12.735 --> 00:09:13.934 You also need to tell
NOTE Confidence: 0.9818685
00:09:13.934 --> 00:09:15.394 the prompt what's the definition
NOTE Confidence: 0.9818685
00:09:15.615 --> 00:09:16.834 of a medical problem,
NOTE Confidence: 0.9294872
00:09:17.295 --> 00:09:18.574 and then you can also
NOTE Confidence: 0.9294872
00:09:18.574 --> 00:09:20.809 give a guideline. So say
NOTE Confidence: 0.9294872
00:09:20.970 --> 00:09:22.250 because in the previous, I
NOTE Confidence: 0.9294872
00:09:22.250 --> 00:09:23.130 also show you what's a
NOTE Confidence: 0.9294872
00:09:23.130 --> 00:09:24.410 boundary. You may say, here,
NOTE Confidence: 0.9294872
00:09:24.410 --> 00:09:25.050 it has to be a
NOTE Confidence: 0.9294872
00:09:25.050 --> 00:09:26.329 noun phrase for that entity.
NOTE Confidence: 0.9294872
00:09:26.329 --> 00:09:27.290 You give that kind of
NOTE Confidence: 0.9294872

00:09:27.290 --> 00:09:28.809 guideline. Then you can also
NOTE Confidence: 0.9294872

00:09:28.809 --> 00:09:30.350 give additional examples.
NOTE Confidence: 0.8501111

00:09:31.130 --> 00:09:32.490 Here's the sentence. Here's the
NOTE Confidence: 0.8501111

00:09:32.490 --> 00:09:34.029 entity I want to extract.
NOTE Confidence: 0.8501111

00:09:34.125 --> 00:09:35.485 So this kind of call
NOTE Confidence: 0.8501111

00:09:35.565 --> 00:09:37.165 few short learning. It gives
NOTE Confidence: 0.8501111

00:09:37.165 --> 00:09:38.684 a three example, like, a
NOTE Confidence: 0.8501111

00:09:38.684 --> 00:09:40.465 few short learning. Right? And
NOTE Confidence: 0.8853078

00:09:41.725 --> 00:09:43.325 we tested all those. We
NOTE Confidence: 0.8853078

00:09:43.325 --> 00:09:44.525 made a framework for the
NOTE Confidence: 0.8853078

00:09:44.525 --> 00:09:46.205 prompt, and we showed the
NOTE Confidence: 0.8853078

00:09:46.205 --> 00:09:48.360 evaluate on the annotate purpose.
NOTE Confidence: 0.8834308

00:09:48.740 --> 00:09:49.860 And then we we show,
NOTE Confidence: 0.8834308

00:09:49.860 --> 00:09:50.360 actually,
NOTE Confidence: 0.9990163

00:09:50.820 --> 00:09:51.880 if you have a lot
NOTE Confidence: 0.9949787

00:09:52.260 --> 00:09:53.640 of annotated data,

NOTE Confidence: 0.90601486
00:09:54.020 --> 00:09:55.460 then BER model, the previous
NOTE Confidence: 0.90601486
00:09:55.460 --> 00:09:57.000 deep learning based approach,
NOTE Confidence: 0.93101937
00:09:57.380 --> 00:09:59.140 still work better. The zero
NOTE Confidence: 0.93101937
00:09:59.140 --> 00:10:00.944 shot performance on GPT is
NOTE Confidence: 0.93101937
00:10:00.944 --> 00:10:02.225 as not as good as
NOTE Confidence: 0.93101937
00:10:02.225 --> 00:10:03.264 the BER model if you
NOTE Confidence: 0.93101937
00:10:03.264 --> 00:10:04.384 have a lot of annotated
NOTE Confidence: 0.93101937
00:10:04.384 --> 00:10:05.584 data. So that's what we
NOTE Confidence: 0.93101937
00:10:05.584 --> 00:10:06.865 found at that time. But
NOTE Confidence: 0.93101937
00:10:06.865 --> 00:10:08.304 it's actually close because the
NOTE Confidence: 0.93101937
00:10:08.304 --> 00:10:08.804 GPT
NOTE Confidence: 0.96370226
00:10:09.425 --> 00:10:11.425 four model can reach to,
NOTE Confidence: 0.96370226
00:10:11.425 --> 00:10:13.745 like, eighty six versus the
NOTE Confidence: 0.96370226
00:10:13.745 --> 00:10:15.105 BER model trained on hundred
NOTE Confidence: 0.96370226
00:10:15.105 --> 00:10:17.130 samples ninety in the relaxing
NOTE Confidence: 0.96370226

00:10:17.130 --> 00:10:17.630 matching.
NOTE Confidence: 0.99493235

00:10:18.010 --> 00:10:19.769 Relaxing matching means if the
NOTE Confidence: 0.99493235

00:10:19.769 --> 00:10:20.269 entity
NOTE Confidence: 0.8810224

00:10:20.730 --> 00:10:22.410 you predicted versus the entity
NOTE Confidence: 0.8810224

00:10:22.410 --> 00:10:23.929 you annotate is overlap but
NOTE Confidence: 0.8810224

00:10:23.929 --> 00:10:25.630 not exactly same. Okay.
NOTE Confidence: 0.7541335

00:10:26.970 --> 00:10:27.290 And,
NOTE Confidence: 0.98931867

00:10:27.850 --> 00:10:29.230 I'll skip this one.
NOTE Confidence: 0.99464303

00:10:29.865 --> 00:10:31.865 Then I'll skip those two.
NOTE Confidence: 0.99464303

00:10:31.865 --> 00:10:33.465 The second exercise we did
NOTE Confidence: 0.99464303

00:10:33.545 --> 00:10:34.045 okay.
NOTE Confidence: 0.95582557

00:10:34.665 --> 00:10:36.905 Later, llama come out. You
NOTE Confidence: 0.95582557

00:10:36.905 --> 00:10:38.265 have all the weights, like
NOTE Confidence: 0.95582557

00:10:38.265 --> 00:10:39.945 I said. You can actually
NOTE Confidence: 0.95582557

00:10:39.945 --> 00:10:40.445 do,
NOTE Confidence: 0.92377937

00:10:41.145 --> 00:10:42.265 fine tune of those using

NOTE Confidence: 0.92377937
00:10:42.265 --> 00:10:44.020 your additional data. So that's
NOTE Confidence: 0.92377937
00:10:44.020 --> 00:10:45.220 what we did in this.
NOTE Confidence: 0.92377937
00:10:45.220 --> 00:10:46.679 We're working on the same,
NOTE Confidence: 0.90869725
00:10:47.300 --> 00:10:49.220 task, extra medical problem treatment
NOTE Confidence: 0.90869725
00:10:49.220 --> 00:10:50.420 test, but now we have
NOTE Confidence: 0.90869725
00:10:50.420 --> 00:10:52.179 the open source LAMA model.
NOTE Confidence: 0.90869725
00:10:52.179 --> 00:10:53.559 We're actually going to use,
NOTE Confidence: 0.96957475
00:10:54.420 --> 00:10:56.440 annotate data from local corpus
NOTE Confidence: 0.9851837
00:10:56.820 --> 00:10:58.179 to fine tune the LAMA
NOTE Confidence: 0.9851837
00:10:58.179 --> 00:10:59.320 model for this task.
NOTE Confidence: 0.9974308
00:10:59.754 --> 00:11:01.035 And this is what we
NOTE Confidence: 0.9974308
00:11:01.035 --> 00:11:01.535 had
NOTE Confidence: 0.9218615
00:11:02.074 --> 00:11:03.995 through the instruction tuning approach.
NOTE Confidence: 0.9218615
00:11:03.995 --> 00:11:05.194 I think, we're not going
NOTE Confidence: 0.9218615
00:11:05.194 --> 00:11:06.074 to talk a little bit
NOTE Confidence: 0.9218615

00:11:06.074 --> 00:11:07.035 about this, so I'm not
NOTE Confidence: 0.9218615

00:11:07.035 --> 00:11:09.115 going to repeat. Basically, you're
NOTE Confidence: 0.9218615

00:11:09.115 --> 00:11:11.035 converting the annotate data, which
NOTE Confidence: 0.9218615

00:11:11.035 --> 00:11:11.694 I showed,
NOTE Confidence: 0.87941825

00:11:12.020 --> 00:11:13.400 to a instruction dataset.
NOTE Confidence: 0.9558562

00:11:13.780 --> 00:11:14.900 Then you fine tune the
NOTE Confidence: 0.9558562

00:11:14.900 --> 00:11:16.339 LAMA model to to,
NOTE Confidence: 0.9642965

00:11:17.300 --> 00:11:18.280 change the weights,
NOTE Confidence: 0.96370775

00:11:18.820 --> 00:11:20.740 for this specific task. And
NOTE Confidence: 0.96370775

00:11:20.740 --> 00:11:22.020 what we found here is
NOTE Confidence: 0.96370775

00:11:22.020 --> 00:11:22.520 actually
NOTE Confidence: 0.8154574

00:11:23.460 --> 00:11:24.760 when you have a
NOTE Confidence: 0.8111755

00:11:25.615 --> 00:11:26.975 a a lot of annotate
NOTE Confidence: 0.8111755

00:11:26.975 --> 00:11:28.495 data, the per the large
NOTE Confidence: 0.8111755

00:11:28.495 --> 00:11:30.015 entry model, llama three, actually
NOTE Confidence: 0.8111755

00:11:30.015 --> 00:11:30.515 start

NOTE Confidence: 0.90621763
00:11:31.375 --> 00:11:32.575 almost same as a per
NOTE Confidence: 0.90621763
00:11:32.575 --> 00:11:33.615 model if you look at
NOTE Confidence: 0.90621763
00:11:33.615 --> 00:11:35.215 those, but slight better than
NOTE Confidence: 0.90621763
00:11:35.215 --> 00:11:37.054 per model. It's this one
NOTE Confidence: 0.90621763
00:11:37.054 --> 00:11:38.275 is more fair compressive
NOTE Confidence: 0.95954686
00:11:38.600 --> 00:11:40.200 because both model use those
NOTE Confidence: 0.95954686
00:11:40.200 --> 00:11:42.600 hundred annotated sample to train
NOTE Confidence: 0.95954686
00:11:42.600 --> 00:11:44.460 it. Okay? But then
NOTE Confidence: 0.91374004
00:11:45.480 --> 00:11:47.000 the last dataset is the
NOTE Confidence: 0.91374004
00:11:47.000 --> 00:11:49.100 unseen dataset for both model.
NOTE Confidence: 0.91374004
00:11:49.320 --> 00:11:50.280 But then you can see
NOTE Confidence: 0.91374004
00:11:50.280 --> 00:11:51.880 actually the LAMA model have
NOTE Confidence: 0.91374004
00:11:51.880 --> 00:11:53.320 much better performance than the
NOTE Confidence: 0.91374004
00:11:53.320 --> 00:11:54.059 BER model.
NOTE Confidence: 0.8878776
00:11:54.545 --> 00:11:56.065 Indicated the LAMA model is
NOTE Confidence: 0.8878776

00:11:56.065 --> 00:11:57.205 more actually generalizable
NOTE Confidence: 0.89877117

00:11:57.665 --> 00:11:59.345 because it's almost have, like,
NOTE Confidence: 0.89877117

00:11:59.345 --> 00:12:00.645 eight percent improvement
NOTE Confidence: 0.96684116

00:12:01.345 --> 00:12:02.705 versus the BER model is
NOTE Confidence: 0.96684116

00:12:02.705 --> 00:12:03.205 around
NOTE Confidence: 0.84568244

00:12:03.585 --> 00:12:05.445 seventy nine. Here's eighty seven.
NOTE Confidence: 0.96216047

00:12:05.905 --> 00:12:07.440 So now we start to
NOTE Confidence: 0.96216047

00:12:07.440 --> 00:12:08.480 actually build this as a
NOTE Confidence: 0.96216047

00:12:08.480 --> 00:12:10.160 LAMA based kind of information
NOTE Confidence: 0.96216047

00:12:10.160 --> 00:12:12.000 instruction system. But one thing
NOTE Confidence: 0.96216047

00:12:12.000 --> 00:12:13.040 I want to point out,
NOTE Confidence: 0.96216047

00:12:13.040 --> 00:12:14.240 at least when we test
NOTE Confidence: 0.96216047

00:12:14.240 --> 00:12:15.380 on the LAMA three,
NOTE Confidence: 0.85616225

00:12:15.839 --> 00:12:17.040 the speed is the issue,
NOTE Confidence: 0.85616225

00:12:17.040 --> 00:12:18.160 actually. If you look at
NOTE Confidence: 0.85616225

00:12:18.160 --> 00:12:19.519 the BER model, it will

NOTE Confidence: 0.85616225
00:12:19.519 --> 00:12:21.120 take us point two second
NOTE Confidence: 0.85616225
00:12:21.120 --> 00:12:22.865 to do a named entity
NOTE Confidence: 0.85616225
00:12:22.865 --> 00:12:24.245 recognition for one document.
NOTE Confidence: 0.80367243
00:12:25.505 --> 00:12:27.445 Three took thirty nine seconds.
NOTE Confidence: 0.9647269
00:12:27.985 --> 00:12:29.745 So if you're processing millions
NOTE Confidence: 0.9647269
00:12:29.745 --> 00:12:31.365 of notes, that's another concern.
NOTE Confidence: 0.9617658
00:12:31.745 --> 00:12:32.785 So there's a a lot
NOTE Confidence: 0.9617658
00:12:32.785 --> 00:12:34.785 of other issues in addition
NOTE Confidence: 0.9617658
00:12:34.785 --> 00:12:36.225 to the performance you want
NOTE Confidence: 0.9617658
00:12:36.225 --> 00:12:37.510 to consider. That's what I
NOTE Confidence: 0.9617658
00:12:37.510 --> 00:12:39.110 want to bring up. So
NOTE Confidence: 0.9617658
00:12:39.110 --> 00:12:40.170 the third approach,
NOTE Confidence: 0.95940226
00:12:41.190 --> 00:12:42.149 you think about the first
NOTE Confidence: 0.95940226
00:12:42.149 --> 00:12:43.990 approach prompt, you don't really
NOTE Confidence: 0.95940226
00:12:43.990 --> 00:12:45.589 need much GPU. Right? It
NOTE Confidence: 0.95940226

00:12:45.589 --> 00:12:46.950 just costs money to call
NOTE Confidence: 0.95940226
00:12:46.950 --> 00:12:47.450 GPT.
NOTE Confidence: 0.8925754
00:12:47.990 --> 00:12:49.589 Second is fine tune. You
NOTE Confidence: 0.8925754
00:12:49.589 --> 00:12:50.790 do need the GPT to
NOTE Confidence: 0.8925754
00:12:50.790 --> 00:12:51.855 load the mod or you
NOTE Confidence: 0.8925754
00:12:51.855 --> 00:12:52.995 You do need a GPU
NOTE Confidence: 0.8925754
00:12:53.135 --> 00:12:54.815 machine to load the model
NOTE Confidence: 0.8925754
00:12:54.815 --> 00:12:55.615 and fine tune. It may
NOTE Confidence: 0.8925754
00:12:55.615 --> 00:12:57.454 take couple hours to couple
NOTE Confidence: 0.8925754
00:12:57.454 --> 00:12:57.954 days.
NOTE Confidence: 0.8246663
00:12:58.334 --> 00:13:00.014 But this one, we call
NOTE Confidence: 0.8246663
00:13:00.014 --> 00:13:01.454 it continue pre training, the
NOTE Confidence: 0.8246663
00:13:01.454 --> 00:13:02.355 LAMA model.
NOTE Confidence: 0.9610689
00:13:02.735 --> 00:13:03.855 You use a lot of
NOTE Confidence: 0.9610689
00:13:03.855 --> 00:13:05.454 clinical data, like all the
NOTE Confidence: 0.9610689
00:13:05.454 --> 00:13:07.480 notes, all the literature. We

NOTE Confidence: 0.9610689
00:13:07.480 --> 00:13:08.920 combined, like, one hundred twenty
NOTE Confidence: 0.9610689
00:13:08.920 --> 00:13:10.059 nine billions of,
NOTE Confidence: 0.94190633
00:13:10.679 --> 00:13:12.779 tokens to continue pretraining,
NOTE Confidence: 0.9500821
00:13:13.559 --> 00:13:15.480 the the the llama model.
NOTE Confidence: 0.9500821
00:13:15.480 --> 00:13:17.160 It took one hundred fifty
NOTE Confidence: 0.9500821
00:13:17.160 --> 00:13:18.860 GPUs running for a month.
NOTE Confidence: 0.94816655
00:13:19.395 --> 00:13:20.755 That will be a lot
NOTE Confidence: 0.94816655
00:13:20.755 --> 00:13:21.635 of money if you go
NOTE Confidence: 0.94816655
00:13:21.635 --> 00:13:22.375 to Amazon.
NOTE Confidence: 0.9890396
00:13:22.755 --> 00:13:23.715 As you can see, the
NOTE Confidence: 0.9890396
00:13:23.715 --> 00:13:24.375 the computational
NOTE Confidence: 0.9260294
00:13:24.755 --> 00:13:26.695 cost for training this model,
NOTE Confidence: 0.92413664
00:13:27.155 --> 00:13:28.755 much bigger compared to a
NOTE Confidence: 0.92413664
00:13:28.755 --> 00:13:30.775 previous one. But the benefit
NOTE Confidence: 0.92413664
00:13:30.835 --> 00:13:32.195 of this is actually the
NOTE Confidence: 0.92413664

00:13:32.195 --> 00:13:34.079 model has become more generalizable.

NOTE Confidence: 0.9777304

00:13:34.459 --> 00:13:35.839 It can work on multiple

NOTE Confidence: 0.83925885

00:13:36.300 --> 00:13:38.220 top clinical NLP task. So

NOTE Confidence: 0.83925885

00:13:38.220 --> 00:13:39.339 that's what we call this

NOTE Confidence: 0.83925885

00:13:39.339 --> 00:13:40.700 MiLama model. We trained on

NOTE Confidence: 0.83925885

00:13:40.700 --> 00:13:42.000 the LAMA two and,

NOTE Confidence: 0.9034455

00:13:43.100 --> 00:13:44.620 show the better performance on

NOTE Confidence: 0.9034455

00:13:44.620 --> 00:13:46.779 actually multiple task, not just

NOTE Confidence: 0.9034455

00:13:46.779 --> 00:13:48.835 on entity recognition, on question

NOTE Confidence: 0.9034455

00:13:48.835 --> 00:13:49.335 answering,

NOTE Confidence: 0.9795872

00:13:49.795 --> 00:13:51.495 inference, and all other tasks.

NOTE Confidence: 0.8517153

00:13:51.875 --> 00:13:53.155 I'm just going to stop

NOTE Confidence: 0.8517153

00:13:53.155 --> 00:13:53.655 here.

NOTE Confidence: 0.98157936

00:13:54.115 --> 00:13:55.175 So in summary,

NOTE Confidence: 0.99966687

00:13:55.795 --> 00:13:57.175 just to quickly

NOTE Confidence: 0.98135006

00:13:57.475 --> 00:13:59.315 talk about what we learned

NOTE Confidence: 0.98135006
00:13:59.315 --> 00:14:00.375 so far. Basically,
NOTE Confidence: 0.9942674
00:14:01.870 --> 00:14:02.370 when
NOTE Confidence: 0.93553776
00:14:02.910 --> 00:14:04.429 you try to extract information
NOTE Confidence: 0.93553776
00:14:04.429 --> 00:14:05.710 out of notes using large
NOTE Confidence: 0.93553776
00:14:05.710 --> 00:14:07.550 engine model, you still can
NOTE Confidence: 0.93553776
00:14:07.550 --> 00:14:09.150 think about, do you really
NOTE Confidence: 0.93553776
00:14:09.150 --> 00:14:10.350 need a large engine model?
NOTE Confidence: 0.93553776
00:14:10.350 --> 00:14:11.309 If the task is simple,
NOTE Confidence: 0.93553776
00:14:11.309 --> 00:14:12.750 I think even sometimes regular
NOTE Confidence: 0.93553776
00:14:12.750 --> 00:14:14.590 expression, even rule based approach
NOTE Confidence: 0.93553776
00:14:14.590 --> 00:14:16.154 still work. And, also, if
NOTE Confidence: 0.93553776
00:14:16.154 --> 00:14:17.115 you already have a lot
NOTE Confidence: 0.93553776
00:14:17.115 --> 00:14:19.035 of annotate data, then deep
NOTE Confidence: 0.93553776
00:14:19.035 --> 00:14:20.715 learning model like BERT still
NOTE Confidence: 0.93553776
00:14:20.715 --> 00:14:22.555 works well. Okay? And, also,
NOTE Confidence: 0.93553776

00:14:22.555 --> 00:14:23.995 it costs less in terms
NOTE Confidence: 0.93553776

00:14:23.995 --> 00:14:25.375 of computational effort.
NOTE Confidence: 0.9985543

00:14:26.315 --> 00:14:26.815 Then
NOTE Confidence: 0.90716463

00:14:27.195 --> 00:14:28.510 if you think g p
NOTE Confidence: 0.90716463

00:14:28.750 --> 00:14:30.269 large engine model does help
NOTE Confidence: 0.90716463

00:14:30.269 --> 00:14:31.790 for that specific, then we
NOTE Confidence: 0.90716463

00:14:31.790 --> 00:14:33.310 also want to discuss, oh,
NOTE Confidence: 0.90716463

00:14:33.310 --> 00:14:34.910 do I should I train
NOTE Confidence: 0.90716463

00:14:34.910 --> 00:14:36.430 my own large engine model
NOTE Confidence: 0.90716463

00:14:36.430 --> 00:14:37.970 based on open source model,
NOTE Confidence: 0.9902116

00:14:38.430 --> 00:14:40.350 or should I go with
NOTE Confidence: 0.9902116

00:14:40.350 --> 00:14:41.490 GPT? Right?
NOTE Confidence: 0.99778897

00:14:41.845 --> 00:14:42.805 Then there's a lot of
NOTE Confidence: 0.99778897

00:14:42.805 --> 00:14:43.305 concern.
NOTE Confidence: 0.968342

00:14:43.685 --> 00:14:45.365 In addition to performance, you
NOTE Confidence: 0.968342

00:14:45.365 --> 00:14:46.725 also will think about the

NOTE Confidence: 0.968342
00:14:46.725 --> 00:14:47.225 cost.
NOTE Confidence: 0.93421525
00:14:47.605 --> 00:14:49.685 Right? The the GPU requirement.
NOTE Confidence: 0.93421525
00:14:49.685 --> 00:14:51.385 Do you have GPUs locally
NOTE Confidence: 0.93421525
00:14:51.445 --> 00:14:52.745 and all those issues?
NOTE Confidence: 0.9661728
00:14:54.350 --> 00:14:56.050 So in the next three
NOTE Confidence: 0.9661728
00:14:56.350 --> 00:14:58.110 to four presentation, we basically
NOTE Confidence: 0.9661728
00:14:58.110 --> 00:14:59.790 will talk about several things,
NOTE Confidence: 0.9661728
00:14:59.790 --> 00:15:00.290 tools
NOTE Confidence: 0.92801034
00:15:00.750 --> 00:15:02.590 available on the CFG safe
NOTE Confidence: 0.92801034
00:15:02.590 --> 00:15:04.590 environment, which will allow you
NOTE Confidence: 0.92801034
00:15:04.590 --> 00:15:05.950 to do this kind of
NOTE Confidence: 0.92801034
00:15:05.950 --> 00:15:06.450 work.
NOTE Confidence: 0.99080545
00:15:06.785 --> 00:15:07.985 The first tool we will
NOTE Confidence: 0.99080545
00:15:07.985 --> 00:15:09.825 talk about is actually about
NOTE Confidence: 0.99080545
00:15:09.825 --> 00:15:11.425 annotation tool. A lot of
NOTE Confidence: 0.99080545

00:15:11.425 --> 00:15:11.925 people
NOTE Confidence: 0.93463266

00:15:12.465 --> 00:15:13.985 didn't really pay much attention
NOTE Confidence: 0.93463266

00:15:13.985 --> 00:15:15.185 to annotation, but you if
NOTE Confidence: 0.93463266

00:15:15.185 --> 00:15:16.865 you really think about look
NOTE Confidence: 0.93463266

00:15:16.865 --> 00:15:18.305 at all the model training,
NOTE Confidence: 0.93463266

00:15:18.305 --> 00:15:20.060 use even within at the
NOTE Confidence: 0.93463266

00:15:20.060 --> 00:15:21.259 area of a large angle,
NOTE Confidence: 0.93463266

00:15:21.259 --> 00:15:22.220 you still need to do
NOTE Confidence: 0.93463266

00:15:22.220 --> 00:15:23.920 some annotation even just for
NOTE Confidence: 0.93463266

00:15:24.139 --> 00:15:24.639 validation
NOTE Confidence: 0.95147157

00:15:25.100 --> 00:15:26.459 evaluation. Then you need a
NOTE Confidence: 0.95147157

00:15:26.459 --> 00:15:27.519 tool to do that.
NOTE Confidence: 0.9843294

00:15:27.980 --> 00:15:28.940 And we we have a
NOTE Confidence: 0.9843294

00:15:28.940 --> 00:15:30.459 tool installed on the CHP
NOTE Confidence: 0.9843294

00:15:30.459 --> 00:15:32.060 for that purpose. Then second
NOTE Confidence: 0.9843294

00:15:32.060 --> 00:15:33.279 one, we showed you

NOTE Confidence: 0.9261147
00:15:33.615 --> 00:15:35.535 a tool we already fine
NOTE Confidence: 0.9261147
00:15:35.535 --> 00:15:37.295 tuned, and we made it
NOTE Confidence: 0.9261147
00:15:37.295 --> 00:15:39.135 available on the, CHP. You
NOTE Confidence: 0.9261147
00:15:39.135 --> 00:15:40.175 can just call it. I
NOTE Confidence: 0.9261147
00:15:40.175 --> 00:15:41.695 think Nate also talked about
NOTE Confidence: 0.9261147
00:15:41.695 --> 00:15:42.435 the services.
NOTE Confidence: 0.9215255
00:15:43.214 --> 00:15:44.595 Third one is really,
NOTE Confidence: 0.8840262
00:15:45.455 --> 00:15:46.895 go deep. I think Lingfue
NOTE Confidence: 0.8840262
00:15:46.895 --> 00:15:48.175 gonna talk about if you
NOTE Confidence: 0.8840262
00:15:48.175 --> 00:15:49.780 have your own data, start
NOTE Confidence: 0.8840262
00:15:49.780 --> 00:15:51.300 with. How can you fine
NOTE Confidence: 0.8840262
00:15:51.300 --> 00:15:53.060 tune that model with your
NOTE Confidence: 0.8840262
00:15:53.060 --> 00:15:54.520 own data on the CHP,
NOTE Confidence: 0.8840262
00:15:54.580 --> 00:15:56.260 man? So let them just
NOTE Confidence: 0.8840262
00:15:56.260 --> 00:15:56.760 start.
NOTE Confidence: 0.9040973

00:15:58.340 --> 00:15:59.300 Do you wanna go ahead?
NOTE Confidence: 0.9040973

00:15:59.300 --> 00:16:00.440 Start with the annotation.
NOTE Confidence: 0.8264268

00:16:01.460 --> 00:16:02.660 Just watch all the time.
NOTE Confidence: 0.8264268

00:16:02.660 --> 00:16:04.040 Maybe just a little fast.
NOTE Confidence: 0.9883891

00:16:14.935 --> 00:16:15.655 Hi, everyone.
NOTE Confidence: 0.9621317

00:16:15.975 --> 00:16:17.540 Today, I'm just, gonna go
NOTE Confidence: 0.9621317

00:16:17.540 --> 00:16:19.399 through why we need annotation
NOTE Confidence: 0.9621317

00:16:19.540 --> 00:16:21.399 and then, try to introduce
NOTE Confidence: 0.9621317

00:16:21.459 --> 00:16:23.320 our annotation tool, Blue.
NOTE Confidence: 0.99413157

00:16:24.740 --> 00:16:25.240 So,
NOTE Confidence: 0.9859944

00:16:25.700 --> 00:16:27.860 annotation is the process of
NOTE Confidence: 0.9859944

00:16:27.860 --> 00:16:28.839 labeling data,
NOTE Confidence: 0.92907965

00:16:29.585 --> 00:16:31.685 marking span of text images
NOTE Confidence: 0.92907965

00:16:31.905 --> 00:16:33.745 or other content with additional
NOTE Confidence: 0.92907965

00:16:33.745 --> 00:16:35.985 information such as, entity types,
NOTE Confidence: 0.92907965

00:16:35.985 --> 00:16:36.485 categories,

NOTE Confidence: 0.9222235
00:16:36.865 --> 00:16:37.685 or relationships.
NOTE Confidence: 0.9549063
00:16:38.225 --> 00:16:39.905 Just as doctor Xu mentioned
NOTE Confidence: 0.9549063
00:16:39.905 --> 00:16:42.240 before showing the graph that
NOTE Confidence: 0.9549063
00:16:42.240 --> 00:16:44.480 there is entities annotated and
NOTE Confidence: 0.9549063
00:16:44.480 --> 00:16:46.500 also the relationship, annotate
NOTE Confidence: 0.85652757
00:16:48.000 --> 00:16:48.740 between those
NOTE Confidence: 0.96907246
00:16:50.320 --> 00:16:53.140 entities. Annotation is, critically important
NOTE Confidence: 0.96907246
00:16:53.360 --> 00:16:54.880 because it serves as the
NOTE Confidence: 0.96907246
00:16:54.880 --> 00:16:57.154 foundation for the machine learning
NOTE Confidence: 0.96907246
00:16:57.154 --> 00:16:58.535 and deep learning models.
NOTE Confidence: 0.9061434
00:16:59.235 --> 00:16:59.975 These models,
NOTE Confidence: 0.9071455
00:17:00.514 --> 00:17:02.855 heavily rely on annotated dataset
NOTE Confidence: 0.9748663
00:17:03.315 --> 00:17:05.635 to learn meaningful patterns and
NOTE Confidence: 0.9748663
00:17:05.635 --> 00:17:06.595 then to make,
NOTE Confidence: 0.9968814
00:17:07.155 --> 00:17:08.135 accurate predictions.
NOTE Confidence: 0.852738

00:17:09.155 --> 00:17:10.294 Annotation remain
NOTE Confidence: 0.54891384

00:17:11.149 --> 00:17:12.609 the large length models.
NOTE Confidence: 0.9213072

00:17:13.149 --> 00:17:15.730 Although LMM is highly capable,
NOTE Confidence: 0.9746174

00:17:16.590 --> 00:17:17.869 it they still depends on
NOTE Confidence: 0.9746174

00:17:17.869 --> 00:17:19.730 the annotation data for
NOTE Confidence: 0.8921592

00:17:20.269 --> 00:17:21.710 fine tuning and the specific
NOTE Confidence: 0.8921592

00:17:21.710 --> 00:17:23.710 to for specific task and
NOTE Confidence: 0.8921592

00:17:23.710 --> 00:17:25.710 for evaluation against the ground
NOTE Confidence: 0.8921592

00:17:25.710 --> 00:17:26.210 truth.
NOTE Confidence: 0.90066063

00:17:27.815 --> 00:17:28.315 Here
NOTE Confidence: 0.8587521

00:17:33.255 --> 00:17:34.695 here you can see a
NOTE Confidence: 0.8587521

00:17:34.695 --> 00:17:35.675 table that,
NOTE Confidence: 0.9907583

00:17:36.295 --> 00:17:37.195 compare performance
NOTE Confidence: 0.92630583

00:17:37.655 --> 00:17:38.155 between,
NOTE Confidence: 0.9922838

00:17:38.615 --> 00:17:39.595 multiple models
NOTE Confidence: 0.8878392

00:17:41.160 --> 00:17:41.660 for,

NOTE Confidence: 0.86685234
00:17:42.520 --> 00:17:44.280 including the lama three variants
NOTE Confidence: 0.86685234
00:17:44.280 --> 00:17:45.560 and also a fine tuned
NOTE Confidence: 0.86685234
00:17:45.560 --> 00:17:47.480 lama model on the language
NOTE Confidence: 0.86685234
00:17:47.480 --> 00:17:48.540 annotation task.
NOTE Confidence: 0.8769667
00:17:50.200 --> 00:17:51.420 As seen in the result,
NOTE Confidence: 0.8769667
00:17:51.480 --> 00:17:52.460 fine tuned model,
NOTE Confidence: 0.79888856
00:17:52.840 --> 00:17:54.619 which trained our well annotated
NOTE Confidence: 0.7536975
00:17:55.000 --> 00:17:55.500 dataset
NOTE Confidence: 0.8551343
00:17:55.924 --> 00:17:57.525 tend to perform better on
NOTE Confidence: 0.8551343
00:17:57.525 --> 00:17:58.025 specific
NOTE Confidence: 0.82373977
00:17:58.804 --> 00:17:59.304 targeted,
NOTE Confidence: 0.9157275
00:18:00.005 --> 00:18:00.505 project.
NOTE Confidence: 0.99655753
00:18:04.405 --> 00:18:06.345 There are several key topics
NOTE Confidence: 0.99655753
00:18:06.565 --> 00:18:08.184 related to the annotation.
NOTE Confidence: 0.96674556
00:18:08.779 --> 00:18:10.700 The process always begin with,
NOTE Confidence: 0.96674556

00:18:11.100 --> 00:18:13.359 developing a clear and detailed
NOTE Confidence: 0.96674556

00:18:13.500 --> 00:18:14.399 annotation guideline.
NOTE Confidence: 0.9453406

00:18:14.940 --> 00:18:17.279 A well developed guideline improves
NOTE Confidence: 0.9453406

00:18:17.340 --> 00:18:17.840 consistency
NOTE Confidence: 0.9710151

00:18:18.220 --> 00:18:18.700 among,
NOTE Confidence: 0.9992384

00:18:19.020 --> 00:18:19.520 annotators
NOTE Confidence: 0.95170295

00:18:20.299 --> 00:18:22.080 and lead to higher annotation
NOTE Confidence: 0.95170295

00:18:22.220 --> 00:18:24.435 quality as the speed up
NOTE Confidence: 0.95170295

00:18:24.595 --> 00:18:26.375 the onboarding of new annotators
NOTE Confidence: 0.9033354

00:18:26.915 --> 00:18:29.155 and also make conflict resolution
NOTE Confidence: 0.9033354

00:18:29.155 --> 00:18:29.655 easier.
NOTE Confidence: 0.95977855

00:18:30.355 --> 00:18:32.135 Once the guideline is,
NOTE Confidence: 0.94053966

00:18:33.155 --> 00:18:34.835 developed, the next step will
NOTE Confidence: 0.94053966

00:18:34.835 --> 00:18:36.535 be select the annotation
NOTE Confidence: 0.9732097

00:18:36.835 --> 00:18:38.615 with appropriate domain knowledge
NOTE Confidence: 0.99653435

00:18:38.940 --> 00:18:40.539 and then train them thoroughly

NOTE Confidence: 0.99653435
00:18:40.539 --> 00:18:41.760 based on the guideline.
NOTE Confidence: 0.99665916
00:18:43.740 --> 00:18:45.100 After the training, it is
NOTE Confidence: 0.99665916
00:18:45.100 --> 00:18:46.399 important to continuously
NOTE Confidence: 0.94221324
00:18:46.940 --> 00:18:48.859 checking the and monitor the
NOTE Confidence: 0.94221324
00:18:48.859 --> 00:18:49.359 annotation
NOTE Confidence: 0.9939876
00:18:49.899 --> 00:18:50.399 quality.
NOTE Confidence: 0.89248914
00:18:51.260 --> 00:18:52.465 This including,
NOTE Confidence: 0.95826244
00:18:53.405 --> 00:18:55.265 checking agreement among annotators,
NOTE Confidence: 0.9967836
00:18:55.965 --> 00:18:56.865 holding discussion
NOTE Confidence: 0.8925627
00:18:57.165 --> 00:18:57.984 to resolve,
NOTE Confidence: 0.92002845
00:18:58.525 --> 00:18:59.025 disagreement
NOTE Confidence: 0.9538897
00:18:59.725 --> 00:19:01.725 and refine the guideline based
NOTE Confidence: 0.9538897
00:19:01.725 --> 00:19:03.744 on common errors or ambiguities
NOTE Confidence: 0.9865798
00:19:04.045 --> 00:19:05.585 identified during the process.
NOTE Confidence: 0.9805972
00:19:06.859 --> 00:19:08.300 And later, I will also
NOTE Confidence: 0.9805972

00:19:08.300 --> 00:19:10.380 introduce the annotation tool that
NOTE Confidence: 0.9805972

00:19:10.380 --> 00:19:12.220 can support and streamline the
NOTE Confidence: 0.9805972

00:19:12.220 --> 00:19:13.200 annotation workflow.
NOTE Confidence: 0.8471476

00:19:16.780 --> 00:19:18.480 For annotation guideline development,
NOTE Confidence: 0.9890651

00:19:19.035 --> 00:19:20.715 the first step is always
NOTE Confidence: 0.9890651

00:19:20.715 --> 00:19:22.315 define the goal of your
NOTE Confidence: 0.9890651

00:19:22.315 --> 00:19:22.815 project.
NOTE Confidence: 0.99320775

00:19:24.395 --> 00:19:25.755 Clearly state what you are
NOTE Confidence: 0.99320775

00:19:25.755 --> 00:19:26.815 trying to achieve,
NOTE Confidence: 0.9984066

00:19:27.195 --> 00:19:28.175 with the project.
NOTE Confidence: 0.96013814

00:19:28.715 --> 00:19:30.975 Next, provide a clear definition
NOTE Confidence: 0.96013814

00:19:31.115 --> 00:19:33.919 for all the concept like
NOTE Confidence: 0.96013814

00:19:33.980 --> 00:19:36.380 entities, relations, or some special
NOTE Confidence: 0.96013814

00:19:36.380 --> 00:19:36.880 terms.
NOTE Confidence: 0.85769576

00:19:38.539 --> 00:19:39.679 After that, develop
NOTE Confidence: 0.84394073

00:19:40.059 --> 00:19:42.320 detailed annotation rule that covers

NOTE Confidence: 0.84394073
00:19:42.539 --> 00:19:43.039 morality
NOTE Confidence: 0.88394725
00:19:43.340 --> 00:19:43.840 scenarios
NOTE Confidence: 0.9898371
00:19:44.140 --> 00:19:46.080 and edge cases to minimize
NOTE Confidence: 0.9898371
00:19:46.299 --> 00:19:47.039 the ambiguity.
NOTE Confidence: 0.9626433
00:19:47.765 --> 00:19:49.845 It is also essential to
NOTE Confidence: 0.9626433
00:19:49.845 --> 00:19:51.845 include many real world examples
NOTE Confidence: 0.9626433
00:19:51.845 --> 00:19:52.665 in the guidelines.
NOTE Confidence: 0.9093584
00:19:53.845 --> 00:19:56.565 Illustrate both, correct annotation and
NOTE Confidence: 0.9093584
00:19:56.565 --> 00:19:57.465 common errors.
NOTE Confidence: 0.8916718
00:19:58.645 --> 00:20:00.325 Guideline development is not a
NOTE Confidence: 0.8916718
00:20:00.325 --> 00:20:02.345 one time effort. It's always
NOTE Confidence: 0.8916718
00:20:02.485 --> 00:20:02.985 need,
NOTE Confidence: 0.98565775
00:20:03.520 --> 00:20:04.500 iterative process.
NOTE Confidence: 0.9832192
00:20:05.359 --> 00:20:07.440 It is important to involve
NOTE Confidence: 0.9832192
00:20:07.440 --> 00:20:07.920 both,
NOTE Confidence: 0.8944893

00:20:08.240 --> 00:20:10.420 domain experts and the linguistics
NOTE Confidence: 0.8944893

00:20:10.640 --> 00:20:11.940 or info implementations
NOTE Confidence: 0.96231383

00:20:12.480 --> 00:20:13.700 to ensure both,
NOTE Confidence: 0.9655188

00:20:14.240 --> 00:20:17.140 technical accuracy and practical usability.
NOTE Confidence: 0.90759677

00:20:18.274 --> 00:20:20.215 After the initial guideline is
NOTE Confidence: 0.7848914

00:20:20.515 --> 00:20:21.575 create created,
NOTE Confidence: 0.9255918

00:20:22.115 --> 00:20:23.554 it should be refined during
NOTE Confidence: 0.9255918

00:20:23.554 --> 00:20:26.135 the annotation training and tested
NOTE Confidence: 0.9255918

00:20:26.274 --> 00:20:27.414 on real world data.
NOTE Confidence: 0.75119877

00:20:28.195 --> 00:20:29.174 Given the variability
NOTE Confidence: 0.96419835

00:20:29.475 --> 00:20:30.294 and complexity
NOTE Confidence: 0.9568897

00:20:30.674 --> 00:20:32.419 of the real world data,
NOTE Confidence: 0.9568897

00:20:32.639 --> 00:20:34.500 new scenarios will always,
NOTE Confidence: 0.6647098

00:20:36.000 --> 00:20:36.500 inevitably
NOTE Confidence: 0.7834834

00:20:37.119 --> 00:20:37.619 raised
NOTE Confidence: 0.8991647

00:20:38.000 --> 00:20:39.859 and may occur further guideline

NOTE Confidence: 0.8991647
00:20:39.919 --> 00:20:40.419 updates.
NOTE Confidence: 0.913407
00:20:41.359 --> 00:20:42.799 Once the guideline is stable
NOTE Confidence: 0.913407
00:20:42.799 --> 00:20:44.720 and robust, the process can
NOTE Confidence: 0.913407
00:20:44.720 --> 00:20:46.544 move to the corpus final
NOTE Confidence: 0.913407
00:20:46.544 --> 00:20:47.684 day final finalization.
NOTE Confidence: 0.8648101
00:20:51.105 --> 00:20:52.784 Here, you can see, example
NOTE Confidence: 0.8648101
00:20:52.784 --> 00:20:54.164 of a annotation guideline.
NOTE Confidence: 0.9967257
00:20:54.625 --> 00:20:56.085 The goal of this guideline
NOTE Confidence: 0.9967257
00:20:56.225 --> 00:20:57.125 is to,
NOTE Confidence: 0.9898429
00:20:57.664 --> 00:20:58.804 identify meaningful,
NOTE Confidence: 0.91760415
00:20:59.424 --> 00:21:01.109 clinical concept from,
NOTE Confidence: 0.97520125
00:21:01.510 --> 00:21:02.570 important patient,
NOTE Confidence: 0.9993994
00:21:03.190 --> 00:21:04.090 medical records
NOTE Confidence: 0.9066288
00:21:04.390 --> 00:21:06.410 and to help extract information
NOTE Confidence: 0.9066288
00:21:06.470 --> 00:21:06.970 like
NOTE Confidence: 0.85468864

00:21:07.350 --> 00:21:09.429 the test problem, drug, and
NOTE Confidence: 0.85468864
00:21:09.429 --> 00:21:09.929 treatment.
NOTE Confidence: 0.98490024
00:21:10.869 --> 00:21:12.309 As shown on the left
NOTE Confidence: 0.98490024
00:21:12.309 --> 00:21:12.809 side,
NOTE Confidence: 0.95480883
00:21:13.270 --> 00:21:15.130 we provide a detailed definition
NOTE Confidence: 0.94533736
00:21:15.510 --> 00:21:16.010 to
NOTE Confidence: 0.9817573
00:21:16.855 --> 00:21:18.934 to ensure the annotator understand
NOTE Confidence: 0.9817573
00:21:18.934 --> 00:21:20.075 what should be labeled.
NOTE Confidence: 0.90039974
00:21:21.014 --> 00:21:22.455 In this guideline, we also
NOTE Confidence: 0.90039974
00:21:22.455 --> 00:21:23.835 introduced the modifiers,
NOTE Confidence: 0.84820986
00:21:24.455 --> 00:21:25.514 a concept that,
NOTE Confidence: 0.6417678
00:21:26.534 --> 00:21:27.034 complement,
NOTE Confidence: 0.91986996
00:21:27.494 --> 00:21:29.494 entity and also extend its
NOTE Confidence: 0.91986996
00:21:29.494 --> 00:21:29.994 mailing.
NOTE Confidence: 0.93745583
00:21:30.520 --> 00:21:31.500 For each modifiers
NOTE Confidence: 0.9115747
00:21:32.200 --> 00:21:34.040 such as the severity and

NOTE Confidence: 0.9115747
00:21:34.040 --> 00:21:34.859 body location,
NOTE Confidence: 0.9525939
00:21:35.880 --> 00:21:37.160 it will also need to
NOTE Confidence: 0.9525939
00:21:37.160 --> 00:21:39.320 be specific on how it
NOTE Confidence: 0.9525939
00:21:39.320 --> 00:21:40.679 need to be annotated and
NOTE Confidence: 0.9525939
00:21:40.679 --> 00:21:42.840 what's the relationship with the
NOTE Confidence: 0.9525939
00:21:42.840 --> 00:21:43.340 entities.
NOTE Confidence: 0.95569813
00:21:44.294 --> 00:21:46.375 Additionally, the guideline need to
NOTE Confidence: 0.95569813
00:21:46.375 --> 00:21:48.875 include, many real world examples,
NOTE Confidence: 0.92251927
00:21:49.734 --> 00:21:51.254 like the diagram show on
NOTE Confidence: 0.92251927
00:21:51.254 --> 00:21:51.835 the bottom,
NOTE Confidence: 0.99755585
00:21:52.294 --> 00:21:53.274 to illustrate
NOTE Confidence: 0.99846554
00:21:53.815 --> 00:21:55.595 the correct annotation practice
NOTE Confidence: 0.9772594
00:21:56.100 --> 00:21:57.400 and for ambiguous
NOTE Confidence: 0.9515268
00:21:57.780 --> 00:21:59.480 phrase or tricky scenarios.
NOTE Confidence: 0.9686359
00:21:59.940 --> 00:22:01.619 Examples will also need to
NOTE Confidence: 0.9686359

00:22:01.619 --> 00:22:03.859 be provided to establish a
NOTE Confidence: 0.9686359

00:22:03.859 --> 00:22:06.359 clear consistent rules for annotator
NOTE Confidence: 0.9686359

00:22:06.500 --> 00:22:07.160 to follow.
NOTE Confidence: 0.93705726

00:22:11.635 --> 00:22:13.255 It is it is important
NOTE Confidence: 0.93705726

00:22:13.315 --> 00:22:15.475 to choose the annotators with
NOTE Confidence: 0.93705726

00:22:15.475 --> 00:22:16.615 a proper it's,
NOTE Confidence: 0.9674798

00:22:17.395 --> 00:22:18.855 background for your task.
NOTE Confidence: 0.98970383

00:22:19.395 --> 00:22:21.175 Depending on the complexity,
NOTE Confidence: 0.9127062

00:22:21.830 --> 00:22:23.450 you might need to choose
NOTE Confidence: 0.9127062

00:22:23.510 --> 00:22:25.369 a domain expert like physicians,
NOTE Confidence: 0.944911

00:22:25.830 --> 00:22:26.330 nurses,
NOTE Confidence: 0.84638005

00:22:27.030 --> 00:22:28.730 or, medical students,
NOTE Confidence: 0.96491796

00:22:29.030 --> 00:22:30.490 or just or some layperson
NOTE Confidence: 0.96491796

00:22:30.630 --> 00:22:32.630 for more general and broad
NOTE Confidence: 0.96491796

00:22:32.630 --> 00:22:33.130 annotation.
NOTE Confidence: 0.9794002

00:22:34.070 --> 00:22:36.150 Training for annotator is a

NOTE Confidence: 0.9794002
00:22:36.150 --> 00:22:37.290 iterative process.
NOTE Confidence: 0.940899
00:22:38.125 --> 00:22:39.885 Annotator should be trained and
NOTE Confidence: 0.940899
00:22:39.885 --> 00:22:41.424 evaluate multiple times,
NOTE Confidence: 0.8823554
00:22:41.965 --> 00:22:43.105 until they achieve
NOTE Confidence: 0.95309144
00:22:43.484 --> 00:22:45.585 the expected level of performance.
NOTE Confidence: 0.9492799
00:22:46.684 --> 00:22:47.664 Quality checking,
NOTE Confidence: 0.988212
00:22:48.205 --> 00:22:50.065 need to be ongoing during
NOTE Confidence: 0.988212
00:22:50.125 --> 00:22:51.505 the annotation progress.
NOTE Confidence: 0.96396536
00:22:52.410 --> 00:22:54.730 Regularly review is always needed
NOTE Confidence: 0.96396536
00:22:54.730 --> 00:22:55.790 during their work.
NOTE Confidence: 0.965456
00:22:56.730 --> 00:22:58.170 And you also need to
NOTE Confidence: 0.965456
00:22:58.170 --> 00:22:59.310 provide the feedback,
NOTE Confidence: 0.9594874
00:23:00.090 --> 00:23:02.110 on time and sometimes additional
NOTE Confidence: 0.9594874
00:23:02.170 --> 00:23:03.630 retraining for the annotators.
NOTE Confidence: 0.9797674
00:23:07.075 --> 00:23:08.455 When managing a project,
NOTE Confidence: 0.9195564

00:23:08.994 --> 00:23:09.734 which contains,
NOTE Confidence: 0.99359465

00:23:10.115 --> 00:23:11.255 multiple annotators,
NOTE Confidence: 0.91868216

00:23:11.635 --> 00:23:12.355 there are several,
NOTE Confidence: 0.9972546

00:23:13.715 --> 00:23:15.475 important steps must be taken
NOTE Confidence: 0.9972546

00:23:15.475 --> 00:23:16.855 to ensure the quality.
NOTE Confidence: 0.9866996

00:23:17.740 --> 00:23:20.080 Before starting the actual annotation,
NOTE Confidence: 0.927501

00:23:21.100 --> 00:23:22.240 train each annotator
NOTE Confidence: 0.9844237

00:23:23.340 --> 00:23:25.340 thoroughly to ensure they can
NOTE Confidence: 0.9844237

00:23:25.340 --> 00:23:28.240 produce consistent and reliable annotation
NOTE Confidence: 0.9844237

00:23:28.299 --> 00:23:28.799 result
NOTE Confidence: 0.9412294

00:23:29.260 --> 00:23:30.720 that align with the guideline
NOTE Confidence: 0.9412294

00:23:30.779 --> 00:23:31.600 you developed.
NOTE Confidence: 0.9546554

00:23:32.505 --> 00:23:33.545 If resource allow,
NOTE Confidence: 0.98452073

00:23:34.505 --> 00:23:35.005 implement
NOTE Confidence: 0.99367166

00:23:35.305 --> 00:23:36.205 double annotation
NOTE Confidence: 0.9252542

00:23:36.665 --> 00:23:37.165 strategy.

NOTE Confidence: 0.9812212
00:23:37.865 --> 00:23:39.465 Ideally, each sample need to
NOTE Confidence: 0.9812212
00:23:39.465 --> 00:23:40.205 be annotated
NOTE Confidence: 0.9545889
00:23:40.585 --> 00:23:42.925 by two annotators independently.
NOTE Confidence: 0.93457323
00:23:43.785 --> 00:23:45.565 Then a third more experienced
NOTE Confidence: 0.93457323
00:23:45.625 --> 00:23:46.125 annotator
NOTE Confidence: 0.9986475
00:23:46.585 --> 00:23:47.244 can review
NOTE Confidence: 0.943538
00:23:47.680 --> 00:23:49.520 any discrepancies and make the
NOTE Confidence: 0.943538
00:23:49.520 --> 00:23:50.420 final decision.
NOTE Confidence: 0.98549706
00:23:50.960 --> 00:23:52.320 This process will have to
NOTE Confidence: 0.98549706
00:23:52.320 --> 00:23:53.680 maintain a high quality of
NOTE Confidence: 0.98549706
00:23:53.680 --> 00:23:54.180 annotation.
NOTE Confidence: 0.9042241
00:23:55.200 --> 00:23:57.200 If double annotation for the
NOTE Confidence: 0.9042241
00:23:57.200 --> 00:23:58.980 intel dataset is not feasible,
NOTE Confidence: 0.9206318
00:23:59.920 --> 00:24:01.300 assign small overlapping
NOTE Confidence: 0.99473923
00:24:01.760 --> 00:24:03.845 subset of data to multiple
NOTE Confidence: 0.99473923

00:24:03.845 --> 00:24:04.345 annotators.
NOTE Confidence: 0.9803714

00:24:05.445 --> 00:24:07.125 This overlap allow you to
NOTE Confidence: 0.9803714

00:24:07.125 --> 00:24:08.345 calculate inter,
NOTE Confidence: 0.7265367

00:24:08.805 --> 00:24:09.305 interagreement
NOTE Confidence: 0.9979175

00:24:10.885 --> 00:24:11.785 of the annotator
NOTE Confidence: 0.97570616

00:24:12.085 --> 00:24:13.525 and then provide a way
NOTE Confidence: 0.97570616

00:24:13.525 --> 00:24:15.385 to monitor and maintain annotation
NOTE Confidence: 0.97570616

00:24:15.445 --> 00:24:15.945 quality.
NOTE Confidence: 0.99526465

00:24:19.800 --> 00:24:21.660 When checking the annotation quality
NOTE Confidence: 0.8735504

00:24:22.119 --> 00:24:23.800 for the NER task, we
NOTE Confidence: 0.8735504

00:24:23.800 --> 00:24:25.320 will focus on two main
NOTE Confidence: 0.8735504

00:24:25.320 --> 00:24:25.820 areas.
NOTE Confidence: 0.91402775

00:24:26.359 --> 00:24:27.880 The first one is entity
NOTE Confidence: 0.91402775

00:24:27.880 --> 00:24:29.800 type agreement and then empty
NOTE Confidence: 0.91402775

00:24:29.800 --> 00:24:30.700 span agreement.
NOTE Confidence: 0.9475457

00:24:31.585 --> 00:24:33.345 For anti type agreement, we

NOTE Confidence: 0.9475457
00:24:33.345 --> 00:24:35.205 verify whether annotators
NOTE Confidence: 0.93709534
00:24:35.585 --> 00:24:37.105 assign the same type for
NOTE Confidence: 0.93709534
00:24:37.105 --> 00:24:37.685 the entity.
NOTE Confidence: 0.95774543
00:24:38.145 --> 00:24:39.505 You can see in this,
NOTE Confidence: 0.95774543
00:24:39.825 --> 00:24:41.685 graph, one of them annotate
NOTE Confidence: 0.8704652
00:24:42.225 --> 00:24:43.985 the Vancom missing HCL as
NOTE Confidence: 0.8704652
00:24:43.985 --> 00:24:45.585 the drug and another one
NOTE Confidence: 0.8704652
00:24:45.585 --> 00:24:46.725 annotate as treatment.
NOTE Confidence: 0.9510615
00:24:47.320 --> 00:24:48.140 Then this,
NOTE Confidence: 0.99818313
00:24:49.160 --> 00:24:50.220 mismatch will
NOTE Confidence: 0.9631529
00:24:50.680 --> 00:24:52.540 need to be discussed when
NOTE Confidence: 0.96569836
00:24:53.080 --> 00:24:55.000 during during the annotation and
NOTE Confidence: 0.96569836
00:24:55.000 --> 00:24:55.820 then correct,
NOTE Confidence: 0.9999011
00:24:56.840 --> 00:24:57.660 for the final
NOTE Confidence: 0.9877126
00:24:58.119 --> 00:24:58.619 step.
NOTE Confidence: 0.88563275

00:24:58.920 --> 00:25:00.315 And also for the
NOTE Confidence: 0.86336863

00:25:00.715 --> 00:25:02.554 anti span agreement, we check
NOTE Confidence: 0.86336863

00:25:02.554 --> 00:25:03.934 whether both annotator
NOTE Confidence: 0.93165106

00:25:04.315 --> 00:25:05.674 select the same portion of
NOTE Confidence: 0.93165106

00:25:05.674 --> 00:25:07.855 text. For the same example,
NOTE Confidence: 0.93165106

00:25:07.994 --> 00:25:08.494 one
NOTE Confidence: 0.91529036

00:25:09.754 --> 00:25:11.375 labeled a lot of emotional
NOTE Confidence: 0.91529036

00:25:11.514 --> 00:25:13.034 stress as a problem and
NOTE Confidence: 0.91529036

00:25:13.034 --> 00:25:14.794 another one annotate just the
NOTE Confidence: 0.91529036

00:25:14.794 --> 00:25:15.855 emotional stress.
NOTE Confidence: 0.9060362

00:25:17.159 --> 00:25:18.700 When such mismatch occurs,
NOTE Confidence: 0.99607

00:25:19.399 --> 00:25:21.080 it is important to refer
NOTE Confidence: 0.99607

00:25:21.080 --> 00:25:22.139 back to the guideline
NOTE Confidence: 0.9172219

00:25:22.440 --> 00:25:24.139 and to and to determine
NOTE Confidence: 0.9172219

00:25:24.200 --> 00:25:26.200 what is the current correct
NOTE Confidence: 0.9172219

00:25:26.200 --> 00:25:27.419 one to move forward.

NOTE Confidence: 0.9012047
00:25:40.085 --> 00:25:42.645 When checking the annotation quality
NOTE Confidence: 0.9012047
00:25:42.645 --> 00:25:43.305 in relation
NOTE Confidence: 0.7833279
00:25:43.619 --> 00:25:44.600 to extraction,
NOTE Confidence: 0.9923493
00:25:45.059 --> 00:25:46.760 there are three main aspects
NOTE Confidence: 0.9923493
00:25:46.900 --> 00:25:47.960 we need to evaluate.
NOTE Confidence: 0.97593874
00:25:48.420 --> 00:25:49.720 The first one is,
NOTE Confidence: 0.9929292
00:25:50.420 --> 00:25:51.720 relation type agreement.
NOTE Confidence: 0.99434996
00:25:52.180 --> 00:25:53.960 We check whether both annotators
NOTE Confidence: 0.9393045
00:25:54.340 --> 00:25:55.859 assign the same type of
NOTE Confidence: 0.9393045
00:25:55.859 --> 00:25:57.240 relation between entities,
NOTE Confidence: 0.992069
00:25:57.835 --> 00:25:59.775 and then we evaluate,
NOTE Confidence: 0.9740135
00:26:00.475 --> 00:26:02.955 entity pair. We verify if
NOTE Confidence: 0.9740135
00:26:02.955 --> 00:26:04.575 the same entity are being
NOTE Confidence: 0.9740135
00:26:04.635 --> 00:26:05.855 linked by the relation.
NOTE Confidence: 0.97218424
00:26:06.235 --> 00:26:07.595 And finally, we need to
NOTE Confidence: 0.97218424

00:26:07.595 --> 00:26:08.895 check the direction
NOTE Confidence: 0.954981

00:26:09.435 --> 00:26:09.935 directionality,
NOTE Confidence: 0.9438983

00:26:10.789 --> 00:26:12.630 which is important for some
NOTE Confidence: 0.9438983

00:26:12.630 --> 00:26:14.789 tasks because the direction may
NOTE Confidence: 0.9438983

00:26:14.789 --> 00:26:15.770 change the meanings.
NOTE Confidence: 0.98370934

00:26:19.109 --> 00:26:19.609 To
NOTE Confidence: 0.9932904

00:26:20.710 --> 00:26:22.470 evaluate, there are several metrics
NOTE Confidence: 0.9932904

00:26:22.470 --> 00:26:23.350 we can use.
NOTE Confidence: 0.93146634

00:26:23.830 --> 00:26:26.015 The common one is precision
NOTE Confidence: 0.93146634

00:26:26.075 --> 00:26:27.615 recall and f one measure,
NOTE Confidence: 0.93146634

00:26:27.755 --> 00:26:30.015 which help quantify how consistently
NOTE Confidence: 0.8766902

00:26:30.395 --> 00:26:32.335 annotate and identify and classify
NOTE Confidence: 0.8766902

00:26:32.475 --> 00:26:32.975 entities.
NOTE Confidence: 0.88953125

00:26:33.595 --> 00:26:35.755 Additionally, we can also use
NOTE Confidence: 0.88953125

00:26:35.755 --> 00:26:38.234 some statistical measures such as
NOTE Confidence: 0.88953125

00:26:38.234 --> 00:26:39.775 Cohen's kappa or.

NOTE Confidence: 0.9993607
00:26:41.190 --> 00:26:42.010 Another important
NOTE Confidence: 0.9664983
00:26:42.549 --> 00:26:44.230 matter is, self train and
NOTE Confidence: 0.9664983
00:26:44.230 --> 00:26:46.309 self test. By training the
NOTE Confidence: 0.9664983
00:26:46.309 --> 00:26:48.409 model on the annotated dataset
NOTE Confidence: 0.9664983
00:26:48.470 --> 00:26:49.990 and then testing on the
NOTE Confidence: 0.9664983
00:26:49.990 --> 00:26:50.730 same dataset,
NOTE Confidence: 0.99972117
00:26:51.750 --> 00:26:52.490 we can
NOTE Confidence: 0.9271745
00:26:53.525 --> 00:26:55.044 check if the model achieve
NOTE Confidence: 0.9271745
00:26:55.044 --> 00:26:55.784 high performance.
NOTE Confidence: 0.9990876
00:26:56.325 --> 00:26:57.225 If the performance
NOTE Confidence: 0.9699376
00:26:57.525 --> 00:26:59.145 is low, it may indicate
NOTE Confidence: 0.9699376
00:26:59.205 --> 00:27:02.025 underlying issue with, annotation inconsistency
NOTE Confidence: 0.8670124
00:27:02.885 --> 00:27:05.465 or, quality that or quality.
NOTE Confidence: 0.9493667
00:27:09.050 --> 00:27:10.910 Oh, here is some examples
NOTE Confidence: 0.9493667
00:27:10.970 --> 00:27:13.470 of, widely used annotation tool.
NOTE Confidence: 0.90872794

00:27:14.570 --> 00:27:15.770 You can see there is,
NOTE Confidence: 0.90872794
00:27:16.090 --> 00:27:16.590 Meditator,
NOTE Confidence: 0.5328799
00:27:16.970 --> 00:27:17.470 Ehost,
NOTE Confidence: 0.9051198
00:27:18.410 --> 00:27:19.609 or a Docana. All of
NOTE Confidence: 0.9051198
00:27:19.609 --> 00:27:21.210 those tools are open source
NOTE Confidence: 0.9051198
00:27:21.210 --> 00:27:22.750 and available on GitHub.
NOTE Confidence: 0.9109878
00:27:23.585 --> 00:27:24.545 Today, I'm gonna,
NOTE Confidence: 0.92101896
00:27:25.185 --> 00:27:27.185 introduce the annotation tool blue
NOTE Confidence: 0.92101896
00:27:27.185 --> 00:27:29.105 is which is implement on
NOTE Confidence: 0.92101896
00:27:29.105 --> 00:27:31.285 the cheap environment, and then
NOTE Confidence: 0.92101896
00:27:31.505 --> 00:27:33.025 each users don't need to
NOTE Confidence: 0.92101896
00:27:33.025 --> 00:27:35.025 install by themselves and could
NOTE Confidence: 0.92101896
00:27:35.025 --> 00:27:36.165 be managed by,
NOTE Confidence: 0.99749565
00:27:36.545 --> 00:27:37.525 admin user.
NOTE Confidence: 0.9129006
00:27:42.240 --> 00:27:43.619 For the Bluetooth,
NOTE Confidence: 0.90882945
00:27:44.320 --> 00:27:45.619 there are several prerequest

NOTE Confidence: 0.9515385
00:27:45.920 --> 00:27:47.200 for the access. The first
NOTE Confidence: 0.9515385
00:27:47.200 --> 00:27:48.320 one, you will need a
NOTE Confidence: 0.9515385
00:27:48.320 --> 00:27:49.460 one HH account
NOTE Confidence: 0.8353273
00:27:49.840 --> 00:27:51.520 and then, adding connect to
NOTE Confidence: 0.8353273
00:27:51.520 --> 00:27:52.020 VPN.
NOTE Confidence: 0.96353203
00:27:52.654 --> 00:27:54.174 So for Mac user, you
NOTE Confidence: 0.96353203
00:27:54.174 --> 00:27:55.294 will need to install a
NOTE Confidence: 0.96353203
00:27:55.294 --> 00:27:56.274 Windows application.
NOTE Confidence: 0.9892286
00:27:56.575 --> 00:27:58.414 And for Windows user, you
NOTE Confidence: 0.9892286
00:27:58.414 --> 00:28:00.195 can use the remote desk,
NOTE Confidence: 0.84028316
00:28:00.654 --> 00:28:02.274 connection to oh, application.
NOTE Confidence: 0.9929996
00:28:04.654 --> 00:28:05.695 First step, you need to
NOTE Confidence: 0.9929996
00:28:05.695 --> 00:28:06.835 connect to the VPN.
NOTE Confidence: 0.993499
00:28:07.630 --> 00:28:08.770 Open the VPN
NOTE Confidence: 0.81415105
00:28:09.309 --> 00:28:11.309 application and then, in the
NOTE Confidence: 0.81415105

00:28:11.309 --> 00:28:13.549 address, type the telecom mute

NOTE Confidence: 0.81415105

00:28:13.549 --> 00:28:15.090 dot y h h dot

NOTE Confidence: 0.88814586

00:28:15.630 --> 00:28:17.410 org backslash y s m.

NOTE Confidence: 0.88814586

00:28:17.549 --> 00:28:19.150 Here, you need to use

NOTE Confidence: 0.88814586

00:28:19.150 --> 00:28:20.750 your Yale Net ID and

NOTE Confidence: 0.88814586

00:28:20.750 --> 00:28:22.130 password to log in.

NOTE Confidence: 0.93926316

00:28:24.335 --> 00:28:25.875 And then once you successfully,

NOTE Confidence: 0.98367715

00:28:26.575 --> 00:28:27.615 log in to the VPN

NOTE Confidence: 0.98367715

00:28:27.615 --> 00:28:29.375 environment, you can open the

NOTE Confidence: 0.98367715

00:28:29.375 --> 00:28:31.934 application and click the add

NOTE Confidence: 0.98367715

00:28:31.934 --> 00:28:32.434 button

NOTE Confidence: 0.9945696

00:28:32.894 --> 00:28:33.955 to add the

NOTE Confidence: 0.9158962

00:28:34.909 --> 00:28:35.730 IP address.

NOTE Confidence: 0.9947412

00:28:36.109 --> 00:28:38.110 It's ten dot forty eight

NOTE Confidence: 0.9947412

00:28:38.110 --> 00:28:39.470 dot one two eight dot,

NOTE Confidence: 0.9947412

00:28:39.710 --> 00:28:40.610 ninety six

NOTE Confidence: 0.9807432
00:28:41.870 --> 00:28:42.850 dot sixty nine.
NOTE Confidence: 0.9992648
00:28:43.950 --> 00:28:44.450 And
NOTE Confidence: 0.9950284
00:28:44.750 --> 00:28:48.029 once the PC successfully added,
NOTE Confidence: 0.9950284
00:28:48.029 --> 00:28:49.809 it will show on the
NOTE Confidence: 0.95813304
00:28:50.495 --> 00:28:52.435 application and then double click
NOTE Confidence: 0.95813304
00:28:52.575 --> 00:28:54.275 to insert your credential.
NOTE Confidence: 0.84003043
00:28:55.215 --> 00:28:57.055 Here, we'll need your one
NOTE Confidence: 0.84003043
00:28:57.055 --> 00:28:58.735 HH ID and the one
NOTE Confidence: 0.84003043
00:28:58.735 --> 00:28:59.555 HH password.
NOTE Confidence: 0.9415577
00:29:03.410 --> 00:29:05.270 Once you, log in to
NOTE Confidence: 0.974134
00:29:06.370 --> 00:29:07.410 the PC, you will see
NOTE Confidence: 0.974134
00:29:07.410 --> 00:29:09.350 a Ubuntu environment.
NOTE Confidence: 0.9973518
00:29:12.929 --> 00:29:14.610 After you get access to
NOTE Confidence: 0.9973518
00:29:14.610 --> 00:29:16.289 that environment, you can use
NOTE Confidence: 0.9973518
00:29:16.289 --> 00:29:17.890 any browser on the left
NOTE Confidence: 0.9973518

00:29:17.890 --> 00:29:19.975 side And on the address,
NOTE Confidence: 0.87591314

00:29:20.435 --> 00:29:22.055 insert the URL, HTTP,
NOTE Confidence: 0.9804883

00:29:22.835 --> 00:29:24.755 local host to open the
NOTE Confidence: 0.9804883

00:29:24.755 --> 00:29:25.815 annotation tool.
NOTE Confidence: 0.99430263

00:29:28.435 --> 00:29:29.875 The first step is to
NOTE Confidence: 0.99430263

00:29:29.875 --> 00:29:31.015 create your account.
NOTE Confidence: 0.95222974

00:29:31.410 --> 00:29:32.770 You always want to have
NOTE Confidence: 0.95222974

00:29:32.770 --> 00:29:35.170 a admin person that creates
NOTE Confidence: 0.95222974

00:29:35.170 --> 00:29:36.070 account first.
NOTE Confidence: 0.985762

00:29:36.530 --> 00:29:37.830 That will be the person
NOTE Confidence: 0.985762

00:29:37.890 --> 00:29:39.090 who can manage the whole
NOTE Confidence: 0.985762

00:29:39.090 --> 00:29:41.090 group and assign the project
NOTE Confidence: 0.985762

00:29:41.090 --> 00:29:42.630 and tasks to each annotators.
NOTE Confidence: 0.9726528

00:29:43.810 --> 00:29:45.570 Please use your email, username,
NOTE Confidence: 0.9726528

00:29:45.570 --> 00:29:46.130 and pass
NOTE Confidence: 0.9900896

00:29:47.245 --> 00:29:49.325 password to sign up. And

NOTE Confidence: 0.9900896
00:29:49.325 --> 00:29:50.465 for the verification
NOTE Confidence: 0.9931561
00:29:50.765 --> 00:29:51.585 code field,
NOTE Confidence: 0.9286913
00:29:52.125 --> 00:29:53.885 we disable that function so
NOTE Confidence: 0.9286913
00:29:53.885 --> 00:29:55.404 you can just enter any
NOTE Confidence: 0.9286913
00:29:55.404 --> 00:29:56.945 four digit number or
NOTE Confidence: 0.99865645
00:29:57.245 --> 00:29:58.544 combination of characters.
NOTE Confidence: 0.9280206
00:30:01.940 --> 00:30:02.900 After you log in to
NOTE Confidence: 0.9280206
00:30:02.900 --> 00:30:04.020 the blue, you will be
NOTE Confidence: 0.9280206
00:30:04.020 --> 00:30:05.460 able to create able to
NOTE Confidence: 0.9280206
00:30:05.460 --> 00:30:07.220 create the project and invite,
NOTE Confidence: 0.9964645
00:30:07.540 --> 00:30:09.320 users to the tool.
NOTE Confidence: 0.94283825
00:30:14.275 --> 00:30:16.835 Once the admin person successfully
NOTE Confidence: 0.94283825
00:30:16.835 --> 00:30:18.515 log in and then he
NOTE Confidence: 0.94283825
00:30:18.515 --> 00:30:19.635 he or she can send
NOTE Confidence: 0.94283825
00:30:19.635 --> 00:30:21.635 the invitation to the other
NOTE Confidence: 0.94283825

00:30:21.635 --> 00:30:22.375 group members,
NOTE Confidence: 0.9915695

00:30:23.475 --> 00:30:25.170 he the person need to
NOTE Confidence: 0.9915695

00:30:25.250 --> 00:30:26.850 click the invite button and
NOTE Confidence: 0.9915695

00:30:26.850 --> 00:30:28.610 then copy the invitation link
NOTE Confidence: 0.9915695

00:30:28.610 --> 00:30:29.510 to each of
NOTE Confidence: 0.9362779

00:30:29.810 --> 00:30:30.470 the annotators.
NOTE Confidence: 0.94595987

00:30:31.410 --> 00:30:32.850 And the annotator need to
NOTE Confidence: 0.94595987

00:30:32.850 --> 00:30:34.530 be use this link to
NOTE Confidence: 0.94595987

00:30:34.530 --> 00:30:36.290 register. Otherwise, they will not
NOTE Confidence: 0.94595987

00:30:36.290 --> 00:30:37.750 be in the same group.
NOTE Confidence: 0.9950011

00:30:41.355 --> 00:30:41.755 And,
NOTE Confidence: 0.9410472

00:30:42.235 --> 00:30:44.175 by click the click on
NOTE Confidence: 0.9410472

00:30:44.475 --> 00:30:45.995 add new project button, you
NOTE Confidence: 0.9410472

00:30:45.995 --> 00:30:47.115 can you will be able
NOTE Confidence: 0.9410472

00:30:47.115 --> 00:30:47.615 to
NOTE Confidence: 0.89295065

00:30:47.915 --> 00:30:50.175 choose your task either NER

NOTE Confidence: 0.89295065
00:30:50.235 --> 00:30:52.415 or NER plus relational extraction.
NOTE Confidence: 0.8959927
00:30:57.270 --> 00:30:59.110 The pro the creative project
NOTE Confidence: 0.8959927
00:30:59.110 --> 00:31:00.549 will show on the front
NOTE Confidence: 0.8959927
00:31:00.549 --> 00:31:01.049 page,
NOTE Confidence: 0.9985282
00:31:01.429 --> 00:31:02.630 and then you will be
NOTE Confidence: 0.9985282
00:31:02.630 --> 00:31:04.090 able to add annotators
NOTE Confidence: 0.98139846
00:31:04.470 --> 00:31:06.330 to the to the project.
NOTE Confidence: 0.9977151
00:31:09.625 --> 00:31:11.245 To add the data source,
NOTE Confidence: 0.9976926
00:31:11.625 --> 00:31:13.065 you can click the data
NOTE Confidence: 0.9976926
00:31:13.065 --> 00:31:13.885 source button
NOTE Confidence: 0.95267946
00:31:14.425 --> 00:31:16.185 button and then choose what
NOTE Confidence: 0.95267946
00:31:16.185 --> 00:31:17.305 kind of format you want
NOTE Confidence: 0.95267946
00:31:17.305 --> 00:31:18.605 to upload to the tool.
NOTE Confidence: 0.95267946
00:31:18.745 --> 00:31:20.505 We access two format. One
NOTE Confidence: 0.95267946
00:31:20.505 --> 00:31:22.265 is txt. That's the plain
NOTE Confidence: 0.95267946

00:31:22.265 --> 00:31:23.730 text without any,
NOTE Confidence: 0.9114534

00:31:24.930 --> 00:31:26.070 entity or relationship,
NOTE Confidence: 0.94469815

00:31:26.850 --> 00:31:29.010 and and pre annotate. You
NOTE Confidence: 0.94469815

00:31:29.010 --> 00:31:30.530 can also choose the blue
NOTE Confidence: 0.94469815

00:31:30.530 --> 00:31:32.130 format. That is a JSON
NOTE Confidence: 0.94469815

00:31:32.130 --> 00:31:34.050 file. You can include the
NOTE Confidence: 0.94469815

00:31:34.050 --> 00:31:35.110 entity or
NOTE Confidence: 0.9061627

00:31:35.410 --> 00:31:35.910 relationship
NOTE Confidence: 0.9908891

00:31:36.290 --> 00:31:37.430 inside that JSON.
NOTE Confidence: 0.99691135

00:31:42.155 --> 00:31:43.674 For each for each project,
NOTE Confidence: 0.99691135

00:31:43.674 --> 00:31:45.355 you can create tasks for
NOTE Confidence: 0.99691135

00:31:45.355 --> 00:31:46.095 the annotators,
NOTE Confidence: 0.9792088

00:31:47.195 --> 00:31:49.674 by adding the by click
NOTE Confidence: 0.9792088

00:31:49.674 --> 00:31:51.470 the add task button.
NOTE Confidence: 0.9549442

00:31:51.770 --> 00:31:53.210 And for each task, you
NOTE Confidence: 0.9549442

00:31:53.210 --> 00:31:55.530 can assign multiple annotators to

NOTE Confidence: 0.9549442
00:31:55.530 --> 00:31:57.130 this one task just as
NOTE Confidence: 0.9549442
00:31:57.130 --> 00:31:58.350 I mentioned before.
NOTE Confidence: 0.982969
00:31:59.370 --> 00:32:00.990 Different annotators can
NOTE Confidence: 0.9483218
00:32:01.690 --> 00:32:04.090 annotate same, sub subgroup of
NOTE Confidence: 0.9483218
00:32:04.090 --> 00:32:05.855 data. That is in order
NOTE Confidence: 0.9489176
00:32:06.395 --> 00:32:08.235 to calculate the agreement among
NOTE Confidence: 0.9489176
00:32:08.235 --> 00:32:08.735 annotators.
NOTE Confidence: 0.66449606
00:32:11.515 --> 00:32:12.335 After Tesla
NOTE Confidence: 0.9517126
00:32:12.635 --> 00:32:14.395 created, you you will be
NOTE Confidence: 0.9517126
00:32:14.395 --> 00:32:15.695 able to start annotation.
NOTE Confidence: 0.99929297
00:32:16.395 --> 00:32:17.135 And then
NOTE Confidence: 0.99088556
00:32:17.740 --> 00:32:19.179 first first thing, you need
NOTE Confidence: 0.99088556
00:32:19.179 --> 00:32:20.779 to define the entity and
NOTE Confidence: 0.99088556
00:32:20.779 --> 00:32:21.279 relationship
NOTE Confidence: 0.99135625
00:32:21.580 --> 00:32:23.100 that you already have in
NOTE Confidence: 0.99135625

00:32:23.100 --> 00:32:24.240 your annotation guideline.
NOTE Confidence: 0.99871886

00:32:25.179 --> 00:32:25.679 And
NOTE Confidence: 0.92935437

00:32:27.500 --> 00:32:29.659 and then after that, highlight
NOTE Confidence: 0.92935437

00:32:29.659 --> 00:32:30.860 the phrase you want to
NOTE Confidence: 0.92935437

00:32:30.860 --> 00:32:32.380 do the annotation and then,
NOTE Confidence: 0.99942076

00:32:33.415 --> 00:32:33.915 choose
NOTE Confidence: 0.9990624

00:32:34.295 --> 00:32:35.495 what kind of entity or
NOTE Confidence: 0.9990624

00:32:35.495 --> 00:32:37.035 relationship you want to
NOTE Confidence: 0.99496365

00:32:37.815 --> 00:32:38.315 annotate.
NOTE Confidence: 0.8821635

00:32:40.455 --> 00:32:41.895 The blue tool can will
NOTE Confidence: 0.8821635

00:32:41.895 --> 00:32:43.575 also provide you a function
NOTE Confidence: 0.8821635

00:32:43.575 --> 00:32:45.735 to calculate the agreement among
NOTE Confidence: 0.8821635

00:32:45.735 --> 00:32:46.475 the annotators.
NOTE Confidence: 0.9619433

00:32:48.490 --> 00:32:50.490 Once the annotator finish and
NOTE Confidence: 0.9619433

00:32:50.570 --> 00:32:51.630 finish the task,
NOTE Confidence: 0.99931264

00:32:52.169 --> 00:32:52.909 you can

NOTE Confidence: 0.96144265
00:32:53.690 --> 00:32:55.529 and finalize them. Then you
NOTE Confidence: 0.96144265
00:32:55.529 --> 00:32:57.470 then you can just use
NOTE Confidence: 0.9908187
00:32:58.169 --> 00:32:59.230 the button to
NOTE Confidence: 0.9475054
00:33:01.345 --> 00:33:02.945 check the agreement among them.
NOTE Confidence: 0.9475054
00:33:02.945 --> 00:33:04.065 It will give you a
NOTE Confidence: 0.9475054
00:33:04.065 --> 00:33:05.505 f one score for both
NOTE Confidence: 0.9475054
00:33:05.505 --> 00:33:06.644 entity and relationship.
NOTE Confidence: 0.9975917
00:33:10.225 --> 00:33:11.664 Then I will have a
NOTE Confidence: 0.9975917
00:33:11.664 --> 00:33:13.845 quick demo for the process.
NOTE Confidence: 0.96204674
00:34:18.180 --> 00:34:19.620 Okay. Okay. As I mentioned,
NOTE Confidence: 0.96204674
00:34:19.620 --> 00:34:21.300 you just, connect to the
NOTE Confidence: 0.96204674
00:34:21.300 --> 00:34:22.600 VPN and then
NOTE Confidence: 0.9680562
00:34:24.815 --> 00:34:25.795 type the password.
NOTE Confidence: 0.8685778
00:34:31.775 --> 00:34:33.394 And here, then open
NOTE Confidence: 0.9553802
00:34:34.974 --> 00:34:36.755 open the Windows app.
NOTE Confidence: 0.99525183

00:34:37.535 --> 00:34:38.755 Click to the
NOTE Confidence: 0.99862397
00:34:42.310 --> 00:34:42.810 server
NOTE Confidence: 0.99026
00:34:43.190 --> 00:34:43.850 we have.
NOTE Confidence: 0.93750894
00:35:08.660 --> 00:35:10.440 And then open the browser.
NOTE Confidence: 0.8941952
00:35:24.475 --> 00:35:25.835 Here, you can sign into
NOTE Confidence: 0.8941952
00:35:25.835 --> 00:35:26.655 your account.
NOTE Confidence: 0.993805
00:35:28.795 --> 00:35:30.155 To create a new project,
NOTE Confidence: 0.993805
00:35:30.155 --> 00:35:31.275 you can just click this
NOTE Confidence: 0.993805
00:35:31.275 --> 00:35:32.255 button and
NOTE Confidence: 0.93605876
00:35:33.195 --> 00:35:34.795 type the project name and
NOTE Confidence: 0.93605876
00:35:34.795 --> 00:35:36.335 select the project type.
NOTE Confidence: 0.95998406
00:35:37.599 --> 00:35:39.040 Here, I already created a
NOTE Confidence: 0.95998406
00:35:39.040 --> 00:35:40.339 demo project project.
NOTE Confidence: 0.9951987
00:35:40.880 --> 00:35:41.839 I want to,
NOTE Confidence: 0.96777415
00:35:42.320 --> 00:35:44.099 import the data source here.
NOTE Confidence: 0.96777415
00:35:44.320 --> 00:35:45.520 So I just click this

NOTE Confidence: 0.96777415
00:35:45.520 --> 00:35:46.020 button,
NOTE Confidence: 0.99879885
00:35:46.480 --> 00:35:47.700 and then I
NOTE Confidence: 0.9042556
00:35:48.320 --> 00:35:50.275 download some The notes I
NOTE Confidence: 0.9042556
00:35:50.275 --> 00:35:51.815 need to annotate it from
NOTE Confidence: 0.9042556
00:35:51.875 --> 00:35:52.855 the chip environment,
NOTE Confidence: 0.98639196
00:35:53.715 --> 00:35:55.155 and then I drag drag
NOTE Confidence: 0.98639196
00:35:55.155 --> 00:35:56.775 it to here.
NOTE Confidence: 0.8387658
00:35:57.555 --> 00:35:58.775 This is the tab,
NOTE Confidence: 0.8175417
00:35:59.235 --> 00:36:00.355 tip c file, so I
NOTE Confidence: 0.8175417
00:36:00.355 --> 00:36:01.575 just choose text,
NOTE Confidence: 0.9967623
00:36:01.955 --> 00:36:03.175 and then I confirm.
NOTE Confidence: 0.98565817
00:36:07.719 --> 00:36:09.160 For each of the project,
NOTE Confidence: 0.98565817
00:36:09.160 --> 00:36:10.859 you can add the annotator,
NOTE Confidence: 0.9038088
00:36:11.560 --> 00:36:13.320 and that's the pit per
NOTE Confidence: 0.9038088
00:36:13.400 --> 00:36:14.920 that's the person within your
NOTE Confidence: 0.9038088

00:36:14.920 --> 00:36:15.420 group.
NOTE Confidence: 0.91483307

00:36:18.825 --> 00:36:20.505 And then you just, create
NOTE Confidence: 0.91483307

00:36:20.505 --> 00:36:22.285 annotation task for them.
NOTE Confidence: 0.96375895

00:36:26.745 --> 00:36:28.685 You can choose, multiple annotators
NOTE Confidence: 0.96375895

00:36:28.905 --> 00:36:29.405 here.
NOTE Confidence: 0.99718744

00:36:33.079 --> 00:36:34.440 And then you go to
NOTE Confidence: 0.99718744

00:36:34.440 --> 00:36:35.420 the file.
NOTE Confidence: 0.9874424

00:36:36.119 --> 00:36:37.719 On this side, you can
NOTE Confidence: 0.9874424

00:36:37.719 --> 00:36:39.719 define the entity. For example,
NOTE Confidence: 0.9874424

00:36:39.719 --> 00:36:41.180 we we want to choose
NOTE Confidence: 0.91484576

00:36:42.119 --> 00:36:43.900 we want to define problem.
NOTE Confidence: 0.9917244

00:36:46.075 --> 00:36:47.355 And then you can,
NOTE Confidence: 0.9980949

00:36:47.755 --> 00:36:49.135 start to do the annotation.
NOTE Confidence: 0.98144066

00:36:51.915 --> 00:36:53.675 Yeah. Basically, that's the whole
NOTE Confidence: 0.98144066

00:36:53.675 --> 00:36:55.195 process for the how you're
NOTE Confidence: 0.98144066

00:36:55.195 --> 00:36:56.475 gonna do the annotation and

NOTE Confidence: 0.98144066
00:36:56.475 --> 00:36:57.855 how to use our tool.
NOTE Confidence: 0.961466
00:37:05.900 --> 00:37:07.119 Yeah. Any questions?
NOTE Confidence: 0.9366959
00:37:08.539 --> 00:37:09.039 So
NOTE Confidence: 0.7439517
00:37:10.299 --> 00:37:12.059 how we can import our
NOTE Confidence: 0.7439517
00:37:12.059 --> 00:37:13.315 own data to this? This
NOTE Confidence: 0.7439517
00:37:13.315 --> 00:37:14.435 because I think this this
NOTE Confidence: 0.7439517
00:37:14.435 --> 00:37:15.815 is your server. Right?
NOTE Confidence: 0.61711377
00:37:17.235 --> 00:37:17.715 Here, as a
NOTE Confidence: 0.9524396
00:37:18.515 --> 00:37:19.975 the we can download
NOTE Confidence: 0.45951208
00:37:21.315 --> 00:37:21.635 the
NOTE Confidence: 0.89796615
00:37:22.114 --> 00:37:23.635 the team also mentioned we
NOTE Confidence: 0.89796615
00:37:23.635 --> 00:37:25.140 use the cheap environment. Right?
NOTE Confidence: 0.89796615
00:37:25.219 --> 00:37:27.000 The Camino, we can upload
NOTE Confidence: 0.89796615
00:37:27.140 --> 00:37:28.739 the own data mod their
NOTE Confidence: 0.89796615
00:37:28.739 --> 00:37:30.260 their own data to their
NOTE Confidence: 0.89796615

00:37:30.260 --> 00:37:32.339 environment, and this server will
NOTE Confidence: 0.89796615

00:37:32.339 --> 00:37:33.640 connect to the Camino.
NOTE Confidence: 0.99442476

00:37:34.099 --> 00:37:35.940 You can download that data
NOTE Confidence: 0.99442476

00:37:35.940 --> 00:37:37.559 from Camino environment.
NOTE Confidence: 0.7042018

00:37:37.940 --> 00:37:38.680 That will.
NOTE Confidence: 0.94260305

00:37:46.745 --> 00:37:48.505 Not yet. Right now, because,
NOTE Confidence: 0.94260305

00:37:48.745 --> 00:37:50.825 this tool will host in
NOTE Confidence: 0.94260305

00:37:50.825 --> 00:37:52.205 a secure environment,
NOTE Confidence: 0.6615233

00:37:52.890 --> 00:37:54.250 while coming to us because
NOTE Confidence: 0.6615233

00:37:54.250 --> 00:37:55.530 there are lots of PHI
NOTE Confidence: 0.6615233

00:37:55.530 --> 00:37:56.030 information.
NOTE Confidence: 0.9988219

00:37:56.650 --> 00:37:57.150 So
NOTE Confidence: 0.98514336

00:37:57.530 --> 00:37:59.069 that's the purpose,
NOTE Confidence: 0.9075419

00:37:59.609 --> 00:38:01.369 we hosted there. So for
NOTE Confidence: 0.9075419

00:38:01.369 --> 00:38:02.489 example, let's say there are
NOTE Confidence: 0.9075419

00:38:02.489 --> 00:38:02.969 other,

NOTE Confidence: 0.7920805
00:38:04.250 --> 00:38:06.030 and publicly available datasets.
NOTE Confidence: 0.8579039
00:38:06.515 --> 00:38:07.555 We really want to annotate
NOTE Confidence: 0.8579039
00:38:07.555 --> 00:38:08.915 them so they're they're able
NOTE Confidence: 0.8579039
00:38:08.915 --> 00:38:10.275 to be for us to
NOTE Confidence: 0.8579039
00:38:10.275 --> 00:38:10.775 ask,
NOTE Confidence: 0.68315446
00:38:11.755 --> 00:38:12.255 to,
NOTE Confidence: 0.78291494
00:38:12.755 --> 00:38:14.375 upload it to and then
NOTE Confidence: 0.78291494
00:38:14.435 --> 00:38:14.935 from
NOTE Confidence: 0.933078
00:38:15.475 --> 00:38:17.175 to this server. Right? Yes.
NOTE Confidence: 0.8365646
00:38:18.114 --> 00:38:19.315 Well, I want to think
NOTE Confidence: 0.8365646
00:38:19.315 --> 00:38:20.760 of if you try to
NOTE Confidence: 0.8365646
00:38:20.760 --> 00:38:22.700 annotate public data, don't use.
NOTE Confidence: 0.8614464
00:38:23.400 --> 00:38:24.680 I I I think what
NOTE Confidence: 0.8614464
00:38:24.680 --> 00:38:25.640 what we can do is
NOTE Confidence: 0.8614464
00:38:25.640 --> 00:38:26.600 we set up a blue
NOTE Confidence: 0.8614464

00:38:26.600 --> 00:38:27.820 in a open,
NOTE Confidence: 0.8541854

00:38:28.680 --> 00:38:30.360 public website, then you can
NOTE Confidence: 0.8541854

00:38:30.360 --> 00:38:31.960 just go to over there.
NOTE Confidence: 0.8541854

00:38:31.960 --> 00:38:32.760 Like, I put it in
NOTE Confidence: 0.8541854

00:38:32.760 --> 00:38:33.945 the spin up, then you
NOTE Confidence: 0.8541854

00:38:33.945 --> 00:38:35.465 can just upload because there's
NOTE Confidence: 0.8541854

00:38:35.465 --> 00:38:37.225 no sensitive data. We can
NOTE Confidence: 0.8541854

00:38:37.225 --> 00:38:38.825 just make another instance of
NOTE Confidence: 0.8541854

00:38:38.825 --> 00:38:40.285 the loop of public data.
NOTE Confidence: 0.93353

00:38:40.585 --> 00:38:41.225 Because this,
NOTE Confidence: 0.8208353

00:38:41.705 --> 00:38:43.245 we install in the communal,
NOTE Confidence: 0.8841377

00:38:43.545 --> 00:38:45.065 in the CHP to support
NOTE Confidence: 0.8841377

00:38:45.065 --> 00:38:46.540 this annotation of of.
NOTE Confidence: 0.8801747

00:38:46.920 --> 00:38:47.800 And if there are public
NOTE Confidence: 0.8801747

00:38:47.800 --> 00:38:49.000 data, well, we can't just
NOTE Confidence: 0.8801747

00:38:49.000 --> 00:38:49.960 set up another because it's

NOTE Confidence: 0.8801747
00:38:49.960 --> 00:38:51.320 a web application. We just
NOTE Confidence: 0.8801747
00:38:51.320 --> 00:38:52.780 set up another web application
NOTE Confidence: 0.8801747
00:38:53.000 --> 00:38:54.280 in this thing, a public
NOTE Confidence: 0.8801747
00:38:54.280 --> 00:38:55.560 space we can. Yeah. We
NOTE Confidence: 0.8801747
00:38:55.560 --> 00:38:56.839 can discuss that. And we
NOTE Confidence: 0.8801747
00:38:56.839 --> 00:38:58.040 should we should ask them
NOTE Confidence: 0.8801747
00:38:58.040 --> 00:38:59.739 to set up that
NOTE Confidence: 0.9589169
00:39:00.055 --> 00:39:00.555 specific
NOTE Confidence: 0.73303306
00:39:00.855 --> 00:39:02.055 list that the public list,
NOTE Confidence: 0.73303306
00:39:02.055 --> 00:39:03.114 or is it available?
NOTE Confidence: 0.8655469
00:39:04.535 --> 00:39:06.375 We have not, but you
NOTE Confidence: 0.8655469
00:39:06.375 --> 00:39:07.815 can contact us. Maybe we
NOTE Confidence: 0.8655469
00:39:07.815 --> 00:39:08.695 just give you a copy.
NOTE Confidence: 0.8655469
00:39:08.695 --> 00:39:09.895 We can stop by ourselves.
NOTE Confidence: 0.8655469
00:39:09.895 --> 00:39:10.855 But right now, we we
NOTE Confidence: 0.8655469

00:39:10.855 --> 00:39:12.855 didn't really distribute this package

NOTE Confidence: 0.8655469

00:39:12.855 --> 00:39:14.860 frame. We're just sitting up

NOTE Confidence: 0.8655469

00:39:14.860 --> 00:39:15.520 before our.

NOTE Confidence: 0.86946565

00:39:34.595 --> 00:39:35.975 It it's just a different,

NOTE Confidence: 0.9848068

00:39:36.435 --> 00:39:36.935 tools.

NOTE Confidence: 0.9950956

00:39:37.555 --> 00:39:38.055 Yeah.

NOTE Confidence: 0.87479395

00:39:50.739 --> 00:39:52.660 Thanks, Silja. I'm gonna be

NOTE Confidence: 0.87479395

00:39:52.660 --> 00:39:53.480 very quick.

NOTE Confidence: 0.7841471

00:39:55.105 --> 00:39:56.805 And machine gun mode on.

NOTE Confidence: 0.7841471

00:39:56.864 --> 00:39:57.364 Okay.

NOTE Confidence: 0.85894346

00:39:58.545 --> 00:40:00.465 Yeah. Doctor Shu already discussed

NOTE Confidence: 0.85894346

00:40:00.465 --> 00:40:01.905 about the difference between,

NOTE Confidence: 0.77699566

00:40:02.305 --> 00:40:03.364 BERT and LAMA.

NOTE Confidence: 0.9959889

00:40:03.825 --> 00:40:04.325 And

NOTE Confidence: 0.8966075

00:40:05.185 --> 00:40:06.245 summarize everything,

NOTE Confidence: 0.9703491

00:40:06.760 --> 00:40:07.720 there is a trade off

NOTE Confidence: 0.9703491
00:40:07.720 --> 00:40:09.500 between performance, computational
NOTE Confidence: 0.99038726
00:40:09.800 --> 00:40:11.100 resources, and time.
NOTE Confidence: 0.96521854
00:40:11.720 --> 00:40:13.640 Okay? So you have, you
NOTE Confidence: 0.96521854
00:40:13.640 --> 00:40:14.860 need better performance.
NOTE Confidence: 0.89365745
00:40:15.640 --> 00:40:17.400 The computational resources are there.
NOTE Confidence: 0.89365745
00:40:17.400 --> 00:40:18.680 Go for llama models, high
NOTE Confidence: 0.89365745
00:40:18.680 --> 00:40:19.500 billion models.
NOTE Confidence: 0.9908056
00:40:19.815 --> 00:40:21.655 Across a wide variety of
NOTE Confidence: 0.9908056
00:40:21.655 --> 00:40:23.594 tasks, they would work well.
NOTE Confidence: 0.9846914
00:40:23.895 --> 00:40:24.855 But if time is a
NOTE Confidence: 0.9846914
00:40:24.855 --> 00:40:26.474 concern, he, projected,
NOTE Confidence: 0.9426048
00:40:26.935 --> 00:40:28.214 you know, issue of speed
NOTE Confidence: 0.9426048
00:40:28.214 --> 00:40:29.575 between BART models and the
NOTE Confidence: 0.9426048
00:40:29.575 --> 00:40:31.015 last language models. It is
NOTE Confidence: 0.9426048
00:40:31.015 --> 00:40:32.375 up to twenty to thirty
NOTE Confidence: 0.9426048

00:40:32.375 --> 00:40:33.594 times slower.
NOTE Confidence: 0.97556
00:40:34.080 --> 00:40:35.120 So if that is a
NOTE Confidence: 0.97556
00:40:35.120 --> 00:40:36.400 concern, you need to switch
NOTE Confidence: 0.97556
00:40:36.400 --> 00:40:37.760 to BERT models. So I'm
NOTE Confidence: 0.97556
00:40:37.760 --> 00:40:39.220 gonna talk about the clinical
NOTE Confidence: 0.97556
00:40:39.280 --> 00:40:40.980 information extraction system
NOTE Confidence: 0.9566314
00:40:41.440 --> 00:40:43.280 where we have developed both
NOTE Confidence: 0.9566314
00:40:43.280 --> 00:40:45.520 the BERT and LAMA based
NOTE Confidence: 0.9566314
00:40:45.520 --> 00:40:47.540 large language models for you
NOTE Confidence: 0.990504
00:40:48.725 --> 00:40:50.325 in such a way that
NOTE Confidence: 0.990504
00:40:50.325 --> 00:40:51.705 you do not have programming
NOTE Confidence: 0.990504
00:40:51.844 --> 00:40:53.864 experience, you have some programming
NOTE Confidence: 0.990504
00:40:54.005 --> 00:40:55.364 experience, or you are a
NOTE Confidence: 0.990504
00:40:55.364 --> 00:40:56.185 pro programmer.
NOTE Confidence: 0.8905734
00:40:56.805 --> 00:40:57.305 Anyway,
NOTE Confidence: 0.99357975
00:40:57.925 --> 00:40:59.364 we have features that will

NOTE Confidence: 0.99357975
00:40:59.364 --> 00:41:00.930 help you take it and
NOTE Confidence: 0.99357975
00:41:00.930 --> 00:41:03.250 customize it to whatever task
NOTE Confidence: 0.99357975
00:41:03.250 --> 00:41:04.130 you want to use it
NOTE Confidence: 0.99357975
00:41:04.130 --> 00:41:04.630 for.
NOTE Confidence: 0.9995298
00:41:05.250 --> 00:41:06.070 And that
NOTE Confidence: 0.99167466
00:41:07.010 --> 00:41:09.190 is what we call Kiwi.
NOTE Confidence: 0.99167466
00:41:09.410 --> 00:41:11.170 Okay? So we are building
NOTE Confidence: 0.99167466
00:41:11.170 --> 00:41:12.609 Kiwi. The one pipeline that
NOTE Confidence: 0.99167466
00:41:12.609 --> 00:41:14.210 I'm currently gonna show you
NOTE Confidence: 0.99167466
00:41:14.210 --> 00:41:15.894 that is set for all
NOTE Confidence: 0.99167466
00:41:15.894 --> 00:41:17.255 these sort of use cases
NOTE Confidence: 0.99167466
00:41:17.255 --> 00:41:18.474 that I'm talking about
NOTE Confidence: 0.9814681
00:41:18.775 --> 00:41:20.875 is a general clinical information
NOTE Confidence: 0.9814681
00:41:21.174 --> 00:41:22.234 extraction pipeline.
NOTE Confidence: 0.97451514
00:41:22.614 --> 00:41:24.134 I also have things coming
NOTE Confidence: 0.97451514

00:41:24.134 --> 00:41:25.174 up for you, and if
NOTE Confidence: 0.97451514

00:41:25.174 --> 00:41:26.694 you have suggestions or something
NOTE Confidence: 0.97451514

00:41:26.694 --> 00:41:27.654 that you have been really
NOTE Confidence: 0.97451514

00:41:27.654 --> 00:41:28.890 working on, it's a real
NOTE Confidence: 0.97451514

00:41:29.130 --> 00:41:30.410 need of the time, let
NOTE Confidence: 0.97451514

00:41:30.410 --> 00:41:31.529 us know, and then we
NOTE Confidence: 0.97451514

00:41:31.529 --> 00:41:33.049 would work on developing those
NOTE Confidence: 0.97451514

00:41:33.049 --> 00:41:33.549 things.
NOTE Confidence: 0.9933422

00:41:34.489 --> 00:41:34.989 Okay.
NOTE Confidence: 0.9912438

00:41:35.930 --> 00:41:37.369 We have the clinical notes.
NOTE Confidence: 0.9912438

00:41:37.369 --> 00:41:38.329 We need to do some
NOTE Confidence: 0.9912438

00:41:38.329 --> 00:41:38.829 preprocessing,
NOTE Confidence: 0.8053

00:41:39.289 --> 00:41:39.789 deidentification,
NOTE Confidence: 0.93608314

00:41:40.410 --> 00:41:42.010 these sort of things. Doctor
NOTE Confidence: 0.93608314

00:41:42.010 --> 00:41:43.950 Shu mentioned named entity recognition
NOTE Confidence: 0.968941

00:41:44.535 --> 00:41:46.694 followed by relation extraction, then

NOTE Confidence: 0.968941
00:41:46.694 --> 00:41:48.135 there is this concept mapping
NOTE Confidence: 0.968941
00:41:48.135 --> 00:41:49.275 or concept normalization.
NOTE Confidence: 0.9705667
00:41:50.214 --> 00:41:52.775 Finally, post process it and
NOTE Confidence: 0.9705667
00:41:52.775 --> 00:41:53.275 get
NOTE Confidence: 0.9949367
00:41:53.655 --> 00:41:55.114 all the structured data
NOTE Confidence: 0.9997191
00:41:55.494 --> 00:41:56.795 from the unstructured
NOTE Confidence: 0.99086565
00:41:57.094 --> 00:41:58.859 clinical notes. So that is
NOTE Confidence: 0.99086565
00:41:58.859 --> 00:42:00.859 the basic block diagram of
NOTE Confidence: 0.99086565
00:42:00.859 --> 00:42:03.279 any clinical information extraction pipeline.
NOTE Confidence: 0.9157201
00:42:05.739 --> 00:42:06.700 I don't need to go
NOTE Confidence: 0.9157201
00:42:06.700 --> 00:42:08.400 over this named entity recognition,
NOTE Confidence: 0.9830013
00:42:08.700 --> 00:42:10.000 identify the boundaries,
NOTE Confidence: 0.96327096
00:42:10.475 --> 00:42:13.035 relation, identify the relationship between
NOTE Confidence: 0.96327096
00:42:13.035 --> 00:42:13.695 the entities,
NOTE Confidence: 0.9220454
00:42:13.995 --> 00:42:14.815 and normalization,
NOTE Confidence: 0.97358716

00:42:15.195 --> 00:42:16.395 doctors write the same thing
NOTE Confidence: 0.97358716

00:42:16.395 --> 00:42:18.235 in hundred different types. High
NOTE Confidence: 0.97358716

00:42:18.235 --> 00:42:20.075 VP, hypertension, all these are
NOTE Confidence: 0.97358716

00:42:20.075 --> 00:42:21.195 the same. Right? So you
NOTE Confidence: 0.97358716

00:42:21.195 --> 00:42:22.155 need to get it to
NOTE Confidence: 0.97358716

00:42:22.155 --> 00:42:23.215 another standardized
NOTE Confidence: 0.94785017

00:42:24.000 --> 00:42:26.500 vocabulary, terminology like ICD, SNOMED,
NOTE Confidence: 0.94785017

00:42:26.560 --> 00:42:27.760 these sort of things. That
NOTE Confidence: 0.94785017

00:42:27.760 --> 00:42:29.060 is what concept normalization
NOTE Confidence: 0.9637255

00:42:29.360 --> 00:42:31.120 does. All these three things
NOTE Confidence: 0.9637255

00:42:31.120 --> 00:42:31.940 comes together.
NOTE Confidence: 0.99621046

00:42:32.239 --> 00:42:33.440 That is where you take
NOTE Confidence: 0.99621046

00:42:33.440 --> 00:42:35.120 unstructured data and get your
NOTE Confidence: 0.99621046

00:42:35.120 --> 00:42:36.739 structured thing out of it.
NOTE Confidence: 0.99494696

00:42:37.485 --> 00:42:39.485 Okay. What does our general
NOTE Confidence: 0.99494696

00:42:39.485 --> 00:42:39.985 clinical,

NOTE Confidence: 0.9980547

00:42:40.605 --> 00:42:42.945 information extraction pipeline give you?

NOTE Confidence: 0.98769367

00:42:43.645 --> 00:42:45.485 We mainly focused on four

NOTE Confidence: 0.98769367

00:42:45.485 --> 00:42:46.305 main entities,

NOTE Confidence: 0.99896747

00:42:46.685 --> 00:42:48.225 medical problem, treatment,

NOTE Confidence: 0.9916277

00:42:48.525 --> 00:42:49.745 drug, and test.

NOTE Confidence: 0.9125634

00:42:50.349 --> 00:42:52.190 Right? So you our Kiwi

NOTE Confidence: 0.9125634

00:42:52.190 --> 00:42:53.550 tool will give you all

NOTE Confidence: 0.9125634

00:42:53.550 --> 00:42:54.450 these four

NOTE Confidence: 0.988354

00:42:54.829 --> 00:42:56.670 main types of entities, but

NOTE Confidence: 0.988354

00:42:56.670 --> 00:42:58.349 these entities are not just

NOTE Confidence: 0.988354

00:42:58.349 --> 00:42:59.730 by themselves. Right?

NOTE Confidence: 0.9827129

00:43:00.109 --> 00:43:01.390 When you are talking about

NOTE Confidence: 0.9827129

00:43:01.390 --> 00:43:02.589 a drug, you have things

NOTE Confidence: 0.9827129

00:43:02.589 --> 00:43:04.415 like the strength, the dosage,

NOTE Confidence: 0.9827129

00:43:04.475 --> 00:43:06.155 the duration, the route, all

NOTE Confidence: 0.9827129

00:43:06.155 --> 00:43:07.835 these things are important. And
NOTE Confidence: 0.9827129

00:43:07.835 --> 00:43:09.035 we need to connect that
NOTE Confidence: 0.9827129

00:43:09.035 --> 00:43:10.815 specific drug to that specific
NOTE Confidence: 0.9827129

00:43:10.875 --> 00:43:11.935 route or specific
NOTE Confidence: 0.9989472

00:43:12.315 --> 00:43:13.535 strength or dosage
NOTE Confidence: 0.9950887

00:43:13.915 --> 00:43:16.075 to actually identify what has
NOTE Confidence: 0.9950887

00:43:16.075 --> 00:43:18.309 the doctor written about giving
NOTE Confidence: 0.9950887

00:43:18.309 --> 00:43:20.089 those information to that patient.
NOTE Confidence: 0.9902346

00:43:20.869 --> 00:43:22.150 So we have a bunch
NOTE Confidence: 0.9902346

00:43:22.150 --> 00:43:23.670 of main entities, and we
NOTE Confidence: 0.9902346

00:43:23.670 --> 00:43:25.130 have a bunch of modifiers
NOTE Confidence: 0.93025255

00:43:25.430 --> 00:43:27.270 that correspond to those main
NOTE Confidence: 0.93025255

00:43:27.270 --> 00:43:27.770 entities.
NOTE Confidence: 0.98502547

00:43:28.309 --> 00:43:28.809 Altogether,
NOTE Confidence: 0.9967164

00:43:29.430 --> 00:43:30.950 this is what Kiwi is
NOTE Confidence: 0.9967164

00:43:30.950 --> 00:43:32.864 gonna extract for you. I

NOTE Confidence: 0.9967164
00:43:32.864 --> 00:43:34.065 know many of the things
NOTE Confidence: 0.9967164
00:43:34.065 --> 00:43:35.285 that you may be needing
NOTE Confidence: 0.9533341
00:43:35.665 --> 00:43:37.525 might be missing from this,
NOTE Confidence: 0.9533341
00:43:37.665 --> 00:43:38.945 but we will. If there
NOTE Confidence: 0.9533341
00:43:38.945 --> 00:43:40.225 are some other cases that
NOTE Confidence: 0.9533341
00:43:40.225 --> 00:43:41.605 you would like to extract,
NOTE Confidence: 0.9533341
00:43:41.745 --> 00:43:43.185 we may in future think
NOTE Confidence: 0.9533341
00:43:43.185 --> 00:43:45.219 about incorporating that. So So
NOTE Confidence: 0.9533341
00:43:45.219 --> 00:43:46.580 for medical problem, you have
NOTE Confidence: 0.9533341
00:43:46.580 --> 00:43:48.500 the severity, the condition, the
NOTE Confidence: 0.9533341
00:43:48.500 --> 00:43:50.100 uncertainty, who is the subject,
NOTE Confidence: 0.9533341
00:43:50.100 --> 00:43:51.300 whether is it really talking
NOTE Confidence: 0.9533341
00:43:51.300 --> 00:43:52.340 about the patient or his
NOTE Confidence: 0.9533341
00:43:52.340 --> 00:43:53.860 family because we can see
NOTE Confidence: 0.9533341
00:43:53.860 --> 00:43:54.980 all these sort of things
NOTE Confidence: 0.9533341

00:43:54.980 --> 00:43:56.200 appearing in the notes.
NOTE Confidence: 0.96693075

00:43:56.500 --> 00:43:58.739 Whether that particular problem is
NOTE Confidence: 0.96693075

00:43:58.739 --> 00:44:00.515 negated or not, So this
NOTE Confidence: 0.96693075

00:44:00.515 --> 00:44:01.555 is how it is. So
NOTE Confidence: 0.96693075

00:44:01.555 --> 00:44:02.594 we have the four main
NOTE Confidence: 0.96693075

00:44:02.594 --> 00:44:04.695 entities and all these modifiers.
NOTE Confidence: 0.9541392

00:44:07.635 --> 00:44:09.875 Going very briefly. So YuJa
NOTE Confidence: 0.9541392

00:44:09.875 --> 00:44:11.475 mentioned about the annotation. Let's
NOTE Confidence: 0.9541392

00:44:11.475 --> 00:44:12.835 think when you annotate, this
NOTE Confidence: 0.9541392

00:44:12.835 --> 00:44:14.035 is on the top figure
NOTE Confidence: 0.9541392

00:44:14.035 --> 00:44:15.495 is something that you get.
NOTE Confidence: 0.9541392

00:44:15.780 --> 00:44:17.140 Now suppose you are using
NOTE Confidence: 0.9541392

00:44:17.140 --> 00:44:18.660 a large language model, it
NOTE Confidence: 0.9541392

00:44:18.660 --> 00:44:20.520 understands the language of prompts.
NOTE Confidence: 0.9541392

00:44:20.660 --> 00:44:22.420 Right? And doctor Xu covered
NOTE Confidence: 0.9541392

00:44:22.420 --> 00:44:23.460 this, how to write a

NOTE Confidence: 0.9541392
00:44:23.460 --> 00:44:24.739 proper prompt for a named
NOTE Confidence: 0.9541392
00:44:24.739 --> 00:44:25.640 entity recognition.
NOTE Confidence: 0.9890417
00:44:26.100 --> 00:44:27.560 So you define the task.
NOTE Confidence: 0.9751695
00:44:28.035 --> 00:44:29.555 We want to identify medical
NOTE Confidence: 0.9751695
00:44:29.555 --> 00:44:31.954 problems, treatment test, and, other
NOTE Confidence: 0.9751695
00:44:31.954 --> 00:44:33.635 things, and then you give
NOTE Confidence: 0.9751695
00:44:33.635 --> 00:44:34.994 us how you need the
NOTE Confidence: 0.9751695
00:44:34.994 --> 00:44:36.055 output. That is
NOTE Confidence: 0.980871
00:44:36.355 --> 00:44:37.954 for making your programming life
NOTE Confidence: 0.980871
00:44:37.954 --> 00:44:39.075 easy to take it in
NOTE Confidence: 0.980871
00:44:39.075 --> 00:44:40.914 a particular output format so
NOTE Confidence: 0.980871
00:44:40.914 --> 00:44:42.035 that you can convert it
NOTE Confidence: 0.980871
00:44:42.035 --> 00:44:43.980 and evaluate it fast. So
NOTE Confidence: 0.980871
00:44:43.980 --> 00:44:45.420 that is the output guideline
NOTE Confidence: 0.980871
00:44:45.420 --> 00:44:45.920 markup.
NOTE Confidence: 0.9837408

00:44:46.380 --> 00:44:48.300 Then you define each entity
NOTE Confidence: 0.9837408

00:44:48.300 --> 00:44:49.580 because we have developed the
NOTE Confidence: 0.9837408

00:44:49.580 --> 00:44:51.420 annotation guidelines, and that is
NOTE Confidence: 0.9837408

00:44:51.420 --> 00:44:53.360 how the humans actually annotate.
NOTE Confidence: 0.9837408

00:44:53.420 --> 00:44:54.700 So the model should also
NOTE Confidence: 0.9837408

00:44:54.700 --> 00:44:55.739 know what is how the
NOTE Confidence: 0.9837408

00:44:55.739 --> 00:44:57.585 humans have annotated. Otherwise, how
NOTE Confidence: 0.9837408

00:44:57.585 --> 00:44:58.864 do you compare that gold
NOTE Confidence: 0.9837408

00:44:58.864 --> 00:45:01.105 human standard annotated data with
NOTE Confidence: 0.9837408

00:45:01.105 --> 00:45:02.385 what is what the model
NOTE Confidence: 0.9837408

00:45:02.385 --> 00:45:04.305 is giving? So whatever information
NOTE Confidence: 0.9837408

00:45:04.305 --> 00:45:05.685 you are giving the human,
NOTE Confidence: 0.99146193

00:45:05.985 --> 00:45:07.505 you also give that to
NOTE Confidence: 0.99146193

00:45:07.505 --> 00:45:08.085 a model
NOTE Confidence: 0.9587802

00:45:08.465 --> 00:45:09.985 in the terms of entity
NOTE Confidence: 0.9587802

00:45:09.985 --> 00:45:10.485 definitions.

NOTE Confidence: 0.96575946
00:45:11.049 --> 00:45:13.210 And then annotation guidelines. We
NOTE Confidence: 0.96575946
00:45:13.210 --> 00:45:14.890 talked about okay. Annotate only
NOTE Confidence: 0.96575946
00:45:14.890 --> 00:45:16.569 complete noun phrases shouldn't be
NOTE Confidence: 0.96575946
00:45:16.569 --> 00:45:18.650 partial, complete adjective phrases. These
NOTE Confidence: 0.96575946
00:45:18.650 --> 00:45:20.089 sort of things that are
NOTE Confidence: 0.96575946
00:45:20.089 --> 00:45:21.609 there in the annotation guideline
NOTE Confidence: 0.96575946
00:45:21.609 --> 00:45:23.130 that you developed is also
NOTE Confidence: 0.96575946
00:45:23.130 --> 00:45:24.430 provided to the model.
NOTE Confidence: 0.99222976
00:45:25.364 --> 00:45:27.125 Now then you build your
NOTE Confidence: 0.99222976
00:45:27.125 --> 00:45:28.805 training data by showing the
NOTE Confidence: 0.99222976
00:45:28.805 --> 00:45:30.505 model a bunch of examples.
NOTE Confidence: 0.96978605
00:45:30.885 --> 00:45:32.485 Suppose your input is at
NOTE Confidence: 0.96978605
00:45:32.485 --> 00:45:33.765 the time of admission, he
NOTE Confidence: 0.96978605
00:45:33.765 --> 00:45:36.005 denied fever, dysphoria, whatever it
NOTE Confidence: 0.96978605
00:45:36.005 --> 00:45:36.665 is. So
NOTE Confidence: 0.97492766

00:45:37.070 --> 00:45:38.270 how does the model provide
NOTE Confidence: 0.97492766

00:45:38.270 --> 00:45:39.550 you the output? So it
NOTE Confidence: 0.97492766

00:45:39.550 --> 00:45:41.150 should say that span class
NOTE Confidence: 0.97492766

00:45:41.150 --> 00:45:42.130 problem fever.
NOTE Confidence: 0.97333664

00:45:42.430 --> 00:45:43.890 That is telling the model,
NOTE Confidence: 0.97333664

00:45:43.950 --> 00:45:44.450 okay,
NOTE Confidence: 0.9615574

00:45:44.830 --> 00:45:47.090 fever is a problem. Whenever
NOTE Confidence: 0.9615574

00:45:47.150 --> 00:45:48.430 see you see a medical
NOTE Confidence: 0.9615574

00:45:48.430 --> 00:45:50.190 problem, put it between the
NOTE Confidence: 0.9615574

00:45:50.190 --> 00:45:50.690 HTML
NOTE Confidence: 0.98120576

00:45:51.070 --> 00:45:53.035 tags, span class is equal
NOTE Confidence: 0.98120576

00:45:53.035 --> 00:45:54.795 to problem, the opening tag,
NOTE Confidence: 0.98120576

00:45:54.795 --> 00:45:56.155 and slash span, which is
NOTE Confidence: 0.98120576

00:45:56.155 --> 00:45:57.915 the closing tag. We did
NOTE Confidence: 0.98120576

00:45:57.915 --> 00:45:59.675 that for our convenience because
NOTE Confidence: 0.98120576

00:45:59.675 --> 00:46:00.795 we were comparing it with

NOTE Confidence: 0.98120576
00:46:00.795 --> 00:46:02.175 the BERT and other models.
NOTE Confidence: 0.97304255
00:46:02.715 --> 00:46:04.315 You can provide the output
NOTE Confidence: 0.97304255
00:46:04.315 --> 00:46:05.195 in the way that you
NOTE Confidence: 0.97304255
00:46:05.195 --> 00:46:06.555 want. You can use JSON
NOTE Confidence: 0.97304255
00:46:06.555 --> 00:46:07.830 format, or if you just
NOTE Confidence: 0.97304255
00:46:07.830 --> 00:46:08.870 want it to be plain
NOTE Confidence: 0.97304255
00:46:08.870 --> 00:46:10.230 text in question answering and
NOTE Confidence: 0.97304255
00:46:10.230 --> 00:46:11.670 things like that, you can
NOTE Confidence: 0.97304255
00:46:11.670 --> 00:46:12.950 give the output in such
NOTE Confidence: 0.97304255
00:46:12.950 --> 00:46:13.910 a way. But at least
NOTE Confidence: 0.97304255
00:46:13.910 --> 00:46:15.210 with the named entity recognition
NOTE Confidence: 0.97304255
00:46:15.430 --> 00:46:17.350 relation extraction, this really helps
NOTE Confidence: 0.97304255
00:46:17.350 --> 00:46:19.190 us. And another thing is
NOTE Confidence: 0.97304255
00:46:19.190 --> 00:46:21.030 that this also prevents or
NOTE Confidence: 0.97304255
00:46:21.030 --> 00:46:22.230 helps us know that the
NOTE Confidence: 0.97304255

00:46:22.230 --> 00:46:23.655 model is not hallucinating.
NOTE Confidence: 0.9684928

00:46:24.195 --> 00:46:25.155 You see? You are giving
NOTE Confidence: 0.9684928

00:46:25.155 --> 00:46:26.515 the input sentence and you
NOTE Confidence: 0.9684928

00:46:26.515 --> 00:46:28.035 are also telling the model
NOTE Confidence: 0.9684928

00:46:28.035 --> 00:46:29.415 to repeat the same sentence
NOTE Confidence: 0.9684928

00:46:29.475 --> 00:46:31.255 but with some tags attached.
NOTE Confidence: 0.9684928

00:46:31.395 --> 00:46:32.935 You can compare your input
NOTE Confidence: 0.9684928

00:46:32.995 --> 00:46:34.195 and your output to see
NOTE Confidence: 0.9684928

00:46:34.195 --> 00:46:36.210 that model is not inserting
NOTE Confidence: 0.9684928

00:46:36.350 --> 00:46:37.630 entities or things that are
NOTE Confidence: 0.9684928

00:46:37.630 --> 00:46:38.989 not already there in the
NOTE Confidence: 0.9684928

00:46:38.989 --> 00:46:40.050 original sentence.
NOTE Confidence: 0.9962768

00:46:42.190 --> 00:46:42.690 Okay.
NOTE Confidence: 0.9893497

00:46:43.310 --> 00:46:45.230 So that is how you
NOTE Confidence: 0.9893497

00:46:45.230 --> 00:46:46.430 create a prompt and do
NOTE Confidence: 0.9893497

00:46:46.430 --> 00:46:48.350 NER with large language models.

NOTE Confidence: 0.9893497
00:46:48.350 --> 00:46:49.890 The next step is relation
NOTE Confidence: 0.9893497
00:46:49.950 --> 00:46:50.450 extraction.
NOTE Confidence: 0.9543477
00:46:51.075 --> 00:46:52.594 There is a particular drug
NOTE Confidence: 0.9543477
00:46:52.594 --> 00:46:54.114 you need to associate its
NOTE Confidence: 0.9543477
00:46:54.114 --> 00:46:56.675 strength, its route, its form,
NOTE Confidence: 0.9543477
00:46:56.675 --> 00:46:57.494 its frequency,
NOTE Confidence: 0.9742312
00:46:57.875 --> 00:46:59.974 everything, and connect that particular
NOTE Confidence: 0.9742312
00:47:00.035 --> 00:47:02.114 drug to whatever is mentioned
NOTE Confidence: 0.9742312
00:47:02.114 --> 00:47:03.970 for it. Right? So for
NOTE Confidence: 0.9742312
00:47:03.970 --> 00:47:05.910 that, for the relation extraction,
NOTE Confidence: 0.9742312
00:47:06.130 --> 00:47:07.650 you need to slightly modify
NOTE Confidence: 0.9742312
00:47:07.650 --> 00:47:08.850 your prompt when you give
NOTE Confidence: 0.9742312
00:47:08.850 --> 00:47:09.890 it to. So here we
NOTE Confidence: 0.9742312
00:47:09.890 --> 00:47:11.330 are saying your task is
NOTE Confidence: 0.9742312
00:47:11.330 --> 00:47:13.430 to mark up modifier entities
NOTE Confidence: 0.9873768

00:47:13.970 --> 00:47:15.750 when given a main entity.
NOTE Confidence: 0.96735585

00:47:16.704 --> 00:47:17.905 So how do we train
NOTE Confidence: 0.96735585

00:47:17.905 --> 00:47:19.364 the model for this task?
NOTE Confidence: 0.96735585

00:47:19.585 --> 00:47:20.885 We will show the model
NOTE Confidence: 0.96735585

00:47:20.944 --> 00:47:22.065 the main entity. That is
NOTE Confidence: 0.96735585

00:47:22.065 --> 00:47:24.145 your input text. Span class
NOTE Confidence: 0.96735585

00:47:24.145 --> 00:47:25.525 drug is equal to this.
NOTE Confidence: 0.96735585

00:47:25.744 --> 00:47:26.785 Then you will ask the
NOTE Confidence: 0.96735585

00:47:26.785 --> 00:47:28.724 model, given this main entity,
NOTE Confidence: 0.96735585

00:47:28.944 --> 00:47:31.204 what are the modifier entities
NOTE Confidence: 0.96735585

00:47:31.424 --> 00:47:33.719 associated with it? And then
NOTE Confidence: 0.96735585

00:47:33.719 --> 00:47:35.160 you give examples in the
NOTE Confidence: 0.96735585

00:47:35.160 --> 00:47:36.600 output where you see now
NOTE Confidence: 0.96735585

00:47:36.600 --> 00:47:38.040 the main entity is not
NOTE Confidence: 0.96735585

00:47:38.040 --> 00:47:40.280 annotated inside the span class
NOTE Confidence: 0.96735585

00:47:40.280 --> 00:47:42.040 tags, whereas you see that

NOTE Confidence: 0.96735585
00:47:42.040 --> 00:47:43.560 point three five mg is
NOTE Confidence: 0.96735585
00:47:43.560 --> 00:47:45.400 within span class is equal
NOTE Confidence: 0.96735585
00:47:45.400 --> 00:47:46.060 to strength.
NOTE Confidence: 0.9413915
00:47:46.535 --> 00:47:48.055 So given a drug, you
NOTE Confidence: 0.9413915
00:47:48.055 --> 00:47:48.875 say this
NOTE Confidence: 0.93750817
00:47:49.415 --> 00:47:51.015 appearance is like this point
NOTE Confidence: 0.93750817
00:47:51.015 --> 00:47:52.935 five milligram or mcg, these
NOTE Confidence: 0.93750817
00:47:52.935 --> 00:47:53.895 sort of things, when it
NOTE Confidence: 0.93750817
00:47:53.895 --> 00:47:54.715 sees repeatedly,
NOTE Confidence: 0.9984606
00:47:55.094 --> 00:47:56.775 it's actually learning that this
NOTE Confidence: 0.9984606
00:47:56.775 --> 00:47:58.535 is the strength associated with
NOTE Confidence: 0.9984606
00:47:58.535 --> 00:48:00.075 that particular main entity.
NOTE Confidence: 0.9784364
00:48:00.500 --> 00:48:01.859 So a lot of examples
NOTE Confidence: 0.9784364
00:48:01.859 --> 00:48:03.299 that are annotated like this
NOTE Confidence: 0.9784364
00:48:03.299 --> 00:48:04.420 is what is helping the
NOTE Confidence: 0.9784364

00:48:04.420 --> 00:48:05.480 model learn.
NOTE Confidence: 0.9755329

00:48:07.059 --> 00:48:08.900 Again, this is another same
NOTE Confidence: 0.9755329

00:48:08.900 --> 00:48:10.420 sort of example. His blood
NOTE Confidence: 0.9755329

00:48:10.420 --> 00:48:12.420 pressure on discharge was one
NOTE Confidence: 0.9755329

00:48:12.420 --> 00:48:13.940 twenty six over sixty three.
NOTE Confidence: 0.9755329

00:48:13.940 --> 00:48:15.505 Heart rate is eighty. You
NOTE Confidence: 0.9755329

00:48:15.505 --> 00:48:16.945 cannot say blood pressure is
NOTE Confidence: 0.9755329

00:48:16.945 --> 00:48:18.225 eighty. Right? It's the same
NOTE Confidence: 0.9755329

00:48:18.225 --> 00:48:19.844 sentence which has two values
NOTE Confidence: 0.9755329

00:48:19.984 --> 00:48:21.665 and two tests. You need
NOTE Confidence: 0.9755329

00:48:21.665 --> 00:48:23.425 to correctly associate blood pressure
NOTE Confidence: 0.9755329

00:48:23.425 --> 00:48:24.705 with hundred and twenty six
NOTE Confidence: 0.9755329

00:48:24.705 --> 00:48:26.305 over sixty three and heart
NOTE Confidence: 0.9755329

00:48:26.305 --> 00:48:27.105 rate with,
NOTE Confidence: 0.998899

00:48:27.425 --> 00:48:27.925 eighty.
NOTE Confidence: 0.97454

00:48:28.630 --> 00:48:30.150 Right? So we give the

NOTE Confidence: 0.97454
00:48:30.150 --> 00:48:31.590 input when we say blood
NOTE Confidence: 0.97454
00:48:31.590 --> 00:48:33.030 pressure is the entity. Its
NOTE Confidence: 0.97454
00:48:33.030 --> 00:48:34.230 value should be hundred and
NOTE Confidence: 0.97454
00:48:34.230 --> 00:48:35.910 twenty six over eighty. If
NOTE Confidence: 0.97454
00:48:35.910 --> 00:48:37.430 we highlight heart rate as
NOTE Confidence: 0.97454
00:48:37.430 --> 00:48:38.950 the entity, then the value
NOTE Confidence: 0.97454
00:48:38.950 --> 00:48:40.010 should be eighty.
NOTE Confidence: 0.93953127
00:48:43.885 --> 00:48:46.125 Again, so now we originally
NOTE Confidence: 0.93953127
00:48:46.125 --> 00:48:47.965 had the annotated data. We
NOTE Confidence: 0.93953127
00:48:47.965 --> 00:48:50.525 converted into these instructions format
NOTE Confidence: 0.93953127
00:48:50.525 --> 00:48:51.645 that had been showing for
NOTE Confidence: 0.93953127
00:48:51.645 --> 00:48:53.325 the named entity recognition, things
NOTE Confidence: 0.93953127
00:48:53.325 --> 00:48:54.285 like this. This is an
NOTE Confidence: 0.93953127
00:48:54.285 --> 00:48:55.265 instruction demonstration
NOTE Confidence: 0.8586649
00:48:55.565 --> 00:48:56.065 sample.
NOTE Confidence: 0.9775078

00:48:56.420 --> 00:48:58.100 And for any so relation
NOTE Confidence: 0.9775078

00:48:58.100 --> 00:48:59.380 extraction, the one that you
NOTE Confidence: 0.9775078

00:48:59.380 --> 00:49:00.980 see down. So you have
NOTE Confidence: 0.9775078

00:49:00.980 --> 00:49:01.480 these
NOTE Confidence: 0.85368085

00:49:01.860 --> 00:49:03.640 and the entire dataset
NOTE Confidence: 0.9392532

00:49:03.940 --> 00:49:05.620 converted into such things is
NOTE Confidence: 0.9392532

00:49:05.620 --> 00:49:07.160 your instruction demonstration.
NOTE Confidence: 0.98938733

00:49:07.940 --> 00:49:09.220 So you have a bunch
NOTE Confidence: 0.98938733

00:49:09.220 --> 00:49:09.960 of these
NOTE Confidence: 0.97851837

00:49:10.335 --> 00:49:11.795 examples that you collectively,
NOTE Confidence: 0.9992448

00:49:13.295 --> 00:49:14.755 call as your instruction
NOTE Confidence: 0.83555615

00:49:15.135 --> 00:49:15.635 dataset.
NOTE Confidence: 0.9487137

00:49:16.015 --> 00:49:17.535 So previously, we have denoted
NOTE Confidence: 0.9487137

00:49:17.535 --> 00:49:18.895 datasets for the other models.
NOTE Confidence: 0.9487137

00:49:18.895 --> 00:49:20.094 They're just slightly different. The
NOTE Confidence: 0.9487137

00:49:20.094 --> 00:49:22.015 term is instruction datasets because

NOTE Confidence: 0.9487137
00:49:22.015 --> 00:49:23.215 the dataset is comprised of
NOTE Confidence: 0.9487137
00:49:23.215 --> 00:49:24.575 a bunch of instructions or
NOTE Confidence: 0.9487137
00:49:24.575 --> 00:49:26.570 prompts with the examples input
NOTE Confidence: 0.9487137
00:49:26.570 --> 00:49:27.550 and output examples.
NOTE Confidence: 0.95779663
00:49:29.450 --> 00:49:30.890 To instruction fine tune a
NOTE Confidence: 0.95779663
00:49:30.890 --> 00:49:32.570 large language model, all you
NOTE Confidence: 0.95779663
00:49:32.570 --> 00:49:34.110 need is such an instruction
NOTE Confidence: 0.95779663
00:49:34.250 --> 00:49:36.350 dataset specific for your task,
NOTE Confidence: 0.95779663
00:49:36.489 --> 00:49:38.605 a base large language model
NOTE Confidence: 0.95779663
00:49:38.665 --> 00:49:40.505 like llama two, llama three,
NOTE Confidence: 0.95779663
00:49:40.505 --> 00:49:41.944 llama four, or whatever it
NOTE Confidence: 0.95779663
00:49:41.944 --> 00:49:42.444 is.
NOTE Confidence: 0.97504616
00:49:42.825 --> 00:49:44.984 And then you give this
NOTE Confidence: 0.97504616
00:49:44.984 --> 00:49:46.825 model, you train it, and
NOTE Confidence: 0.97504616
00:49:46.825 --> 00:49:48.605 finally, you get an instruction
NOTE Confidence: 0.96528244

00:49:48.984 --> 00:49:51.219 tuned large language model. So
NOTE Confidence: 0.96528244

00:49:51.219 --> 00:49:52.100 if it is a LAMA
NOTE Confidence: 0.96528244

00:49:52.100 --> 00:49:53.300 model as your base, you
NOTE Confidence: 0.96528244

00:49:53.300 --> 00:49:54.739 will get an instruction tuned
NOTE Confidence: 0.96528244

00:49:54.739 --> 00:49:55.239 LAMA,
NOTE Confidence: 0.99837226

00:49:55.700 --> 00:49:56.200 but
NOTE Confidence: 0.9285842

00:49:56.980 --> 00:49:58.760 the one that is actually
NOTE Confidence: 0.9879419

00:49:59.140 --> 00:49:59.640 adapted
NOTE Confidence: 0.9253226

00:50:00.100 --> 00:50:01.239 for those tasks.
NOTE Confidence: 0.9673259

00:50:01.540 --> 00:50:02.739 So when you just take
NOTE Confidence: 0.9673259

00:50:02.739 --> 00:50:04.739 the originally available LAMA model,
NOTE Confidence: 0.9673259

00:50:04.739 --> 00:50:06.180 it's a general trained model.
NOTE Confidence: 0.9673259

00:50:06.180 --> 00:50:07.935 Right? It's not adapted or
NOTE Confidence: 0.9673259

00:50:08.095 --> 00:50:09.775 it is not domain adapted
NOTE Confidence: 0.9673259

00:50:09.775 --> 00:50:11.315 for your specific task.
NOTE Confidence: 0.99293715

00:50:11.615 --> 00:50:13.055 By fine tuning a large

NOTE Confidence: 0.99293715
00:50:13.055 --> 00:50:14.815 language model, what you are
NOTE Confidence: 0.99293715
00:50:14.815 --> 00:50:16.835 doing is making its capabilities
NOTE Confidence: 0.9916635
00:50:17.295 --> 00:50:19.635 much more lean towards whatever
NOTE Confidence: 0.9916635
00:50:19.775 --> 00:50:21.315 task you want to perform
NOTE Confidence: 0.9916635
00:50:21.609 --> 00:50:22.730 by showing it a lot
NOTE Confidence: 0.9916635
00:50:22.730 --> 00:50:24.750 of such examples and modifying
NOTE Confidence: 0.9916635
00:50:24.810 --> 00:50:26.089 its weights in such a
NOTE Confidence: 0.9916635
00:50:26.089 --> 00:50:27.930 way that it adapts to
NOTE Confidence: 0.9916635
00:50:27.930 --> 00:50:29.230 that specific task.
NOTE Confidence: 0.98872167
00:50:29.690 --> 00:50:31.609 That particular model, if you
NOTE Confidence: 0.98872167
00:50:31.609 --> 00:50:32.890 now go and test back
NOTE Confidence: 0.98872167
00:50:32.890 --> 00:50:34.585 on some general task, it
NOTE Confidence: 0.98872167
00:50:34.585 --> 00:50:36.344 might not perform the way
NOTE Confidence: 0.98872167
00:50:36.344 --> 00:50:37.325 that it previously
NOTE Confidence: 0.9997243
00:50:37.705 --> 00:50:38.205 performed
NOTE Confidence: 0.9686251

00:50:38.505 --> 00:50:39.864 because you have changed the
NOTE Confidence: 0.9686251

00:50:39.864 --> 00:50:42.025 model weights and adapted it
NOTE Confidence: 0.9686251

00:50:42.025 --> 00:50:43.645 to that specific task.
NOTE Confidence: 0.9928769

00:50:44.025 --> 00:50:45.405 Okay. So this is basically
NOTE Confidence: 0.9625918

00:50:45.705 --> 00:50:47.305 fine tuning and then you
NOTE Confidence: 0.9625918

00:50:47.305 --> 00:50:48.765 would evaluate the model.
NOTE Confidence: 0.9433341

00:50:49.460 --> 00:50:50.660 Now going back, I said
NOTE Confidence: 0.9433341

00:50:50.660 --> 00:50:51.619 we also we had the
NOTE Confidence: 0.9433341

00:50:51.619 --> 00:50:53.539 BERT based models too. There
NOTE Confidence: 0.9433341

00:50:53.539 --> 00:50:54.900 also, as you just shown,
NOTE Confidence: 0.9433341

00:50:54.900 --> 00:50:56.440 you would annotate the dataset,
NOTE Confidence: 0.9877961

00:50:56.739 --> 00:50:58.420 but for a large language
NOTE Confidence: 0.9877961

00:50:58.420 --> 00:51:00.039 model, you would give prompt.
NOTE Confidence: 0.9877961

00:51:00.260 --> 00:51:02.039 For BERT, you would convert
NOTE Confidence: 0.9877961

00:51:02.099 --> 00:51:03.319 it in such a way.
NOTE Confidence: 0.9774993

00:51:03.665 --> 00:51:05.685 So BERT model is basically

NOTE Confidence: 0.9774993
00:51:05.744 --> 00:51:07.765 sequence tagging. So you convert
NOTE Confidence: 0.97231877
00:51:08.145 --> 00:51:09.425 each of the sentence into
NOTE Confidence: 0.97231877
00:51:09.425 --> 00:51:11.105 tokens. Let's say a token
NOTE Confidence: 0.97231877
00:51:11.105 --> 00:51:12.625 is a word. So vital
NOTE Confidence: 0.97231877
00:51:12.625 --> 00:51:14.724 sign remains stable. And then
NOTE Confidence: 0.97231877
00:51:14.945 --> 00:51:16.410 doctor Xu has covered this
NOTE Confidence: 0.97231877
00:51:16.890 --> 00:51:18.569 BIO tagging is what we
NOTE Confidence: 0.97231877
00:51:18.569 --> 00:51:20.190 call beginning of an entity,
NOTE Confidence: 0.97231877
00:51:20.329 --> 00:51:22.170 inside of an entity, outside
NOTE Confidence: 0.97231877
00:51:22.170 --> 00:51:23.849 of an entity. So vital
NOTE Confidence: 0.97231877
00:51:23.849 --> 00:51:25.130 sign is a test here,
NOTE Confidence: 0.97231877
00:51:25.130 --> 00:51:26.750 so you say b test.
NOTE Confidence: 0.97231877
00:51:27.049 --> 00:51:28.250 If you have a problem,
NOTE Confidence: 0.97231877
00:51:28.250 --> 00:51:30.109 you would say acute carcinoma
NOTE Confidence: 0.97231877
00:51:30.329 --> 00:51:31.455 or something. I'm making this
NOTE Confidence: 0.97231877

00:51:31.455 --> 00:51:32.975 up. So acute is gonna
NOTE Confidence: 0.97231877

00:51:32.975 --> 00:51:33.955 be b problem
NOTE Confidence: 0.94583786

00:51:34.255 --> 00:51:35.775 and carcinoma is gonna be
NOTE Confidence: 0.94583786

00:51:35.775 --> 00:51:37.135 I problem. If it is
NOTE Confidence: 0.94583786

00:51:37.135 --> 00:51:38.575 not within the four main
NOTE Confidence: 0.94583786

00:51:38.575 --> 00:51:40.655 entities and four modifiers that
NOTE Confidence: 0.94583786

00:51:40.655 --> 00:51:42.175 we have, we tag it
NOTE Confidence: 0.94583786

00:51:42.175 --> 00:51:44.175 as o, which means outside
NOTE Confidence: 0.94583786

00:51:44.175 --> 00:51:45.760 of an entity. So take
NOTE Confidence: 0.94583786

00:51:45.760 --> 00:51:47.760 the same annotated dataset, convert
NOTE Confidence: 0.94583786

00:51:47.760 --> 00:51:49.920 into two different formats. One,
NOTE Confidence: 0.94583786

00:51:49.920 --> 00:51:51.360 safe for the llama, another
NOTE Confidence: 0.94583786

00:51:51.360 --> 00:51:52.180 for the bird.
NOTE Confidence: 0.96982765

00:51:52.480 --> 00:51:53.760 And for bird, this is
NOTE Confidence: 0.96982765

00:51:53.760 --> 00:51:55.600 token classification. So given a
NOTE Confidence: 0.96982765

00:51:55.600 --> 00:51:57.360 content given a sentence, you

NOTE Confidence: 0.96982765
00:51:57.360 --> 00:51:59.200 are basically predicting whether vital
NOTE Confidence: 0.96982765
00:51:59.200 --> 00:52:00.974 is among the b test,
NOTE Confidence: 0.96982765
00:52:00.974 --> 00:52:02.414 I test, b problem, I
NOTE Confidence: 0.96982765
00:52:02.414 --> 00:52:03.934 for problem, b value, I
NOTE Confidence: 0.96982765
00:52:03.934 --> 00:52:05.714 value, whatever is the corresponding
NOTE Confidence: 0.9672559
00:52:06.015 --> 00:52:07.454 label that should be for
NOTE Confidence: 0.9672559
00:52:07.454 --> 00:52:09.055 that particular token. So it
NOTE Confidence: 0.9672559
00:52:09.055 --> 00:52:11.295 is token classification task what
NOTE Confidence: 0.9672559
00:52:11.295 --> 00:52:11.954 we do.
NOTE Confidence: 0.96572065
00:52:13.100 --> 00:52:15.200 How does relation extraction correspond
NOTE Confidence: 0.96572065
00:52:15.340 --> 00:52:16.640 in the case of BERT?
NOTE Confidence: 0.96572065
00:52:16.940 --> 00:52:18.380 So here also the same
NOTE Confidence: 0.96572065
00:52:18.380 --> 00:52:19.980 thing. Now it becomes sentence
NOTE Confidence: 0.96572065
00:52:19.980 --> 00:52:20.480 classification.
NOTE Confidence: 0.85382193
00:52:20.860 --> 00:52:22.620 You have two classes. One
NOTE Confidence: 0.85382193

00:52:22.620 --> 00:52:24.480 that has value, which is
NOTE Confidence: 0.98473275

00:52:24.825 --> 00:52:26.364 which is a positive class
NOTE Confidence: 0.98473275

00:52:26.425 --> 00:52:27.385 and the other, it is
NOTE Confidence: 0.98473275

00:52:27.385 --> 00:52:29.065 a negative class. So if
NOTE Confidence: 0.98473275

00:52:29.065 --> 00:52:31.065 you show blood pressure and
NOTE Confidence: 0.98473275

00:52:31.065 --> 00:52:32.665 eighty, that is a negative
NOTE Confidence: 0.98473275

00:52:32.665 --> 00:52:34.185 sample. You should say label
NOTE Confidence: 0.98473275

00:52:34.185 --> 00:52:36.025 that sentence as negative. If
NOTE Confidence: 0.98473275

00:52:36.025 --> 00:52:37.464 you have blood pressure and
NOTE Confidence: 0.98473275

00:52:37.464 --> 00:52:38.984 hundred and twenty six over
NOTE Confidence: 0.98473275

00:52:38.984 --> 00:52:40.344 sixty three, then it is
NOTE Confidence: 0.98473275

00:52:40.344 --> 00:52:42.480 a positive sample. So it
NOTE Confidence: 0.98473275

00:52:42.480 --> 00:52:44.020 becomes a sentence classification
NOTE Confidence: 0.9550411

00:52:44.320 --> 00:52:46.080 task. And so many patterns
NOTE Confidence: 0.9550411

00:52:46.080 --> 00:52:47.460 like this and seeing repeated
NOTE Confidence: 0.9550411

00:52:47.600 --> 00:52:49.280 sentences like that, the model

NOTE Confidence: 0.9550411
00:52:49.280 --> 00:52:51.200 is learning that particular pattern
NOTE Confidence: 0.9550411
00:52:51.200 --> 00:52:52.960 and identifying it. Another time
NOTE Confidence: 0.9550411
00:52:52.960 --> 00:52:54.320 that sort of sentence appears,
NOTE Confidence: 0.9550411
00:52:54.320 --> 00:52:55.780 okay. This is a positive
NOTE Confidence: 0.9550411
00:52:56.015 --> 00:52:57.215 or has value or this
NOTE Confidence: 0.9550411
00:52:57.215 --> 00:52:58.094 is a negative or a
NOTE Confidence: 0.9550411
00:52:58.094 --> 00:52:59.235 negative class there.
NOTE Confidence: 0.9959546
00:53:01.055 --> 00:53:02.975 This is the entire Kiwi
NOTE Confidence: 0.9959546
00:53:02.975 --> 00:53:03.475 pipeline.
NOTE Confidence: 0.99472463
00:53:04.255 --> 00:53:04.755 Okay.
NOTE Confidence: 0.9955434
00:53:06.335 --> 00:53:06.835 Having
NOTE Confidence: 0.9997563
00:53:07.135 --> 00:53:07.635 data
NOTE Confidence: 0.98391664
00:53:07.935 --> 00:53:09.955 from multiple sources is important.
NOTE Confidence: 0.98650616
00:53:10.539 --> 00:53:12.140 Something that works on your
NOTE Confidence: 0.98650616
00:53:12.140 --> 00:53:13.760 specific data at a particular
NOTE Confidence: 0.98650616

00:53:13.819 --> 00:53:15.500 hospital setting written by one
NOTE Confidence: 0.98650616

00:53:15.500 --> 00:53:17.339 specific doctor in one particular
NOTE Confidence: 0.98650616

00:53:17.339 --> 00:53:19.920 setting might not generalize well
NOTE Confidence: 0.98650616

00:53:19.980 --> 00:53:21.339 when you try to use
NOTE Confidence: 0.98650616

00:53:21.339 --> 00:53:23.260 that same pipeline in another
NOTE Confidence: 0.98650616

00:53:23.260 --> 00:53:23.760 hospital,
NOTE Confidence: 0.9486602

00:53:24.255 --> 00:53:25.775 in another node that is
NOTE Confidence: 0.9486602

00:53:25.775 --> 00:53:27.454 written by another physician or
NOTE Confidence: 0.9486602

00:53:27.454 --> 00:53:28.515 health care provider.
NOTE Confidence: 0.9951988

00:53:29.135 --> 00:53:30.974 So for Kiwi, we actually
NOTE Confidence: 0.9951988

00:53:30.974 --> 00:53:32.914 have data from four sources
NOTE Confidence: 0.9951988

00:53:32.974 --> 00:53:34.194 so that we can
NOTE Confidence: 0.9996481

00:53:34.494 --> 00:53:36.015 make the model much more
NOTE Confidence: 0.9996481

00:53:36.015 --> 00:53:36.515 generalizable
NOTE Confidence: 0.91792935

00:53:37.280 --> 00:53:38.400 and make it see the
NOTE Confidence: 0.91792935

00:53:38.400 --> 00:53:39.760 patterns that happens in a

NOTE Confidence: 0.91792935

00:53:39.760 --> 00:53:41.520 wide variety of data. We

NOTE Confidence: 0.91792935

00:53:41.520 --> 00:53:42.719 have the UTP that is

NOTE Confidence: 0.91792935

00:53:42.719 --> 00:53:44.560 UT Physicians, empty samples that

NOTE Confidence: 0.91792935

00:53:44.560 --> 00:53:46.820 is an, publicly available dataset.

NOTE Confidence: 0.91792935

00:53:46.880 --> 00:53:48.339 MIMIC three, you might know.

NOTE Confidence: 0.91792935

00:53:48.480 --> 00:53:50.154 And all these data from

NOTE Confidence: 0.91792935

00:53:50.154 --> 00:53:52.255 these different sources are incorporated

NOTE Confidence: 0.91792935

00:53:52.395 --> 00:53:53.694 for our training process.

NOTE Confidence: 0.9593701

00:53:54.154 --> 00:53:55.855 And as I mentioned, instruction

NOTE Confidence: 0.9593701

00:53:55.914 --> 00:53:58.174 format for llama BERT format

NOTE Confidence: 0.9593701

00:53:58.234 --> 00:53:59.934 for training the BERT models.

NOTE Confidence: 0.9593701

00:54:00.154 --> 00:54:01.755 Then you fine tune both

NOTE Confidence: 0.9593701

00:54:01.755 --> 00:54:03.355 the models, and then you

NOTE Confidence: 0.9593701

00:54:03.355 --> 00:54:05.250 test the models out. So

NOTE Confidence: 0.9593701

00:54:05.250 --> 00:54:06.290 you test it on a

NOTE Confidence: 0.9593701

00:54:06.290 --> 00:54:07.969 subset of the UTP, empty
NOTE Confidence: 0.9593701

00:54:07.969 --> 00:54:09.570 samples, and mimic three and
NOTE Confidence: 0.9593701

00:54:09.570 --> 00:54:10.690 also on I two b
NOTE Confidence: 0.9593701

00:54:10.690 --> 00:54:12.050 two. Again, doctor Xu mentioned
NOTE Confidence: 0.9593701

00:54:12.050 --> 00:54:12.930 that I two b two
NOTE Confidence: 0.9593701

00:54:12.930 --> 00:54:14.690 is unseen data. Right? It's
NOTE Confidence: 0.9593701

00:54:14.690 --> 00:54:16.050 not in your training data.
NOTE Confidence: 0.9593701

00:54:16.050 --> 00:54:17.090 That's how we are testing
NOTE Confidence: 0.9593701

00:54:17.090 --> 00:54:17.750 the generalizability
NOTE Confidence: 0.99931574

00:54:18.290 --> 00:54:19.330 to see whether it is
NOTE Confidence: 0.99931574

00:54:19.330 --> 00:54:21.884 actually performing on unseen data.
NOTE Confidence: 0.9390398

00:54:22.505 --> 00:54:23.325 Post process
NOTE Confidence: 0.75591785

00:54:23.625 --> 00:54:25.244 separate entities relationship
NOTE Confidence: 0.89431804

00:54:25.545 --> 00:54:27.464 calculate precision recall and f
NOTE Confidence: 0.89431804

00:54:27.464 --> 00:54:28.664 one, and that is your
NOTE Confidence: 0.89431804

00:54:28.664 --> 00:54:29.164 evaluation.

NOTE Confidence: 0.99571294
00:54:31.065 --> 00:54:32.025 So this is,
NOTE Confidence: 0.9680526
00:54:32.344 --> 00:54:34.025 quickly the composition of the
NOTE Confidence: 0.9680526
00:54:34.025 --> 00:54:36.150 Kiwi dataset, means the Kiwi
NOTE Confidence: 0.9680526
00:54:36.150 --> 00:54:37.590 model that we are giving
NOTE Confidence: 0.9680526
00:54:37.590 --> 00:54:39.270 out currently. It has it
NOTE Confidence: 0.9680526
00:54:39.270 --> 00:54:40.630 has been trained on about
NOTE Confidence: 0.9680526
00:54:40.630 --> 00:54:42.570 one thousand four hundred documents
NOTE Confidence: 0.9610982
00:54:42.950 --> 00:54:44.710 and then tested on some,
NOTE Confidence: 0.9776361
00:54:45.349 --> 00:54:46.950 four different types, each having
NOTE Confidence: 0.9776361
00:54:46.950 --> 00:54:48.230 fifty documents or,
NOTE Confidence: 0.99965477
00:54:48.550 --> 00:54:49.609 twenty five documents.
NOTE Confidence: 0.9616748
00:54:52.015 --> 00:54:53.955 So evaluation, I mentioned precision
NOTE Confidence: 0.9616748
00:54:54.015 --> 00:54:55.375 recall and f one, and
NOTE Confidence: 0.9616748
00:54:55.375 --> 00:54:56.895 it is exact match and
NOTE Confidence: 0.9616748
00:54:56.895 --> 00:54:58.495 relaxed match. To be clear,
NOTE Confidence: 0.9616748

00:54:58.495 --> 00:55:00.495 exact match, the entity type
NOTE Confidence: 0.9616748

00:55:00.495 --> 00:55:02.255 should match and the boundary
NOTE Confidence: 0.9616748

00:55:02.255 --> 00:55:03.910 should also match. But when
NOTE Confidence: 0.9616748

00:55:03.910 --> 00:55:05.430 it is relaxed match, the
NOTE Confidence: 0.9616748

00:55:05.430 --> 00:55:07.270 entity type should still match,
NOTE Confidence: 0.9616748

00:55:07.270 --> 00:55:08.710 but the boundary can be
NOTE Confidence: 0.9616748

00:55:08.710 --> 00:55:09.210 overlapping.
NOTE Confidence: 0.96118855

00:55:12.390 --> 00:55:13.910 How did we perform? Llama
NOTE Confidence: 0.96118855

00:55:13.910 --> 00:55:15.670 three seventy billion was sort
NOTE Confidence: 0.96118855

00:55:15.670 --> 00:55:18.010 of better for NER task
NOTE Confidence: 0.96118855

00:55:18.155 --> 00:55:19.855 and, again, for relation extraction,
NOTE Confidence: 0.96118855

00:55:19.915 --> 00:55:21.114 but you also see that
NOTE Confidence: 0.96118855

00:55:21.114 --> 00:55:23.055 some smaller models still performed
NOTE Confidence: 0.9508643

00:55:23.675 --> 00:55:25.195 on par with it. Sometimes
NOTE Confidence: 0.9508643

00:55:25.195 --> 00:55:26.315 you do not have much
NOTE Confidence: 0.9508643

00:55:26.315 --> 00:55:27.835 difference with the BERT model,

NOTE Confidence: 0.9508643
00:55:27.835 --> 00:55:29.275 but here we saw that
NOTE Confidence: 0.9508643
00:55:29.275 --> 00:55:31.214 at least some statistical significance
NOTE Confidence: 0.9508643
00:55:31.435 --> 00:55:32.920 was there. And I two
NOTE Confidence: 0.9508643
00:55:32.920 --> 00:55:33.960 b two is the unseen
NOTE Confidence: 0.9508643
00:55:33.960 --> 00:55:35.560 data, and doctor Xu mentioned
NOTE Confidence: 0.9508643
00:55:35.560 --> 00:55:36.780 again about how,
NOTE Confidence: 0.9737618
00:55:37.160 --> 00:55:39.080 you know, last language models
NOTE Confidence: 0.9737618
00:55:39.080 --> 00:55:40.600 are better on unseen data
NOTE Confidence: 0.9737618
00:55:40.600 --> 00:55:42.200 compared to the BERT. And
NOTE Confidence: 0.9737618
00:55:42.200 --> 00:55:44.120 BERT models, again, definitely need
NOTE Confidence: 0.9737618
00:55:44.120 --> 00:55:45.239 a lot more data to
NOTE Confidence: 0.9737618
00:55:45.239 --> 00:55:46.060 train on.
NOTE Confidence: 0.91960067
00:55:48.575 --> 00:55:50.255 Now what about the memory
NOTE Confidence: 0.91960067
00:55:50.255 --> 00:55:52.515 usage, total GPU hours GPU
NOTE Confidence: 0.91960067
00:55:52.575 --> 00:55:54.674 hours per epoch, energy consumption,
NOTE Confidence: 0.9848243

00:55:55.135 --> 00:55:56.035 carbon emission?
NOTE Confidence: 0.97025067

00:55:56.414 --> 00:55:57.934 That is where a lot
NOTE Confidence: 0.97025067

00:55:57.934 --> 00:55:59.775 of these computational resources and
NOTE Confidence: 0.97025067

00:55:59.775 --> 00:56:01.454 things comes into play. You
NOTE Confidence: 0.97025067

00:56:01.454 --> 00:56:03.230 need huge amount of memory.
NOTE Confidence: 0.97025067

00:56:03.230 --> 00:56:04.109 As you know, we are
NOTE Confidence: 0.97025067

00:56:04.109 --> 00:56:05.469 comparing a BART model that
NOTE Confidence: 0.97025067

00:56:05.469 --> 00:56:06.910 is about hundred million, three
NOTE Confidence: 0.97025067

00:56:06.910 --> 00:56:08.829 hundred million parameters to something
NOTE Confidence: 0.97025067

00:56:08.829 --> 00:56:10.270 that is seven billion, eight
NOTE Confidence: 0.97025067

00:56:10.270 --> 00:56:12.109 billion, seventy billion, and that
NOTE Confidence: 0.97025067

00:56:12.109 --> 00:56:12.609 difference
NOTE Confidence: 0.9814347

00:56:12.910 --> 00:56:13.630 really shows,
NOTE Confidence: 0.98489046

00:56:14.635 --> 00:56:15.995 in the amount of compute
NOTE Confidence: 0.98489046

00:56:15.995 --> 00:56:17.355 and the, hours that you
NOTE Confidence: 0.98489046

00:56:17.355 --> 00:56:19.195 require for training these models

NOTE Confidence: 0.98489046
00:56:19.195 --> 00:56:20.635 and the memory that they
NOTE Confidence: 0.98489046
00:56:20.635 --> 00:56:21.135 utilize.
NOTE Confidence: 0.9535904
00:56:22.555 --> 00:56:24.315 So, if you want to,
NOTE Confidence: 0.9535904
00:56:24.555 --> 00:56:25.755 fine tune the model, if
NOTE Confidence: 0.9535904
00:56:25.755 --> 00:56:27.215 you are using this parameter
NOTE Confidence: 0.9535904
00:56:27.275 --> 00:56:29.114 efficient fine tuning approaches like
NOTE Confidence: 0.9535904
00:56:29.114 --> 00:56:31.150 Laura, Melingfei is gonna discuss
NOTE Confidence: 0.9535904
00:56:31.150 --> 00:56:31.810 on that,
NOTE Confidence: 0.99255174
00:56:32.110 --> 00:56:32.910 then it is,
NOTE Confidence: 0.99476606
00:56:33.550 --> 00:56:34.770 you need one,
NOTE Confidence: 0.9787066
00:56:35.230 --> 00:56:36.910 a one hundred eighty gigabyte
NOTE Confidence: 0.9787066
00:56:36.910 --> 00:56:38.690 GPU. But if you,
NOTE Confidence: 0.983032
00:56:39.390 --> 00:56:40.830 need to do the inference
NOTE Confidence: 0.983032
00:56:40.830 --> 00:56:42.350 for the seventy billion model,
NOTE Confidence: 0.983032
00:56:42.350 --> 00:56:43.695 you need two a one
NOTE Confidence: 0.983032

00:56:43.695 --> 00:56:44.175 hundred,
NOTE Confidence: 0.99291307
00:56:44.495 --> 00:56:45.795 eighty gigabyte GPU.
NOTE Confidence: 0.9875859
00:56:46.815 --> 00:56:48.175 Okay. Again, our paper has
NOTE Confidence: 0.9875859
00:56:48.175 --> 00:56:48.975 a lot of things. I
NOTE Confidence: 0.9875859
00:56:48.975 --> 00:56:50.335 can skip through this. Just
NOTE Confidence: 0.9875859
00:56:50.335 --> 00:56:51.614 want to talk about concept
NOTE Confidence: 0.9875859
00:56:51.614 --> 00:56:52.114 normalization.
NOTE Confidence: 0.99825424
00:56:52.735 --> 00:56:53.855 The actual way we do
NOTE Confidence: 0.99825424
00:56:53.855 --> 00:56:54.755 concept normalization
NOTE Confidence: 0.9832494
00:56:55.135 --> 00:56:57.215 is actually having elastic search,
NOTE Confidence: 0.9832494
00:56:57.215 --> 00:56:59.060 which basically does exact
NOTE Confidence: 0.9878063
00:56:59.520 --> 00:57:00.100 match and,
NOTE Confidence: 0.92170346
00:57:00.400 --> 00:57:02.000 partial match of a z
NOTE Confidence: 0.92170346
00:57:02.000 --> 00:57:03.440 match and then b m
NOTE Confidence: 0.92170346
00:57:03.440 --> 00:57:04.820 twenty five to rerank
NOTE Confidence: 0.99902713
00:57:05.120 --> 00:57:05.620 those

NOTE Confidence: 0.9221044
00:57:06.080 --> 00:57:06.820 things extracted.
NOTE Confidence: 0.9565772
00:57:07.280 --> 00:57:08.640 So here we in the
NOTE Confidence: 0.9565772
00:57:08.640 --> 00:57:10.000 Kiwi, we have mapped it
NOTE Confidence: 0.9565772
00:57:10.000 --> 00:57:12.340 to UMLS concept unique identifiers.
NOTE Confidence: 0.9717762
00:57:12.735 --> 00:57:14.175 For anyone who's not familiar
NOTE Confidence: 0.9717762
00:57:14.175 --> 00:57:14.915 with UMLS,
NOTE Confidence: 0.90430826
00:57:15.295 --> 00:57:16.675 it is a meta UMLS
NOTE Confidence: 0.90430826
00:57:16.735 --> 00:57:17.235 metatasaurus
NOTE Confidence: 0.87763715
00:57:17.535 --> 00:57:18.675 basically incorporates
NOTE Confidence: 0.98224324
00:57:19.135 --> 00:57:21.235 about hundred and, some vocabularies
NOTE Confidence: 0.9950762
00:57:21.855 --> 00:57:23.135 and gives it a unique
NOTE Confidence: 0.9950762
00:57:23.135 --> 00:57:25.295 identity. Same concepts from all
NOTE Confidence: 0.9950762
00:57:25.295 --> 00:57:26.275 different vocabularies
NOTE Confidence: 0.96936905
00:57:26.655 --> 00:57:27.935 are mapped to a unique
NOTE Confidence: 0.96936905
00:57:27.935 --> 00:57:29.810 concept ID. So here,
NOTE Confidence: 0.9951137

00:57:30.190 --> 00:57:31.390 this is a large language
NOTE Confidence: 0.9951137

00:57:31.390 --> 00:57:33.330 model utilized concept normalization
NOTE Confidence: 0.9991157

00:57:33.630 --> 00:57:34.130 pipeline.
NOTE Confidence: 0.9771481

00:57:34.590 --> 00:57:35.970 Once you do the NER,
NOTE Confidence: 0.9771481

00:57:36.030 --> 00:57:37.470 you get that query. On
NOTE Confidence: 0.9771481

00:57:37.470 --> 00:57:38.670 the left side, you see
NOTE Confidence: 0.9771481

00:57:38.670 --> 00:57:40.590 left atrium dilated. So that
NOTE Confidence: 0.9771481

00:57:40.590 --> 00:57:42.515 is your query entity with
NOTE Confidence: 0.9771481

00:57:42.515 --> 00:57:43.875 its context. Let's say the
NOTE Confidence: 0.9771481

00:57:43.875 --> 00:57:45.015 sentence that has
NOTE Confidence: 0.9605486

00:57:45.315 --> 00:57:46.355 that. That you give to
NOTE Confidence: 0.9605486

00:57:46.355 --> 00:57:47.955 a last language model and
NOTE Confidence: 0.9605486

00:57:47.955 --> 00:57:49.575 ask it to generate multiple
NOTE Confidence: 0.9605486

00:57:49.635 --> 00:57:50.755 synonyms of it. Why are
NOTE Confidence: 0.9605486

00:57:50.755 --> 00:57:51.955 we doing that? Because the
NOTE Confidence: 0.9605486

00:57:51.955 --> 00:57:53.235 exact thing might not be

NOTE Confidence: 0.9605486
00:57:53.235 --> 00:57:54.435 appearing in any of the
NOTE Confidence: 0.9605486
00:57:54.435 --> 00:57:55.970 standardized vocabularies. So
NOTE Confidence: 0.6857144
00:57:56.350 --> 00:57:56.850 we
NOTE Confidence: 0.8022595
00:57:57.230 --> 00:57:58.990 generate as many variations of
NOTE Confidence: 0.8022595
00:57:58.990 --> 00:58:00.270 that particular entity so that
NOTE Confidence: 0.8022595
00:58:00.270 --> 00:58:01.410 we can do and
NOTE Confidence: 0.69586235
00:58:01.710 --> 00:58:02.210 match,
NOTE Confidence: 0.91109043
00:58:02.670 --> 00:58:04.109 that elastic search and b
NOTE Confidence: 0.91109043
00:58:04.109 --> 00:58:05.470 m twenty five actually does
NOTE Confidence: 0.91109043
00:58:05.470 --> 00:58:06.349 that. So you give the
NOTE Confidence: 0.91109043
00:58:06.349 --> 00:58:08.109 original utterance and all the
NOTE Confidence: 0.91109043
00:58:08.109 --> 00:58:09.890 synonyms and actually check-in
NOTE Confidence: 0.98868716
00:58:10.190 --> 00:58:10.590 your,
NOTE Confidence: 0.9954951
00:58:11.464 --> 00:58:12.085 you know,
NOTE Confidence: 0.9401489
00:58:12.625 --> 00:58:14.224 database that you have created
NOTE Confidence: 0.9401489

00:58:14.224 --> 00:58:15.984 whether that entity is actually
NOTE Confidence: 0.9401489

00:58:15.984 --> 00:58:16.885 present there.
NOTE Confidence: 0.9623802

00:58:17.744 --> 00:58:18.244 So,
NOTE Confidence: 0.9540374

00:58:18.785 --> 00:58:19.744 you search, you will get
NOTE Confidence: 0.9540374

00:58:19.744 --> 00:58:20.865 a bunch of concepts that
NOTE Confidence: 0.9540374

00:58:20.865 --> 00:58:21.905 are sort of similar to
NOTE Confidence: 0.9540374

00:58:21.905 --> 00:58:23.184 that, and then you again
NOTE Confidence: 0.9540374

00:58:23.184 --> 00:58:24.724 use a large language model
NOTE Confidence: 0.9540374

00:58:24.890 --> 00:58:26.810 to find among those concept
NOTE Confidence: 0.9540374

00:58:26.810 --> 00:58:28.410 which is the best one
NOTE Confidence: 0.9540374

00:58:28.410 --> 00:58:29.150 that actually
NOTE Confidence: 0.77716625

00:58:29.610 --> 00:58:32.110 represents the originally redeemed entity.
NOTE Confidence: 0.992579

00:58:32.890 --> 00:58:34.250 I know I'm going so
NOTE Confidence: 0.992579

00:58:34.250 --> 00:58:35.150 fast, but
NOTE Confidence: 0.96261364

00:58:35.610 --> 00:58:37.130 the slides will be available,
NOTE Confidence: 0.96261364

00:58:37.130 --> 00:58:38.170 and we will also think

NOTE Confidence: 0.96261364
00:58:38.170 --> 00:58:39.850 of making the recordings available
NOTE Confidence: 0.96261364
00:58:39.850 --> 00:58:40.785 on the YBIG website.
NOTE Confidence: 0.9432206
00:58:41.744 --> 00:58:43.125 Okay. Last step,
NOTE Confidence: 0.98525965
00:58:43.825 --> 00:58:45.265 the Kiwi usually gives you
NOTE Confidence: 0.98525965
00:58:45.265 --> 00:58:46.704 output in a JSON format,
NOTE Confidence: 0.98525965
00:58:46.704 --> 00:58:48.145 but we also have scripts
NOTE Confidence: 0.98525965
00:58:48.145 --> 00:58:49.105 to make it easy for
NOTE Confidence: 0.98525965
00:58:49.105 --> 00:58:50.305 you so that that JSON
NOTE Confidence: 0.98525965
00:58:50.305 --> 00:58:51.505 can be converted into a
NOTE Confidence: 0.98525965
00:58:51.505 --> 00:58:52.005 CSV.
NOTE Confidence: 0.8962111
00:58:52.305 --> 00:58:53.585 And what I have actually
NOTE Confidence: 0.8962111
00:58:53.585 --> 00:58:55.025 highlighted is you see in
NOTE Confidence: 0.8962111
00:58:55.025 --> 00:58:55.560 the first,
NOTE Confidence: 0.97613305
00:58:56.440 --> 00:58:58.040 column, it's the entity, the
NOTE Confidence: 0.97613305
00:58:58.040 --> 00:58:59.240 term that we have actually
NOTE Confidence: 0.97613305

00:58:59.240 --> 00:59:00.920 extracted, and the highlighted one
NOTE Confidence: 0.97613305

00:59:00.920 --> 00:59:02.280 is the concept ID for
NOTE Confidence: 0.97613305

00:59:02.280 --> 00:59:03.420 that, which is basically
NOTE Confidence: 0.95581996

00:59:03.880 --> 00:59:06.220 quiz or concept unique identifiers
NOTE Confidence: 0.9951272

00:59:06.600 --> 00:59:08.280 of that particular thing from
NOTE Confidence: 0.9951272

00:59:08.280 --> 00:59:08.780 UMLS.
NOTE Confidence: 0.97305673

00:59:09.215 --> 00:59:10.175 And if you ask why
NOTE Confidence: 0.97305673

00:59:10.175 --> 00:59:10.675 UMLS,
NOTE Confidence: 0.9540199

00:59:11.055 --> 00:59:12.175 if there is a concept
NOTE Confidence: 0.9540199

00:59:12.175 --> 00:59:14.015 unique identifier, you can actually
NOTE Confidence: 0.9540199

00:59:14.015 --> 00:59:15.295 map it back to SNOMED,
NOTE Confidence: 0.9540199

00:59:15.295 --> 00:59:17.695 ICD, mesh because UMLS includes
NOTE Confidence: 0.9540199

00:59:17.695 --> 00:59:19.135 all those things. That's a
NOTE Confidence: 0.9540199

00:59:19.135 --> 00:59:20.275 very easy task.
NOTE Confidence: 0.9537706

00:59:20.735 --> 00:59:22.015 Where can you find Kiwi?
NOTE Confidence: 0.9537706

00:59:22.015 --> 00:59:23.250 This is our website, You

NOTE Confidence: 0.9537706
00:59:23.250 --> 00:59:25.090 know? Kiwi dot clinical n
NOTE Confidence: 0.9537706
00:59:25.090 --> 00:59:26.310 l p dot org.
NOTE Confidence: 0.97680146
00:59:26.850 --> 00:59:28.050 The QR code will take
NOTE Confidence: 0.97680146
00:59:28.050 --> 00:59:29.830 you right there. You press
NOTE Confidence: 0.97680146
00:59:29.890 --> 00:59:31.110 the live demo,
NOTE Confidence: 0.99875045
00:59:31.490 --> 00:59:32.150 and then
NOTE Confidence: 0.93590766
00:59:32.610 --> 00:59:34.790 you, get a prepopulated,
NOTE Confidence: 0.8551228
00:59:35.890 --> 00:59:37.635 note, a few sentences.
NOTE Confidence: 0.99720865
00:59:38.335 --> 00:59:39.935 Click submit. It will show
NOTE Confidence: 0.99720865
00:59:39.935 --> 00:59:41.375 you the entities and the
NOTE Confidence: 0.99720865
00:59:41.375 --> 00:59:41.875 relation
NOTE Confidence: 0.9784693
00:59:42.335 --> 00:59:43.695 extracted. You can remove that
NOTE Confidence: 0.9784693
00:59:43.695 --> 00:59:45.215 text, add your own text.
NOTE Confidence: 0.9784693
00:59:45.215 --> 00:59:47.055 No programming experience. You can
NOTE Confidence: 0.9784693
00:59:47.055 --> 00:59:48.415 put something in there and
NOTE Confidence: 0.9784693

00:59:48.415 --> 00:59:49.455 get to see what are
NOTE Confidence: 0.9784693

00:59:49.455 --> 00:59:51.295 the entities. Just play around
NOTE Confidence: 0.9784693

00:59:51.295 --> 00:59:51.960 with that.
NOTE Confidence: 0.97033405

00:59:52.440 --> 00:59:54.119 Okay. Now if you want
NOTE Confidence: 0.97033405

00:59:54.119 --> 00:59:55.640 to download the models, we
NOTE Confidence: 0.97033405

00:59:55.640 --> 00:59:57.080 have this another page called
NOTE Confidence: 0.97033405

00:59:57.080 --> 00:59:58.359 download. You need to fill
NOTE Confidence: 0.97033405

00:59:58.359 --> 00:59:59.099 a form,
NOTE Confidence: 0.996563

00:59:59.800 --> 01:00:00.839 and then we will send
NOTE Confidence: 0.996563

01:00:00.839 --> 01:00:01.339 you,
NOTE Confidence: 0.93134636

01:00:01.800 --> 01:00:03.320 the Docker images for that.
NOTE Confidence: 0.93134636

01:00:03.320 --> 01:00:04.380 Now how is Docker,
NOTE Confidence: 0.9967054

01:00:04.839 --> 01:00:05.339 different?
NOTE Confidence: 0.9931013

01:00:05.775 --> 01:00:06.994 Everything is prepackaged
NOTE Confidence: 0.99521136

01:00:07.295 --> 01:00:08.815 into a container. You do
NOTE Confidence: 0.99521136

01:00:08.815 --> 01:00:10.414 not need to install things

NOTE Confidence: 0.99521136
01:00:10.414 --> 01:00:12.015 separately. You just and the
NOTE Confidence: 0.99521136
01:00:12.015 --> 01:00:13.474 Docker comes with instructions
NOTE Confidence: 0.9749034
01:00:13.775 --> 01:00:14.895 as to what to do.
NOTE Confidence: 0.9749034
01:00:14.895 --> 01:00:16.355 It's just like an executable
NOTE Confidence: 0.9749034
01:00:16.575 --> 01:00:17.315 you run,
NOTE Confidence: 0.9376887
01:00:17.855 --> 01:00:19.375 select, okay, one, two, three
NOTE Confidence: 0.9376887
01:00:19.375 --> 01:00:20.740 numbers. It also has a
NOTE Confidence: 0.9376887
01:00:20.740 --> 01:00:22.280 readme file, which gives
NOTE Confidence: 0.9303634
01:00:22.660 --> 01:00:23.700 you the ways how to
NOTE Confidence: 0.9303634
01:00:23.700 --> 01:00:25.220 run about it, and then
NOTE Confidence: 0.9303634
01:00:25.220 --> 01:00:26.500 the output will be stored
NOTE Confidence: 0.9303634
01:00:26.500 --> 01:00:27.460 in the you need to
NOTE Confidence: 0.9303634
01:00:27.460 --> 01:00:28.980 give this, where your input
NOTE Confidence: 0.9303634
01:00:28.980 --> 01:00:30.180 data is and the output,
NOTE Confidence: 0.9303634
01:00:30.420 --> 01:00:31.640 where you want the output
NOTE Confidence: 0.9303634

01:00:31.780 --> 01:00:32.980 which to be, and it
NOTE Confidence: 0.9303634

01:00:32.980 --> 01:00:34.180 will give you run the
NOTE Confidence: 0.9303634

01:00:34.180 --> 01:00:35.460 entire Kiwi and give you
NOTE Confidence: 0.9303634

01:00:35.460 --> 01:00:36.335 the output there. Output
NOTE Confidence: 0.8357408

01:00:36.755 --> 01:00:37.255 there.
NOTE Confidence: 0.93385893

01:00:37.675 --> 01:00:39.075 Okay? So easy to install
NOTE Confidence: 0.93385893

01:00:39.075 --> 01:00:40.155 Docker images. All the dependencies,
NOTE Confidence: 0.93385893

01:00:40.155 --> 01:00:41.755 everything is taken care of.
NOTE Confidence: 0.93385893

01:00:41.755 --> 01:00:43.055 Can be run on Linux,
NOTE Confidence: 0.93385893

01:00:43.115 --> 01:00:45.055 Mac, Windows. You have CPU.
NOTE Confidence: 0.93385893

01:00:45.115 --> 01:00:46.075 You have we we have
NOTE Confidence: 0.93385893

01:00:46.075 --> 01:00:47.275 versions for that. You have
NOTE Confidence: 0.93385893

01:00:47.275 --> 01:00:48.839 GPU. We have versions for
NOTE Confidence: 0.93385893

01:00:48.839 --> 01:00:49.339 that.
NOTE Confidence: 0.97157735

01:00:49.799 --> 01:00:51.480 And, we have both the
NOTE Confidence: 0.97157735

01:00:51.480 --> 01:00:53.160 BERT based and LAMA based

NOTE Confidence: 0.97157735
01:00:53.160 --> 01:00:54.440 models that does the thing
NOTE Confidence: 0.97157735
01:00:54.440 --> 01:00:55.900 that I was talking about.
NOTE Confidence: 0.99944425
01:00:56.839 --> 01:00:57.339 Finally,
NOTE Confidence: 0.95296997
01:00:57.960 --> 01:00:58.460 what,
NOTE Confidence: 0.972604
01:00:59.400 --> 01:01:01.065 Vincent is gonna demo is
NOTE Confidence: 0.972604
01:01:01.224 --> 01:01:02.665 forget about all these things.
NOTE Confidence: 0.972604
01:01:02.665 --> 01:01:03.704 Your data is on the
NOTE Confidence: 0.972604
01:01:03.704 --> 01:01:04.984 chip. You want to directly
NOTE Confidence: 0.972604
01:01:04.984 --> 01:01:06.744 use it just with an
NOTE Confidence: 0.972604
01:01:06.744 --> 01:01:08.984 API call. Currently, you need
NOTE Confidence: 0.972604
01:01:08.984 --> 01:01:10.505 to contact Chris Gilman, who's
NOTE Confidence: 0.972604
01:01:10.505 --> 01:01:12.025 a senior software engineer, to
NOTE Confidence: 0.972604
01:01:12.025 --> 01:01:13.545 get that API for calling
NOTE Confidence: 0.972604
01:01:13.545 --> 01:01:15.099 the Kiwi. But in future,
NOTE Confidence: 0.972604
01:01:15.099 --> 01:01:15.980 we are gonna come up
NOTE Confidence: 0.972604

01:01:15.980 --> 01:01:16.940 with a system where you
NOTE Confidence: 0.972604

01:01:16.940 --> 01:01:18.380 can submit the tickets, get
NOTE Confidence: 0.972604

01:01:18.380 --> 01:01:19.740 the API key. So get
NOTE Confidence: 0.972604

01:01:19.740 --> 01:01:20.940 your API key, put it
NOTE Confidence: 0.972604

01:01:20.940 --> 01:01:22.060 into a program that we
NOTE Confidence: 0.972604

01:01:22.060 --> 01:01:23.099 are gonna give you, run
NOTE Confidence: 0.972604

01:01:23.099 --> 01:01:24.460 it. That's as easy as
NOTE Confidence: 0.972604

01:01:24.460 --> 01:01:25.200 it gets.
NOTE Confidence: 0.9668862

01:01:26.934 --> 01:01:28.855 A growing database, about thirty
NOTE Confidence: 0.9668862

01:01:28.855 --> 01:01:29.895 two requests so far since
NOTE Confidence: 0.9668862

01:01:29.895 --> 01:01:32.135 we released, and, that's it.
NOTE Confidence: 0.9668862

01:01:32.135 --> 01:01:33.115 I don't want to
NOTE Confidence: 0.9551652

01:01:33.655 --> 01:01:35.015 go more on that. What's
NOTE Confidence: 0.9551652

01:01:35.015 --> 01:01:37.015 coming up? We have more
NOTE Confidence: 0.9551652

01:01:37.015 --> 01:01:38.615 packages that we have actually
NOTE Confidence: 0.9551652

01:01:38.615 --> 01:01:39.895 built but not made it

NOTE Confidence: 0.9551652
01:01:39.895 --> 01:01:41.160 available in as a docker
NOTE Confidence: 0.9551652
01:01:41.160 --> 01:01:42.440 or a service or something
NOTE Confidence: 0.9551652
01:01:42.440 --> 01:01:43.800 like that. One of those
NOTE Confidence: 0.9551652
01:01:43.800 --> 01:01:44.680 that we are thinking of
NOTE Confidence: 0.9551652
01:01:44.680 --> 01:01:46.040 making it available on the
NOTE Confidence: 0.9551652
01:01:46.040 --> 01:01:47.800 chip or as, Kiwi is
NOTE Confidence: 0.9551652
01:01:47.800 --> 01:01:48.300 currently
NOTE Confidence: 0.9185796
01:01:48.680 --> 01:01:50.280 is the resist pipeline, which
NOTE Confidence: 0.9185796
01:01:50.280 --> 01:01:52.839 is extracting systematic anticancer therapy
NOTE Confidence: 0.9185796
01:01:52.839 --> 01:01:54.599 and the responses based on
NOTE Confidence: 0.9185796
01:01:54.599 --> 01:01:56.595 the, RECIST guidelines. Again, I'm
NOTE Confidence: 0.9185796
01:01:56.595 --> 01:01:57.555 not a clinician, so I'm
NOTE Confidence: 0.9185796
01:01:57.555 --> 01:01:58.755 not going on to it.
NOTE Confidence: 0.9185796
01:01:58.755 --> 01:02:00.115 So probably you can see
NOTE Confidence: 0.9185796
01:02:00.115 --> 01:02:01.635 something like the similar in
NOTE Confidence: 0.9185796

01:02:01.635 --> 01:02:02.135 future
NOTE Confidence: 0.978937

01:02:02.435 --> 01:02:02.935 available,
NOTE Confidence: 0.9626005

01:02:03.395 --> 01:02:04.835 like the stalker images or
NOTE Confidence: 0.9626005

01:02:04.835 --> 01:02:06.275 API services or something that
NOTE Confidence: 0.9626005

01:02:06.275 --> 01:02:07.475 you can download and play
NOTE Confidence: 0.9626005

01:02:07.475 --> 01:02:07.975 with.
NOTE Confidence: 0.9630013

01:02:09.280 --> 01:02:10.799 My main area of research,
NOTE Confidence: 0.9630013

01:02:10.799 --> 01:02:12.480 social determinants of health. This
NOTE Confidence: 0.9630013

01:02:12.480 --> 01:02:13.920 is another pipeline that I
NOTE Confidence: 0.9630013

01:02:13.920 --> 01:02:15.520 have built. Twenty one social
NOTE Confidence: 0.9630013

01:02:15.520 --> 01:02:16.900 determinants of health,
NOTE Confidence: 0.97470105

01:02:17.359 --> 01:02:18.799 four one, two, three, four.
NOTE Confidence: 0.97470105

01:02:18.799 --> 01:02:20.579 Yeah. Four different models.
NOTE Confidence: 0.9900526

01:02:20.985 --> 01:02:22.205 You start from XGBoost,
NOTE Confidence: 0.78630096

01:02:22.905 --> 01:02:23.405 TextCNN,
NOTE Confidence: 0.8884776

01:02:24.425 --> 01:02:24.925 SentenceBird,

NOTE Confidence: 0.9061602
01:02:25.305 --> 01:02:27.725 llama. After that, that actually
NOTE Confidence: 0.9842223
01:02:28.345 --> 01:02:30.185 can take your notes and
NOTE Confidence: 0.9842223
01:02:30.185 --> 01:02:31.865 annotate it on two levels,
NOTE Confidence: 0.9842223
01:02:31.865 --> 01:02:33.565 on twenty one social determinants,
NOTE Confidence: 0.9614623
01:02:34.105 --> 01:02:36.180 determinant factors. And let's say,
NOTE Confidence: 0.9915275
01:02:36.900 --> 01:02:37.859 it does a sort of
NOTE Confidence: 0.9915275
01:02:37.859 --> 01:02:38.839 sentence classification.
NOTE Confidence: 0.97791505
01:02:39.300 --> 01:02:40.500 It takes your note. It
NOTE Confidence: 0.97791505
01:02:40.500 --> 01:02:42.260 divides it into sentences, tells
NOTE Confidence: 0.97791505
01:02:42.260 --> 01:02:43.460 you, okay. This sentence is
NOTE Confidence: 0.97791505
01:02:43.460 --> 01:02:45.319 talking about race, sex, gender.
NOTE Confidence: 0.97791505
01:02:45.540 --> 01:02:46.819 This sentence is talking about
NOTE Confidence: 0.97791505
01:02:46.819 --> 01:02:48.280 the insurance of the person.
NOTE Confidence: 0.97791505
01:02:48.425 --> 01:02:49.545 So this sentence is talking
NOTE Confidence: 0.97791505
01:02:49.545 --> 01:02:50.984 about their education. Now we
NOTE Confidence: 0.97791505

01:02:50.984 --> 01:02:52.345 go one more level. You
NOTE Confidence: 0.97791505

01:02:52.345 --> 01:02:53.865 will also have models that
NOTE Confidence: 0.97791505

01:02:53.865 --> 01:02:55.484 tells you, okay. This education
NOTE Confidence: 0.97791505

01:02:55.625 --> 01:02:57.224 of this person is high
NOTE Confidence: 0.97791505

01:02:57.224 --> 01:02:59.005 school or below. The insurance,
NOTE Confidence: 0.97791505

01:02:59.224 --> 01:03:00.425 it's yes. The person has
NOTE Confidence: 0.97791505

01:03:00.585 --> 01:03:02.265 having an insurance or no.
NOTE Confidence: 0.97791505

01:03:02.265 --> 01:03:04.010 So high level on the
NOTE Confidence: 0.97791505

01:03:04.010 --> 01:03:05.930 twenty one factors, both the
NOTE Confidence: 0.97791505

01:03:05.930 --> 01:03:07.150 values and attributes
NOTE Confidence: 0.9341416

01:03:07.690 --> 01:03:08.910 on another
NOTE Confidence: 0.92397285

01:03:09.450 --> 01:03:11.210 digging another level deep. So
NOTE Confidence: 0.92397285

01:03:11.210 --> 01:03:12.330 that's all we have here.
NOTE Confidence: 0.92397285

01:03:12.330 --> 01:03:14.170 And, also, I just wanna,
NOTE Confidence: 0.966931

01:03:14.730 --> 01:03:15.550 forgot that.
NOTE Confidence: 0.9979821

01:03:16.235 --> 01:03:16.735 Yeah.

NOTE Confidence: 0.9942718
01:03:17.515 --> 01:03:18.975 When you sign a DUA
NOTE Confidence: 0.9942718
01:03:19.035 --> 01:03:19.695 with us,
NOTE Confidence: 0.94981
01:03:20.155 --> 01:03:21.275 we are gonna give you
NOTE Confidence: 0.94981
01:03:21.275 --> 01:03:22.815 the model weights of KB
NOTE Confidence: 0.94981
01:03:22.875 --> 01:03:23.835 that is still in the
NOTE Confidence: 0.94981
01:03:23.835 --> 01:03:25.615 pipeline. It will come here
NOTE Confidence: 0.94981
01:03:25.675 --> 01:03:27.035 in the form. So in
NOTE Confidence: 0.94981
01:03:27.035 --> 01:03:28.395 this form that I'm asking
NOTE Confidence: 0.94981
01:03:28.395 --> 01:03:29.355 you to fill to get
NOTE Confidence: 0.94981
01:03:29.355 --> 01:03:31.000 the docker images would also
NOTE Confidence: 0.94981
01:03:31.000 --> 01:03:32.040 be if you are good
NOTE Confidence: 0.94981
01:03:32.040 --> 01:03:32.780 in programming,
NOTE Confidence: 0.9660935
01:03:33.240 --> 01:03:34.220 take our model,
NOTE Confidence: 0.96351975
01:03:35.000 --> 01:03:37.000 you know, continuously fine tune
NOTE Confidence: 0.96351975
01:03:37.000 --> 01:03:38.360 on it with your data,
NOTE Confidence: 0.96351975

01:03:38.360 --> 01:03:39.720 make it whatever you want
NOTE Confidence: 0.96351975

01:03:39.720 --> 01:03:40.860 to do with it. So,
NOTE Confidence: 0.95925736

01:03:41.240 --> 01:03:42.440 that is another thing, but
NOTE Confidence: 0.95925736

01:03:42.440 --> 01:03:43.240 you need to sign a
NOTE Confidence: 0.95925736

01:03:43.240 --> 01:03:44.440 DUA with us, and that
NOTE Confidence: 0.95925736

01:03:44.440 --> 01:03:45.880 form will be available soon
NOTE Confidence: 0.95925736

01:03:45.880 --> 01:03:47.765 there. With that, Vincent, take
NOTE Confidence: 0.95925736

01:03:47.765 --> 01:03:48.885 it over for a Kiwi
NOTE Confidence: 0.95925736

01:03:48.885 --> 01:03:49.785 API demo.
NOTE Confidence: 0.7829319

01:03:56.485 --> 01:03:57.925 Not taking questions because of
NOTE Confidence: 0.7829319

01:03:57.925 --> 01:03:59.365 the time that we focus
NOTE Confidence: 0.7829319

01:03:59.365 --> 01:03:59.865 on.
NOTE Confidence: 0.8796835

01:04:06.780 --> 01:04:08.619 So good afternoon, everyone. My
NOTE Confidence: 0.8796835

01:04:08.619 --> 01:04:10.220 name is Vincent, and I'm
NOTE Confidence: 0.8796835

01:04:10.220 --> 01:04:12.220 a software developer in doctor.
NOTE Confidence: 0.8796835

01:04:12.220 --> 01:04:12.960 She labs.

NOTE Confidence: 0.90669197
01:04:13.260 --> 01:04:14.780 And today, I will talk
NOTE Confidence: 0.90669197
01:04:14.780 --> 01:04:16.140 about how to use the
NOTE Confidence: 0.90669197
01:04:16.140 --> 01:04:17.440 Kiwi API service.
NOTE Confidence: 0.92170286
01:04:18.085 --> 01:04:19.845 And the core concept is
NOTE Confidence: 0.92170286
01:04:20.005 --> 01:04:21.685 of of the Kiwi have
NOTE Confidence: 0.92170286
01:04:21.685 --> 01:04:23.965 already become discussed with, by,
NOTE Confidence: 0.92170286
01:04:23.965 --> 01:04:25.845 so I will talk very
NOTE Confidence: 0.92170286
01:04:25.845 --> 01:04:26.345 quick.
NOTE Confidence: 0.9445003
01:04:27.605 --> 01:04:29.865 So what is the Kiwi's
NOTE Confidence: 0.784585
01:04:30.950 --> 01:04:32.170 so API service?
NOTE Confidence: 0.8952173
01:04:32.549 --> 01:04:34.650 The Kiwi API service provide
NOTE Confidence: 0.8952173
01:04:34.710 --> 01:04:36.549 an API as a service
NOTE Confidence: 0.8952173
01:04:36.549 --> 01:04:37.049 interface
NOTE Confidence: 0.8114218
01:04:37.509 --> 01:04:39.130 that allow user within
NOTE Confidence: 0.28079024
01:04:39.910 --> 01:04:40.410 chips
NOTE Confidence: 0.858168

01:04:40.710 --> 01:04:42.309 in internal network to access
NOTE Confidence: 0.858168

01:04:42.309 --> 01:04:43.849 a Kiwi without a request,
NOTE Confidence: 0.8910007

01:04:44.435 --> 01:04:45.955 high performance GPU and having
NOTE Confidence: 0.8910007

01:04:45.955 --> 01:04:46.775 to install
NOTE Confidence: 0.9923417

01:04:47.395 --> 01:04:49.255 or manage the model locally.
NOTE Confidence: 0.9069159

01:04:50.115 --> 01:04:52.535 User simply request a API
NOTE Confidence: 0.9069159

01:04:52.595 --> 01:04:55.095 key and make standard HTTP
NOTE Confidence: 0.9069159

01:04:55.155 --> 01:04:56.755 API API calls to use
NOTE Confidence: 0.9069159

01:04:56.755 --> 01:04:57.415 the service.
NOTE Confidence: 0.902738

01:04:58.370 --> 01:04:59.190 All computational
NOTE Confidence: 0.9272547

01:04:59.490 --> 01:05:01.010 resource are running on chips,
NOTE Confidence: 0.9272547

01:05:01.330 --> 01:05:02.690 so and, we don't need
NOTE Confidence: 0.9272547

01:05:02.690 --> 01:05:04.950 to request the local GPU,
NOTE Confidence: 0.99504554

01:05:05.410 --> 01:05:05.910 GPU.
NOTE Confidence: 0.5395022

01:05:06.450 --> 01:05:08.310 Does this setup of streamline
NOTE Confidence: 0.5395022

01:05:08.450 --> 01:05:10.150 assess TV functionality

NOTE Confidence: 0.9547893
01:05:10.770 --> 01:05:13.365 and, makes it more accessible
NOTE Confidence: 0.9547893
01:05:13.425 --> 01:05:14.725 in resource constraint
NOTE Confidence: 0.98630464
01:05:15.025 --> 01:05:15.525 environment.
NOTE Confidence: 0.9745125
01:05:19.505 --> 01:05:20.865 So how does the Kiwi
NOTE Confidence: 0.9745125
01:05:20.865 --> 01:05:22.965 API service actually work?
NOTE Confidence: 0.94779617
01:05:23.345 --> 01:05:25.345 The process follow a simple
NOTE Confidence: 0.94779617
01:05:25.345 --> 01:05:26.305 request and,
NOTE Confidence: 0.72709507
01:05:26.705 --> 01:05:27.285 a response
NOTE Confidence: 0.99211264
01:05:28.760 --> 01:05:29.240 response.
NOTE Confidence: 0.8826221
01:05:29.560 --> 01:05:31.240 User can send a request
NOTE Confidence: 0.8826221
01:05:31.240 --> 01:05:33.400 to the Kiwi API server
NOTE Confidence: 0.8826221
01:05:33.400 --> 01:05:35.080 in chips environment such as
NOTE Confidence: 0.8826221
01:05:35.080 --> 01:05:36.880 the cam Camino, which include
NOTE Confidence: 0.8826221
01:05:36.880 --> 01:05:37.720 the either,
NOTE Confidence: 0.83220273
01:05:38.040 --> 01:05:40.040 the clinical notes or other
NOTE Confidence: 0.83220273

01:05:40.040 --> 01:05:41.500 tests related data.
NOTE Confidence: 0.91571224

01:05:42.055 --> 01:05:43.895 Once the API service received
NOTE Confidence: 0.91571224

01:05:43.895 --> 01:05:45.755 the request, it's determined
NOTE Confidence: 0.9738865

01:05:46.375 --> 01:05:47.895 the task type, based on
NOTE Confidence: 0.9738865

01:05:47.895 --> 01:05:48.555 the specific
NOTE Confidence: 0.7741092

01:05:49.015 --> 01:05:51.015 endpoint and then return the
NOTE Confidence: 0.7741092

01:05:51.015 --> 01:05:53.595 up appropriate appreciate the response.
NOTE Confidence: 0.98214275

01:05:54.680 --> 01:05:56.760 Most tasks are handled by
NOTE Confidence: 0.98214275

01:05:56.760 --> 01:05:58.440 the background process on the
NOTE Confidence: 0.98214275

01:05:58.440 --> 01:05:59.500 API server.
NOTE Confidence: 0.94073963

01:05:59.960 --> 01:06:01.800 All incoming requests are and
NOTE Confidence: 0.94073963

01:06:01.800 --> 01:06:02.300 queued
NOTE Confidence: 0.84834355

01:06:02.600 --> 01:06:04.620 is queued and processed sequentially
NOTE Confidence: 0.8549565

01:06:05.560 --> 01:06:07.560 to issue the efficient using
NOTE Confidence: 0.8549565

01:06:07.560 --> 01:06:08.780 the lim and imitate,
NOTE Confidence: 0.98751163

01:06:09.654 --> 01:06:10.714 computational resource.

NOTE Confidence: 0.99940443
01:06:11.255 --> 01:06:11.755 So
NOTE Confidence: 0.9987999
01:06:13.494 --> 01:06:14.234 and then
NOTE Confidence: 0.93707514
01:06:14.535 --> 01:06:16.775 let's take a closer look
NOTE Confidence: 0.93707514
01:06:16.775 --> 01:06:18.535 at how to use the
NOTE Confidence: 0.93707514
01:06:18.535 --> 01:06:19.674 QApi service.
NOTE Confidence: 0.8634327
01:06:20.375 --> 01:06:21.835 Before we getting start,
NOTE Confidence: 0.97198963
01:06:23.030 --> 01:06:24.550 there are something you need
NOTE Confidence: 0.97198963
01:06:24.550 --> 01:06:26.010 to prepare. First,
NOTE Confidence: 0.8105896
01:06:26.550 --> 01:06:28.089 obviously, you need,
NOTE Confidence: 0.9468236
01:06:28.550 --> 01:06:30.630 access in the chips environment
NOTE Confidence: 0.9468236
01:06:30.630 --> 01:06:31.690 like the Camino.
NOTE Confidence: 0.8192455
01:06:32.150 --> 01:06:33.430 You need to have a
NOTE Confidence: 0.8192455
01:06:33.430 --> 01:06:34.250 one HX
NOTE Confidence: 0.697598
01:06:34.550 --> 01:06:35.050 account.
NOTE Confidence: 0.9987297
01:06:35.755 --> 01:06:36.494 Then you
NOTE Confidence: 0.83041006

01:06:36.795 --> 01:06:38.255 need to answer right API
NOTE Confidence: 0.83041006

01:06:38.395 --> 01:06:39.835 key, for the detail just
NOTE Confidence: 0.83041006

01:06:39.835 --> 01:06:40.734 mentioned by,
NOTE Confidence: 0.85277486

01:06:42.474 --> 01:06:43.755 You need to ask Chris
NOTE Confidence: 0.85277486

01:06:43.755 --> 01:06:44.955 Gellman to get the API
NOTE Confidence: 0.85277486

01:06:44.955 --> 01:06:45.455 key.
NOTE Confidence: 0.9537569

01:06:47.035 --> 01:06:48.715 For the user who have
NOTE Confidence: 0.9537569

01:06:48.715 --> 01:06:50.830 some coding experience, they can
NOTE Confidence: 0.9537569

01:06:50.890 --> 01:06:52.410 write their own script to
NOTE Confidence: 0.9537569

01:06:52.410 --> 01:06:54.010 access the API, but we
NOTE Confidence: 0.9537569

01:06:54.010 --> 01:06:55.610 also provide the API launch
NOTE Confidence: 0.9537569

01:06:55.610 --> 01:06:57.130 script and some use case
NOTE Confidence: 0.9537569

01:06:57.130 --> 01:06:57.630 include
NOTE Confidence: 0.8410722

01:06:58.010 --> 01:07:00.010 in Jupyter Notebook provided under
NOTE Confidence: 0.8410722

01:07:00.010 --> 01:07:01.070 this, GitHub,
NOTE Confidence: 0.9981334

01:07:01.370 --> 01:07:01.870 link.

NOTE Confidence: 0.8907243
01:07:06.055 --> 01:07:07.655 Now assuming you, you already
NOTE Confidence: 0.8907243
01:07:07.655 --> 01:07:09.095 have the access in into
NOTE Confidence: 0.8907243
01:07:09.095 --> 01:07:10.215 the formula and then you
NOTE Confidence: 0.8907243
01:07:10.215 --> 01:07:11.915 you open a Jupyter notebook,
NOTE Confidence: 0.9185614
01:07:12.375 --> 01:07:13.495 and then you have a
NOTE Confidence: 0.9185614
01:07:13.495 --> 01:07:15.655 a key API key. First,
NOTE Confidence: 0.9185614
01:07:15.895 --> 01:07:17.015 you need to define a
NOTE Confidence: 0.9185614
01:07:17.015 --> 01:07:18.235 variable to instantiate
NOTE Confidence: 0.97580075
01:07:18.615 --> 01:07:20.160 the the class, that I
NOTE Confidence: 0.97580075
01:07:20.160 --> 01:07:21.140 provide in the script.
NOTE Confidence: 0.9803024
01:07:21.440 --> 01:07:22.799 At this step, you need
NOTE Confidence: 0.9803024
01:07:22.799 --> 01:07:23.460 to insert
NOTE Confidence: 0.9324588
01:07:24.000 --> 01:07:25.920 your API key and into
NOTE Confidence: 0.9324588
01:07:25.920 --> 01:07:26.660 this instance.
NOTE Confidence: 0.9747385
01:07:27.520 --> 01:07:28.660 For the first time,
NOTE Confidence: 0.9932249

01:07:29.119 --> 01:07:30.480 for using the API key
NOTE Confidence: 0.9932249

01:07:30.480 --> 01:07:30.980 server,
NOTE Confidence: 0.87869626

01:07:31.920 --> 01:07:33.039 you can use the key
NOTE Confidence: 0.87869626

01:07:33.039 --> 01:07:34.880 info function to test your
NOTE Confidence: 0.87869626

01:07:34.880 --> 01:07:35.380 connection.
NOTE Confidence: 0.92963815

01:07:36.995 --> 01:07:39.075 This will give some, result.
NOTE Confidence: 0.92963815

01:07:39.075 --> 01:07:39.575 Yeah.
NOTE Confidence: 0.86020446

01:07:39.955 --> 01:07:42.115 It, responds as a JSON
NOTE Confidence: 0.86020446

01:07:42.115 --> 01:07:42.615 format,
NOTE Confidence: 0.99909425

01:07:42.995 --> 01:07:43.495 information.
NOTE Confidence: 0.9098372

01:07:44.435 --> 01:07:46.135 There are three main component
NOTE Confidence: 0.9098372

01:07:46.195 --> 01:07:47.655 in the in this response.
NOTE Confidence: 0.9098372

01:07:47.715 --> 01:07:49.255 You can see the usage
NOTE Confidence: 0.9098372

01:07:49.475 --> 01:07:51.580 count and, which tells you
NOTE Confidence: 0.9098372

01:07:51.580 --> 01:07:52.780 how many tokens you use
NOTE Confidence: 0.9098372

01:07:52.780 --> 01:07:54.300 since you create the API

NOTE Confidence: 0.9098372
01:07:54.300 --> 01:07:56.400 key. And the token remain,
NOTE Confidence: 0.9448445
01:07:56.780 --> 01:07:58.220 which tells you how many
NOTE Confidence: 0.9448445
01:07:58.220 --> 01:07:59.900 token you still have in
NOTE Confidence: 0.9448445
01:07:59.900 --> 01:08:01.660 the API key. Finally, the
NOTE Confidence: 0.9448445
01:08:01.660 --> 01:08:03.580 expire at, tells you when
NOTE Confidence: 0.9448445
01:08:03.580 --> 01:08:04.720 the key is expired.
NOTE Confidence: 0.94172776
01:08:05.805 --> 01:08:07.244 For the token, the expire
NOTE Confidence: 0.94172776
01:08:07.244 --> 01:08:09.085 state, you can contact our
NOTE Confidence: 0.94172776
01:08:09.085 --> 01:08:10.445 team to add add in
NOTE Confidence: 0.94172776
01:08:10.445 --> 01:08:12.065 the usage in the future.
NOTE Confidence: 0.92272633
01:08:13.484 --> 01:08:15.244 Our main function process is
NOTE Confidence: 0.92272633
01:08:15.244 --> 01:08:16.385 the batch prediction,
NOTE Confidence: 0.9303894
01:08:16.765 --> 01:08:18.685 which allow user process their
NOTE Confidence: 0.9303894
01:08:18.685 --> 01:08:20.385 clinical node in the bark.
NOTE Confidence: 0.9980365
01:08:21.330 --> 01:08:22.550 To use this function,
NOTE Confidence: 0.9676613

01:08:23.170 --> 01:08:24.290 you need to provide the
NOTE Confidence: 0.9676613

01:08:24.290 --> 01:08:25.810 path of your files in
NOTE Confidence: 0.9676613

01:08:25.889 --> 01:08:28.469 into in the communal environment.
NOTE Confidence: 0.8491341

01:08:29.010 --> 01:08:30.929 Currently, the upload format are
NOTE Confidence: 0.8491341

01:08:30.929 --> 01:08:32.369 support to compress the files
NOTE Confidence: 0.8491341

01:08:32.369 --> 01:08:33.489 such as the deep or
NOTE Confidence: 0.8491341

01:08:33.489 --> 01:08:35.514 tar or the single text
NOTE Confidence: 0.8491341

01:08:35.514 --> 01:08:36.554 t file, and you can
NOTE Confidence: 0.8491341

01:08:36.554 --> 01:08:37.354 compress your,
NOTE Confidence: 0.94896877

01:08:37.835 --> 01:08:39.434 node into a single text
NOTE Confidence: 0.94896877

01:08:39.434 --> 01:08:40.494 t file as well.
NOTE Confidence: 0.8357911

01:08:40.875 --> 01:08:41.854 If the sit,
NOTE Confidence: 0.9040675

01:08:42.395 --> 01:08:43.594 your note will on the
NOTE Confidence: 0.9040675

01:08:43.594 --> 01:08:44.954 Kiwi server as a task
NOTE Confidence: 0.9040675

01:08:44.954 --> 01:08:46.715 in the queue. The function
NOTE Confidence: 0.9040675

01:08:46.715 --> 01:08:48.155 will return the text status

NOTE Confidence: 0.9040675
01:08:48.155 --> 01:08:50.150 including the test ID for,
NOTE Confidence: 0.933035
01:08:50.530 --> 01:08:51.650 for this test as a
NOTE Confidence: 0.933035
01:08:51.650 --> 01:08:52.550 JSON format.
NOTE Confidence: 0.93430734
01:08:53.090 --> 01:08:54.290 All all test ID are
NOTE Confidence: 0.93430734
01:08:54.290 --> 01:08:56.229 related to the API key,
NOTE Confidence: 0.93430734
01:08:56.530 --> 01:08:57.810 which means all the user
NOTE Confidence: 0.93430734
01:08:57.810 --> 01:09:00.210 data is isolate isolated by
NOTE Confidence: 0.93430734
01:09:00.210 --> 01:09:01.250 the API key and the
NOTE Confidence: 0.93430734
01:09:01.250 --> 01:09:02.150 test ID.
NOTE Confidence: 0.9574709
01:09:04.025 --> 01:09:05.865 Here, is an example when
NOTE Confidence: 0.9574709
01:09:05.865 --> 01:09:08.045 you do a batch prediction.
NOTE Confidence: 0.9711923
01:09:09.145 --> 01:09:10.345 As you can see, it's
NOTE Confidence: 0.9711923
01:09:10.345 --> 01:09:12.185 a return, JSON format of
NOTE Confidence: 0.9711923
01:09:12.185 --> 01:09:13.245 the test information.
NOTE Confidence: 0.8740358
01:09:14.104 --> 01:09:16.265 It's including the test ID
NOTE Confidence: 0.8740358

01:09:16.265 --> 01:09:17.385 and the message that shows
NOTE Confidence: 0.8740358

01:09:17.385 --> 01:09:19.520 the state, status and how
NOTE Confidence: 0.8740358

01:09:19.520 --> 01:09:21.860 many tokens using this, task
NOTE Confidence: 0.8740358

01:09:22.080 --> 01:09:23.860 and, how many token remaining
NOTE Confidence: 0.8740358

01:09:24.000 --> 01:09:24.479 your,
NOTE Confidence: 0.7898215

01:09:25.040 --> 01:09:25.540 account.
NOTE Confidence: 0.9504511

01:09:26.160 --> 01:09:28.160 Finally, the estimate time for
NOTE Confidence: 0.9504511

01:09:28.160 --> 01:09:29.600 your task is calculated based
NOTE Confidence: 0.9504511

01:09:29.600 --> 01:09:31.380 under your task queue position
NOTE Confidence: 0.9504511

01:09:31.440 --> 01:09:32.800 and the progress of the
NOTE Confidence: 0.9504511

01:09:32.800 --> 01:09:34.695 task, ahead of it.
NOTE Confidence: 0.9336694

01:09:37.014 --> 01:09:38.534 After you submit a task
NOTE Confidence: 0.9336694

01:09:38.534 --> 01:09:40.395 and receive a task ID,
NOTE Confidence: 0.937109

01:09:40.775 --> 01:09:42.135 you can use your task
NOTE Confidence: 0.937109

01:09:42.135 --> 01:09:43.494 ID to check the current
NOTE Confidence: 0.937109

01:09:43.494 --> 01:09:45.094 status of your task at

NOTE Confidence: 0.937109
01:09:45.094 --> 01:09:46.614 anytime using the task ID
NOTE Confidence: 0.937109
01:09:46.614 --> 01:09:47.514 status function.
NOTE Confidence: 0.88635296
01:09:48.600 --> 01:09:50.140 It will provide your,
NOTE Confidence: 0.9064552
01:09:50.520 --> 01:09:52.220 task information and the detail.
NOTE Confidence: 0.99260366
01:09:53.479 --> 01:09:55.320 Typically, there are three main,
NOTE Confidence: 0.98540896
01:09:55.720 --> 01:09:56.860 status of your,
NOTE Confidence: 0.7789945
01:09:58.040 --> 01:09:59.240 the test data. So first
NOTE Confidence: 0.7789945
01:09:59.240 --> 01:10:00.540 is the queue status.
NOTE Confidence: 0.5696389
01:10:01.054 --> 01:10:01.854 When a state,
NOTE Confidence: 0.9653566
01:10:02.255 --> 01:10:03.534 task in the queue, the
NOTE Confidence: 0.9653566
01:10:03.534 --> 01:10:05.275 system will provide the,
NOTE Confidence: 0.8900483
01:10:06.175 --> 01:10:08.014 test current queue position as
NOTE Confidence: 0.8900483
01:10:08.014 --> 01:10:09.775 well as the estimate time
NOTE Confidence: 0.8900483
01:10:09.775 --> 01:10:10.594 for processing.
NOTE Confidence: 0.86444855
01:10:11.295 --> 01:10:12.594 And next is processing.
NOTE Confidence: 0.8901205

01:10:13.135 --> 01:10:14.574 When no no one attack
NOTE Confidence: 0.8901205

01:10:14.574 --> 01:10:15.074 you,
NOTE Confidence: 0.9113512

01:10:16.140 --> 01:10:17.740 your task will put into
NOTE Confidence: 0.9113512

01:10:17.740 --> 01:10:19.580 a process or pull. It
NOTE Confidence: 0.9113512

01:10:19.580 --> 01:10:21.260 will in indicate how many
NOTE Confidence: 0.9113512

01:10:21.260 --> 01:10:23.500 files into your task and,
NOTE Confidence: 0.9130504

01:10:24.140 --> 01:10:25.820 include how many not yet
NOTE Confidence: 0.9130504

01:10:25.820 --> 01:10:26.320 processed
NOTE Confidence: 0.9432637

01:10:26.860 --> 01:10:28.300 and, how many have you
NOTE Confidence: 0.9432637

01:10:28.300 --> 01:10:29.040 been proceed
NOTE Confidence: 0.97265935

01:10:29.495 --> 01:10:31.175 and, also, remaining time based
NOTE Confidence: 0.97265935

01:10:31.175 --> 01:10:32.314 on the remaining files.
NOTE Confidence: 0.9446643

01:10:32.695 --> 01:10:34.314 Finally, the incomplete
NOTE Confidence: 0.96978146

01:10:35.175 --> 01:10:36.615 status, you can use the
NOTE Confidence: 0.96978146

01:10:36.615 --> 01:10:38.055 test ID in the next
NOTE Confidence: 0.96978146

01:10:38.055 --> 01:10:39.915 function to download your result.

NOTE Confidence: 0.8054798
01:10:42.229 --> 01:10:43.670 Once your, test is start
NOTE Confidence: 0.8054798
01:10:43.670 --> 01:10:44.170 complete,
NOTE Confidence: 0.9988915
01:10:44.550 --> 01:10:46.410 you can download your
NOTE Confidence: 0.7438547
01:10:46.790 --> 01:10:48.650 test using the type download
NOTE Confidence: 0.7438547
01:10:48.710 --> 01:10:49.210 function.
NOTE Confidence: 0.9195943
01:10:49.670 --> 01:10:50.870 In this function, you need
NOTE Confidence: 0.9195943
01:10:50.870 --> 01:10:52.710 to give output path, and,
NOTE Confidence: 0.9195943
01:10:52.949 --> 01:10:54.469 then you want to save
NOTE Confidence: 0.9195943
01:10:54.469 --> 01:10:56.064 your, file as a add
NOTE Confidence: 0.9195943
01:10:56.064 --> 01:10:58.224 local and the output type
NOTE Confidence: 0.9195943
01:10:58.224 --> 01:10:59.284 you you prefer.
NOTE Confidence: 0.95380616
01:10:59.824 --> 01:11:01.425 By default, the path are
NOTE Confidence: 0.95380616
01:11:01.425 --> 01:11:02.645 the working directory,
NOTE Confidence: 0.8554151
01:11:03.264 --> 01:11:04.645 and the the output,
NOTE Confidence: 0.8684583
01:11:05.025 --> 01:11:06.564 output is the JSON format.
NOTE Confidence: 0.8684583

01:11:06.704 --> 01:11:08.224 For the file saving, it
NOTE Confidence: 0.8684583

01:11:08.224 --> 01:11:10.309 supported three types of the
NOTE Confidence: 0.8684583

01:11:10.309 --> 01:11:12.789 tip that typically use. First
NOTE Confidence: 0.8684583

01:11:12.789 --> 01:11:13.929 is seed file,
NOTE Confidence: 0.9138216

01:11:14.389 --> 01:11:15.829 which compress the all result
NOTE Confidence: 0.9138216

01:11:15.829 --> 01:11:17.429 into a separate JSON file
NOTE Confidence: 0.9138216

01:11:17.429 --> 01:11:18.889 for your input files.
NOTE Confidence: 0.92107606

01:11:19.269 --> 01:11:21.110 And then it, insert the
NOTE Confidence: 0.92107606

01:11:21.110 --> 01:11:23.030 JSON, which combined all the
NOTE Confidence: 0.92107606

01:11:23.030 --> 01:11:24.630 JSON result into a single
NOTE Confidence: 0.92107606

01:11:24.630 --> 01:11:25.485 JSON file.
NOTE Confidence: 0.80277455

01:11:25.885 --> 01:11:27.185 Finally, it's the CSVs,
NOTE Confidence: 0.6121961

01:11:27.885 --> 01:11:28.844 where we've seen our mission.
NOTE Confidence: 0.6121961

01:11:28.844 --> 01:11:30.225 We have a con convert,
NOTE Confidence: 0.7086864

01:11:30.844 --> 01:11:32.764 integrate into the j the
NOTE Confidence: 0.7086864

01:11:32.764 --> 01:11:34.284 eight QVAP service so they

NOTE Confidence: 0.7086864
01:11:34.284 --> 01:11:34.925 can just,
NOTE Confidence: 0.99463415
01:11:35.485 --> 01:11:35.985 simply,
NOTE Confidence: 0.8339813
01:11:36.604 --> 01:11:37.824 output CSV.
NOTE Confidence: 0.9536475
01:11:39.060 --> 01:11:40.740 Each file can only be
NOTE Confidence: 0.9536475
01:11:40.740 --> 01:11:43.220 download once. After you download,
NOTE Confidence: 0.9536475
01:11:43.220 --> 01:11:44.260 you cannot access,
NOTE Confidence: 0.9304161
01:11:44.660 --> 01:11:45.640 again because,
NOTE Confidence: 0.95906
01:11:46.100 --> 01:11:47.240 for the some,
NOTE Confidence: 0.91830313
01:11:47.860 --> 01:11:49.880 privacy issue, they will delete
NOTE Confidence: 0.9482759
01:11:50.260 --> 01:11:51.700 the delete the re record
NOTE Confidence: 0.9482759
01:11:51.700 --> 01:11:52.520 of the data.
NOTE Confidence: 0.97199553
01:11:53.564 --> 01:11:55.265 And, here is a example
NOTE Confidence: 0.97199553
01:11:55.485 --> 01:11:57.325 for download. Typically, it will
NOTE Confidence: 0.97199553
01:11:57.325 --> 01:11:59.325 save your file into local
NOTE Confidence: 0.97199553
01:11:59.325 --> 01:12:00.765 directory and give you some
NOTE Confidence: 0.97199553

01:12:00.765 --> 01:12:01.885 message to tell you if
NOTE Confidence: 0.97199553

01:12:01.885 --> 01:12:02.945 this, succeed.
NOTE Confidence: 0.91699857

01:12:03.564 --> 01:12:04.685 And the left side, it
NOTE Confidence: 0.91699857

01:12:04.685 --> 01:12:06.380 should be the, like, the
NOTE Confidence: 0.91699857

01:12:06.380 --> 01:12:07.920 typical re re
NOTE Confidence: 0.8329768

01:12:08.220 --> 01:12:08.960 result format.
NOTE Confidence: 0.9878973

01:12:10.460 --> 01:12:10.939 And,
NOTE Confidence: 0.91330796

01:12:11.340 --> 01:12:13.180 in some case, you might
NOTE Confidence: 0.91330796

01:12:13.180 --> 01:12:15.840 submit multiple task and or
NOTE Confidence: 0.83872336

01:12:16.220 --> 01:12:17.439 forget a specific
NOTE Confidence: 0.94763196

01:12:18.025 --> 01:12:20.105 task ID, this function allow
NOTE Confidence: 0.94763196

01:12:20.105 --> 01:12:21.865 you to quickly review the
NOTE Confidence: 0.94763196

01:12:21.865 --> 01:12:23.405 status of each text.
NOTE Confidence: 0.90146637

01:12:24.185 --> 01:12:25.785 This this function will list
NOTE Confidence: 0.90146637

01:12:25.785 --> 01:12:27.385 all tests that not yet
NOTE Confidence: 0.90146637

01:12:27.385 --> 01:12:27.885 downloaded.

NOTE Confidence: 0.9811064
01:12:28.345 --> 01:12:29.405 Yeah. Yeah.
NOTE Confidence: 0.88451487
01:12:30.310 --> 01:12:32.710 And, if you exit as
NOTE Confidence: 0.88451487
01:12:33.030 --> 01:12:33.530 accidentally
NOTE Confidence: 0.84797645
01:12:33.989 --> 01:12:35.750 submit your task, if the
NOTE Confidence: 0.84797645
01:12:35.750 --> 01:12:36.550 task is not,
NOTE Confidence: 0.96604604
01:12:37.030 --> 01:12:38.090 into a process,
NOTE Confidence: 0.9409562
01:12:38.710 --> 01:12:39.929 you can still using,
NOTE Confidence: 0.99459124
01:12:40.390 --> 01:12:41.530 this, this
NOTE Confidence: 0.9532242
01:12:41.830 --> 01:12:43.989 function to cancel your task
NOTE Confidence: 0.9532242
01:12:43.989 --> 01:12:45.610 be, before the
NOTE Confidence: 0.9339288
01:12:46.104 --> 01:12:46.925 task is into
NOTE Confidence: 0.91275865
01:12:47.545 --> 01:12:48.525 the process pool,
NOTE Confidence: 0.9973192
01:12:48.985 --> 01:12:50.824 and, they will give you
NOTE Confidence: 0.9973192
01:12:50.824 --> 01:12:51.965 the token back.
NOTE Confidence: 0.9953171
01:12:52.985 --> 01:12:54.264 So I will show you
NOTE Confidence: 0.9953171

01:12:54.264 --> 01:12:55.804 a quick demo for them.
NOTE Confidence: 0.88430554

01:13:16.615 --> 01:13:17.735 Okay. Sure. So
NOTE Confidence: 0.8985248

01:13:18.215 --> 01:13:19.415 So I think we'll skip
NOTE Confidence: 0.8985248

01:13:19.415 --> 01:13:20.455 the demo and move to
NOTE Confidence: 0.8985248

01:13:20.455 --> 01:13:22.055 the next speaker. It's the
NOTE Confidence: 0.8985248

01:13:22.055 --> 01:13:23.335 same things, but you see
NOTE Confidence: 0.8985248

01:13:23.335 --> 01:13:24.615 it on coming on. Like,
NOTE Confidence: 0.8985248

01:13:24.615 --> 01:13:25.814 I'll actually showed you the
NOTE Confidence: 0.8985248

01:13:25.814 --> 01:13:26.875 program there. Right.
NOTE Confidence: 0.77266836

01:13:27.495 --> 01:13:27.995 Yeah.
NOTE Confidence: 0.9232161

01:13:28.979 --> 01:13:29.640 Thank you.
NOTE Confidence: 0.9197787

01:13:35.380 --> 01:13:36.900 Oh, hi, everyone. My name
NOTE Confidence: 0.9197787

01:13:36.900 --> 01:13:38.580 is Lingfei Chen. I am
NOTE Confidence: 0.9197787

01:13:38.820 --> 01:13:39.939 and, I am a postdoc
NOTE Confidence: 0.9197787

01:13:39.939 --> 01:13:41.860 at doctor Vashu's group. Today,
NOTE Confidence: 0.9197787

01:13:41.860 --> 01:13:43.220 I'm going to show you

NOTE Confidence: 0.9197787
01:13:43.220 --> 01:13:44.475 how to develop customized models
NOTE Confidence: 0.9197787
01:13:44.555 --> 01:13:46.314 customized models for some specific
NOTE Confidence: 0.9197787
01:13:46.314 --> 01:13:46.814 applications.
NOTE Confidence: 0.99367774
01:13:47.915 --> 01:13:49.354 So at the beginning, I
NOTE Confidence: 0.99367774
01:13:49.354 --> 01:13:50.795 would like to introduce why
NOTE Confidence: 0.99367774
01:13:50.795 --> 01:13:52.575 we need those customized models.
NOTE Confidence: 0.94804865
01:13:53.035 --> 01:13:54.075 We all know that large
NOTE Confidence: 0.94804865
01:13:54.075 --> 01:13:55.675 language models like LLAMA and
NOTE Confidence: 0.94804865
01:13:55.675 --> 01:13:57.915 GPT series have shown great
NOTE Confidence: 0.94804865
01:13:57.915 --> 01:13:59.295 potential in many domains
NOTE Confidence: 0.88620436
01:13:59.689 --> 01:14:00.429 as they,
NOTE Confidence: 0.85839736
01:14:00.729 --> 01:14:02.570 portrayed on large scale of
NOTE Confidence: 0.85839736
01:14:02.570 --> 01:14:04.510 text, and they have strong
NOTE Confidence: 0.85839736
01:14:04.650 --> 01:14:06.110 instruction following abilities
NOTE Confidence: 0.9812746
01:14:06.650 --> 01:14:08.590 across different tasks, and they
NOTE Confidence: 0.9812746

01:14:08.650 --> 01:14:10.090 have a wide coverage of
NOTE Confidence: 0.9812746

01:14:10.090 --> 01:14:10.989 general knowledge.
NOTE Confidence: 0.91490614

01:14:11.290 --> 01:14:13.130 However, they may not fully
NOTE Confidence: 0.91490614

01:14:13.130 --> 01:14:14.955 capture the analysis of some
NOTE Confidence: 0.91490614

01:14:14.955 --> 01:14:16.875 specific tasks tasks or user
NOTE Confidence: 0.91490614

01:14:16.875 --> 01:14:19.535 needs, especially when some, like,
NOTE Confidence: 0.9270753

01:14:20.155 --> 01:14:21.695 the task involves some,
NOTE Confidence: 0.9950713

01:14:22.075 --> 01:14:22.575 like,
NOTE Confidence: 0.9960558

01:14:23.035 --> 01:14:24.575 specific design definitions
NOTE Confidence: 0.92282975

01:14:25.595 --> 01:14:26.975 or the task is actually
NOTE Confidence: 0.92282975

01:14:27.275 --> 01:14:28.735 real in common users.
NOTE Confidence: 0.9985993

01:14:29.400 --> 01:14:30.760 And that's why we need
NOTE Confidence: 0.9985993

01:14:30.760 --> 01:14:32.760 to develop customized models for
NOTE Confidence: 0.9985993

01:14:32.760 --> 01:14:33.260 ourselves.
NOTE Confidence: 0.89043653

01:14:33.800 --> 01:14:35.160 We could in enhance the
NOTE Confidence: 0.89043653

01:14:35.160 --> 01:14:36.760 model with the domain specific

NOTE Confidence: 0.89043653
01:14:36.760 --> 01:14:37.260 expertise
NOTE Confidence: 0.9276926
01:14:37.800 --> 01:14:39.720 in this task and improve
NOTE Confidence: 0.9276926
01:14:39.720 --> 01:14:41.560 the performance of existing large
NOTE Confidence: 0.9276926
01:14:41.560 --> 01:14:42.300 range models.
NOTE Confidence: 0.9710523
01:14:43.005 --> 01:14:45.245 And in the process of
NOTE Confidence: 0.9710523
01:14:45.245 --> 01:14:47.245 improving the performance, we could
NOTE Confidence: 0.9710523
01:14:47.245 --> 01:14:49.165 actually, like, let the small
NOTE Confidence: 0.9710523
01:14:49.165 --> 01:14:51.185 size smaller size models to
NOTE Confidence: 0.9710523
01:14:51.405 --> 01:14:53.405 get comparable performance with those
NOTE Confidence: 0.9710523
01:14:53.405 --> 01:14:55.325 large larger size models, and
NOTE Confidence: 0.9710523
01:14:55.325 --> 01:14:56.604 we could get more, like,
NOTE Confidence: 0.9710523
01:14:56.604 --> 01:14:58.145 efficient and cost effective.
NOTE Confidence: 0.9484232
01:14:58.900 --> 01:15:00.820 And, also, we could better
NOTE Confidence: 0.9484232
01:15:00.820 --> 01:15:01.960 user experience
NOTE Confidence: 0.93871564
01:15:02.500 --> 01:15:04.500 be experienced by reducing some
NOTE Confidence: 0.93871564

01:15:04.500 --> 01:15:06.180 of the hallucinations in those
NOTE Confidence: 0.93871564

01:15:06.180 --> 01:15:07.720 existing large language models.
NOTE Confidence: 0.99890345

01:15:08.740 --> 01:15:09.940 And here are some key
NOTE Confidence: 0.99890345

01:15:09.940 --> 01:15:12.360 steps of developing customized models.
NOTE Confidence: 0.9528486

01:15:12.985 --> 01:15:14.344 So the first is to
NOTE Confidence: 0.9528486

01:15:14.344 --> 01:15:16.265 actually define your what is
NOTE Confidence: 0.9528486

01:15:16.265 --> 01:15:18.104 your NRP task, and the
NOTE Confidence: 0.9528486

01:15:18.104 --> 01:15:19.945 second is to prepare those
NOTE Confidence: 0.9528486

01:15:19.945 --> 01:15:20.844 data to,
NOTE Confidence: 0.8879828

01:15:21.465 --> 01:15:23.385 train and evaluate the large
NOTE Confidence: 0.8879828

01:15:23.385 --> 01:15:24.285 language model.
NOTE Confidence: 0.6367761

01:15:26.080 --> 01:15:26.580 Prevent,
NOTE Confidence: 0.84170234

01:15:26.960 --> 01:15:28.479 the prep the preparation of
NOTE Confidence: 0.84170234

01:15:28.479 --> 01:15:30.640 data involving some, like, steps
NOTE Confidence: 0.84170234

01:15:30.640 --> 01:15:32.660 that you have, like, introduced
NOTE Confidence: 0.84170234

01:15:32.720 --> 01:15:34.500 before do the data annotation

NOTE Confidence: 0.84170234
01:15:34.560 --> 01:15:36.479 and data pre processing to
NOTE Confidence: 0.84170234
01:15:36.479 --> 01:15:37.780 afford any further models.
NOTE Confidence: 0.9543318
01:15:38.355 --> 01:15:39.955 And after we get to
NOTE Confidence: 0.9543318
01:15:39.955 --> 01:15:41.235 the data, we could start
NOTE Confidence: 0.9543318
01:15:41.235 --> 01:15:42.935 model training is to enhance
NOTE Confidence: 0.9543318
01:15:42.995 --> 01:15:44.435 the performance of the models
NOTE Confidence: 0.9543318
01:15:44.435 --> 01:15:46.775 with those task specific data.
NOTE Confidence: 0.99334896
01:15:47.155 --> 01:15:47.895 And then
NOTE Confidence: 0.9851353
01:15:48.275 --> 01:15:50.035 once we finish the the
NOTE Confidence: 0.9851353
01:15:50.035 --> 01:15:51.635 model training, we could actually
NOTE Confidence: 0.9851353
01:15:51.635 --> 01:15:53.220 use another set of this
NOTE Confidence: 0.9851353
01:15:53.220 --> 01:15:55.220 annotated data to evaluate the
NOTE Confidence: 0.9851353
01:15:55.220 --> 01:15:57.000 performance of our developed
NOTE Confidence: 0.96866566
01:15:57.380 --> 01:15:59.060 model to see if the
NOTE Confidence: 0.96866566
01:15:59.060 --> 01:16:01.700 performance actually gained when compared
NOTE Confidence: 0.96866566

01:16:01.700 --> 01:16:03.000 with the backbone model.
NOTE Confidence: 0.9091143

01:16:03.300 --> 01:16:04.920 And then once we confirm
NOTE Confidence: 0.9091143

01:16:04.980 --> 01:16:06.500 that the performance of the
NOTE Confidence: 0.9091143

01:16:06.500 --> 01:16:07.000 model
NOTE Confidence: 0.88972765

01:16:07.305 --> 01:16:09.225 improved, we could actually use
NOTE Confidence: 0.88972765

01:16:09.225 --> 01:16:10.925 this model, do this customized
NOTE Confidence: 0.88972765

01:16:11.064 --> 01:16:12.765 models for those production.
NOTE Confidence: 0.9744394

01:16:13.945 --> 01:16:15.385 And here is a general
NOTE Confidence: 0.9744394

01:16:15.385 --> 01:16:17.064 workflow of model training and
NOTE Confidence: 0.9744394

01:16:17.064 --> 01:16:17.564 evaluation.
NOTE Confidence: 0.98981774

01:16:18.344 --> 01:16:19.864 Once we define our task
NOTE Confidence: 0.98981774

01:16:19.864 --> 01:16:21.660 and prepare our data, we
NOTE Confidence: 0.98981774

01:16:21.660 --> 01:16:23.020 need to, like, split data
NOTE Confidence: 0.98981774

01:16:23.020 --> 01:16:24.640 into different subsets.
NOTE Confidence: 0.9814004

01:16:25.500 --> 01:16:26.860 Usually, we would have three
NOTE Confidence: 0.9814004

01:16:26.860 --> 01:16:27.979 subsets. The first is the

NOTE Confidence: 0.9814004
01:16:27.979 --> 01:16:29.500 training data to develop the
NOTE Confidence: 0.9814004
01:16:29.500 --> 01:16:31.020 model, and the second would
NOTE Confidence: 0.9814004
01:16:31.020 --> 01:16:31.920 be the validation
NOTE Confidence: 0.99664706
01:16:32.460 --> 01:16:34.320 data to validate the effectiveness
NOTE Confidence: 0.96355575
01:16:34.699 --> 01:16:36.135 of the training model. But
NOTE Confidence: 0.96355575
01:16:36.135 --> 01:16:37.814 for simplification, we just use
NOTE Confidence: 0.96355575
01:16:37.814 --> 01:16:39.914 that data to, to instead,
NOTE Confidence: 0.96928746
01:16:40.854 --> 01:16:42.454 like, the we use the
NOTE Confidence: 0.96928746
01:16:42.454 --> 01:16:44.474 test data to, like, evaluate
NOTE Confidence: 0.96928746
01:16:44.534 --> 01:16:45.974 the training model. If the
NOTE Confidence: 0.96928746
01:16:45.974 --> 01:16:48.534 training model is, effective compared
NOTE Confidence: 0.96928746
01:16:48.534 --> 01:16:49.814 with the backbone model, we
NOTE Confidence: 0.96928746
01:16:49.814 --> 01:16:51.014 then could use it to
NOTE Confidence: 0.96928746
01:16:51.710 --> 01:16:53.469 for the production model to,
NOTE Confidence: 0.96928746
01:16:53.710 --> 01:16:55.810 like, process the production data.
NOTE Confidence: 0.96928746

01:16:55.949 --> 01:16:57.710 And if the training model
NOTE Confidence: 0.96928746

01:16:57.710 --> 01:16:59.889 actually the performance might decrease,
NOTE Confidence: 0.96928746

01:17:00.110 --> 01:17:01.389 once it's decreased, we might
NOTE Confidence: 0.96928746

01:17:01.389 --> 01:17:03.170 need to adjust the training
NOTE Confidence: 0.96928746

01:17:03.310 --> 01:17:03.810 process
NOTE Confidence: 0.9720229

01:17:04.270 --> 01:17:05.010 to, like,
NOTE Confidence: 0.8430713

01:17:06.195 --> 01:17:08.215 we redo the training part.
NOTE Confidence: 0.9939248

01:17:09.875 --> 01:17:10.375 And,
NOTE Confidence: 0.98590124

01:17:10.994 --> 01:17:12.275 next, I will show more
NOTE Confidence: 0.98590124

01:17:12.275 --> 01:17:13.955 details of each step. I
NOTE Confidence: 0.98590124

01:17:13.955 --> 01:17:15.875 will, start from how to
NOTE Confidence: 0.98590124

01:17:15.875 --> 01:17:16.695 define tasks.
NOTE Confidence: 0.8688981

01:17:17.870 --> 01:17:19.470 This is actually a real
NOTE Confidence: 0.8688981

01:17:19.470 --> 01:17:19.970 example.
NOTE Confidence: 0.9941807

01:17:20.350 --> 01:17:22.270 Start from, like, task task
NOTE Confidence: 0.9941807

01:17:22.270 --> 01:17:23.710 design. I will show you

NOTE Confidence: 0.9941807
01:17:23.710 --> 01:17:25.810 how to, like, develop customized
NOTE Confidence: 0.9941807
01:17:25.950 --> 01:17:27.410 models step by step.
NOTE Confidence: 0.9763521
01:17:27.950 --> 01:17:29.470 So let's say that we
NOTE Confidence: 0.9763521
01:17:29.470 --> 01:17:31.170 have a research to investigate
NOTE Confidence: 0.9927699
01:17:31.470 --> 01:17:33.535 about the impact of bilingual
NOTE Confidence: 0.99974
01:17:34.075 --> 01:17:34.575 bilingualism
NOTE Confidence: 0.9775998
01:17:34.955 --> 01:17:36.655 and ADRD pro progression.
NOTE Confidence: 0.99582213
01:17:37.995 --> 01:17:39.375 So the first step
NOTE Confidence: 0.9587607
01:17:39.755 --> 01:17:41.595 is, like, in different clinical
NOTE Confidence: 0.9587607
01:17:41.595 --> 01:17:43.275 researches, is to find eligible
NOTE Confidence: 0.9587607
01:17:43.275 --> 01:17:43.775 patients.
NOTE Confidence: 0.99680185
01:17:44.409 --> 01:17:45.530 But once we find those
NOTE Confidence: 0.99680185
01:17:45.530 --> 01:17:47.289 eligible patients, we need to
NOTE Confidence: 0.99680185
01:17:47.289 --> 01:17:49.150 identify those bilingual or monolingual
NOTE Confidence: 0.99680185
01:17:49.210 --> 01:17:51.389 patients within these patients.
NOTE Confidence: 0.99623895

01:17:52.090 --> 01:17:52.590 And,
NOTE Confidence: 0.9545582

01:17:53.130 --> 01:17:54.750 the first thought is to,
NOTE Confidence: 0.9545582

01:17:54.889 --> 01:17:57.354 like, check the structured data
NOTE Confidence: 0.9545582

01:17:57.415 --> 01:17:59.495 about those, preferred language or
NOTE Confidence: 0.9545582

01:17:59.495 --> 01:18:00.314 written language,
NOTE Confidence: 0.8743466

01:18:01.415 --> 01:18:02.775 areas to see what the
NOTE Confidence: 0.8743466

01:18:02.775 --> 01:18:04.795 language the patient will prefer.
NOTE Confidence: 0.9460714

01:18:05.175 --> 01:18:07.015 But where where we, like,
NOTE Confidence: 0.9460714

01:18:07.015 --> 01:18:08.455 check the actual data, we
NOTE Confidence: 0.9460714

01:18:08.455 --> 01:18:10.375 find that the structured data
NOTE Confidence: 0.9460714

01:18:10.375 --> 01:18:10.875 actually,
NOTE Confidence: 0.93054295

01:18:11.189 --> 01:18:13.110 the there might not be
NOTE Confidence: 0.93054295

01:18:13.110 --> 01:18:14.550 enough this kind of data
NOTE Confidence: 0.93054295

01:18:14.550 --> 01:18:16.010 to support our research,
NOTE Confidence: 0.9969899

01:18:16.310 --> 01:18:17.510 and some of them might
NOTE Confidence: 0.9969899

01:18:17.510 --> 01:18:19.130 not be even accurate.

NOTE Confidence: 0.9807402
01:18:19.670 --> 01:18:21.510 But we noticed that there
NOTE Confidence: 0.9807402
01:18:21.510 --> 01:18:22.810 are a lot of language
NOTE Confidence: 0.98468
01:18:23.189 --> 01:18:25.130 information contained in the clinical
NOTE Confidence: 0.98468
01:18:25.189 --> 01:18:26.725 notes. For example, many notes
NOTE Confidence: 0.98468
01:18:26.725 --> 01:18:28.345 would record what the patient,
NOTE Confidence: 0.9886432
01:18:28.885 --> 01:18:30.165 speak, what is the preferred
NOTE Confidence: 0.9886432
01:18:30.165 --> 01:18:31.525 language, and how well do
NOTE Confidence: 0.9886432
01:18:31.525 --> 01:18:33.864 they speak. So we might,
NOTE Confidence: 0.98958015
01:18:34.165 --> 01:18:35.064 like, comprehensively
NOTE Confidence: 0.95369023
01:18:36.005 --> 01:18:37.765 extract all the language speaking
NOTE Confidence: 0.95369023
01:18:37.765 --> 01:18:39.525 status from all the clinical
NOTE Confidence: 0.95369023
01:18:39.525 --> 01:18:41.989 notes, like using those OLP,
NOTE Confidence: 0.970317
01:18:42.530 --> 01:18:43.030 models.
NOTE Confidence: 0.9588801
01:18:44.130 --> 01:18:45.889 There are two targets that
NOTE Confidence: 0.9588801
01:18:45.889 --> 01:18:47.170 we want to extract. The
NOTE Confidence: 0.9588801

01:18:47.170 --> 01:18:48.610 first is what language does
NOTE Confidence: 0.9588801

01:18:48.610 --> 01:18:50.050 the patient speak and how
NOTE Confidence: 0.9588801

01:18:50.050 --> 01:18:51.889 well do they speak. So,
NOTE Confidence: 0.9588801

01:18:52.210 --> 01:18:53.189 for these two
NOTE Confidence: 0.9673767

01:18:53.604 --> 01:18:54.485 specific tasks,
NOTE Confidence: 0.96507186

01:18:55.285 --> 01:18:56.505 aims, we, like,
NOTE Confidence: 0.90667534

01:18:56.805 --> 01:18:58.645 could formulate the task as
NOTE Confidence: 0.90667534

01:18:58.645 --> 01:18:59.305 a task.
NOTE Confidence: 0.9574227

01:19:00.564 --> 01:19:01.925 The first the first thing
NOTE Confidence: 0.9574227

01:19:01.925 --> 01:19:02.965 we want to do is
NOTE Confidence: 0.9574227

01:19:02.965 --> 01:19:04.564 to identify all the language
NOTE Confidence: 0.9574227

01:19:04.564 --> 01:19:06.240 entities in the clinical notes,
NOTE Confidence: 0.9574227

01:19:06.480 --> 01:19:07.920 and then we could assign
NOTE Confidence: 0.9574227

01:19:07.920 --> 01:19:10.000 different tags based on different
NOTE Confidence: 0.9574227

01:19:10.000 --> 01:19:10.500 context
NOTE Confidence: 0.9758833

01:19:10.800 --> 01:19:12.980 to indicate different speaking status

NOTE Confidence: 0.9758833
01:19:13.040 --> 01:19:14.340 of the patient.
NOTE Confidence: 0.8291229
01:19:16.080 --> 01:19:18.100 Once we, like, formulate the
NOTE Confidence: 0.93445116
01:19:18.400 --> 01:19:20.395 task as a task, we
NOTE Confidence: 0.93445116
01:19:20.395 --> 01:19:22.155 need to further refine the
NOTE Confidence: 0.93445116
01:19:22.155 --> 01:19:23.775 details of the task.
NOTE Confidence: 0.94660175
01:19:24.075 --> 01:19:25.835 We might need to review
NOTE Confidence: 0.94660175
01:19:25.835 --> 01:19:27.435 some of the clinical notes
NOTE Confidence: 0.94660175
01:19:27.435 --> 01:19:29.755 and design different tasks for
NOTE Confidence: 0.94660175
01:19:29.755 --> 01:19:30.575 the task.
NOTE Confidence: 0.96578884
01:19:31.755 --> 01:19:33.675 In this, like, data reviewing
NOTE Confidence: 0.96578884
01:19:33.915 --> 01:19:35.195 after the data review, we,
NOTE Confidence: 0.91782933
01:19:36.270 --> 01:19:38.030 designed four different tasks for
NOTE Confidence: 0.91782933
01:19:38.030 --> 01:19:39.310 this task. The first is
NOTE Confidence: 0.91782933
01:19:39.310 --> 01:19:40.990 language fluent, which indicate the
NOTE Confidence: 0.91782933
01:19:40.990 --> 01:19:42.430 patient speaks some kind of
NOTE Confidence: 0.91782933

01:19:42.430 --> 01:19:43.330 language fluently.
NOTE Confidence: 0.91600287

01:19:43.950 --> 01:19:45.890 And there is another language
NOTE Confidence: 0.91600287

01:19:45.950 --> 01:19:47.310 sum to indicate the patient
NOTE Confidence: 0.91600287

01:19:47.310 --> 01:19:48.770 speaks some of the languages,
NOTE Confidence: 0.91600287

01:19:49.070 --> 01:19:50.110 and that they are language
NOTE Confidence: 0.91600287

01:19:50.110 --> 01:19:51.570 no and the language other.
NOTE Confidence: 0.94294953

01:19:52.405 --> 01:19:54.104 And here are some examples.
NOTE Confidence: 0.94294953

01:19:54.245 --> 01:19:55.845 For, for the first one,
NOTE Confidence: 0.94294953

01:19:55.845 --> 01:19:56.745 language fluent,
NOTE Confidence: 0.9788008

01:19:57.445 --> 01:19:58.965 some of the sentences says
NOTE Confidence: 0.9788008

01:19:58.965 --> 01:20:01.385 that the patient speaks Italian
NOTE Confidence: 0.6839928

01:20:01.925 --> 01:20:02.425 primary
NOTE Confidence: 0.6515143

01:20:02.805 --> 01:20:03.305 primarily.
NOTE Confidence: 0.93567795

01:20:04.510 --> 01:20:06.189 So this indicate that the
NOTE Confidence: 0.93567795

01:20:06.189 --> 01:20:08.530 person is has lang fluent
NOTE Confidence: 0.93567795

01:20:08.590 --> 01:20:09.090 Italian

NOTE Confidence: 0.9691571
01:20:09.790 --> 01:20:10.290 abilities.
NOTE Confidence: 0.89324147
01:20:10.590 --> 01:20:12.270 And for the language some,
NOTE Confidence: 0.89324147
01:20:12.270 --> 01:20:14.130 here's the sentence. She speaks
NOTE Confidence: 0.89324147
01:20:14.189 --> 01:20:14.930 some English.
NOTE Confidence: 0.96928227
01:20:15.390 --> 01:20:17.150 And, no, the patient does
NOTE Confidence: 0.96928227
01:20:17.150 --> 01:20:19.125 not speak English. And for
NOTE Confidence: 0.96928227
01:20:19.125 --> 01:20:20.824 the language other, we actually,
NOTE Confidence: 0.8849359
01:20:22.085 --> 01:20:23.764 when when we reviewing data,
NOTE Confidence: 0.8849359
01:20:23.764 --> 01:20:25.045 we found a lot of,
NOTE Confidence: 0.8849359
01:20:25.045 --> 01:20:27.145 like, languages used in other
NOTE Confidence: 0.9885655
01:20:27.525 --> 01:20:29.784 individuals. For example, the patient's
NOTE Confidence: 0.9885655
01:20:29.844 --> 01:20:31.385 family or the
NOTE Confidence: 0.9569616
01:20:31.710 --> 01:20:33.630 written language. So this kind
NOTE Confidence: 0.9569616
01:20:33.630 --> 01:20:35.810 of language do not, indicate
NOTE Confidence: 0.9569616
01:20:35.870 --> 01:20:37.630 the language speaking status of
NOTE Confidence: 0.9569616

01:20:37.630 --> 01:20:39.250 the patient, so we categorize
NOTE Confidence: 0.9569616

01:20:39.390 --> 01:20:40.850 them as language other.
NOTE Confidence: 0.9730098

01:20:41.870 --> 01:20:43.810 Once we refine the details
NOTE Confidence: 0.9730098

01:20:43.870 --> 01:20:45.310 of the EER task, we
NOTE Confidence: 0.9730098

01:20:45.310 --> 01:20:47.354 need to, like, start to
NOTE Confidence: 0.9730098

01:20:47.354 --> 01:20:48.795 prepare the data for model
NOTE Confidence: 0.9730098

01:20:48.795 --> 01:20:50.335 training and the model evaluation.
NOTE Confidence: 0.9894747

01:20:51.915 --> 01:20:53.775 Here is the overall flow
NOTE Confidence: 0.9894747

01:20:53.835 --> 01:20:55.455 to prepare the data for,
NOTE Confidence: 0.9894747

01:20:55.515 --> 01:20:56.735 like, model developing.
NOTE Confidence: 0.99468017

01:20:57.755 --> 01:20:59.115 We first need to get
NOTE Confidence: 0.99468017

01:20:59.115 --> 01:21:00.575 some raw data and
NOTE Confidence: 0.97104883

01:21:01.189 --> 01:21:03.030 do the annotation with the
NOTE Confidence: 0.97104883

01:21:03.030 --> 01:21:05.050 annotation guideline that we developed
NOTE Confidence: 0.97104883

01:21:05.110 --> 01:21:07.429 before. And once we get
NOTE Confidence: 0.97104883

01:21:07.429 --> 01:21:09.189 the annotated results, we may

NOTE Confidence: 0.97104883
01:21:09.189 --> 01:21:10.469 need to, like, design a
NOTE Confidence: 0.97104883
01:21:10.469 --> 01:21:11.929 prompt for this task.
NOTE Confidence: 0.96341604
01:21:12.469 --> 01:21:13.670 Some of the prompt has
NOTE Confidence: 0.96341604
01:21:13.670 --> 01:21:14.790 been, like, discussed,
NOTE Confidence: 0.8052125
01:21:15.755 --> 01:21:17.375 by doctor Shui and Vipina.
NOTE Confidence: 0.96610016
01:21:18.155 --> 01:21:19.434 And once we get the
NOTE Confidence: 0.96610016
01:21:19.434 --> 01:21:20.814 prompts and all the annotated
NOTE Confidence: 0.96610016
01:21:20.875 --> 01:21:22.155 results, we need to process
NOTE Confidence: 0.96610016
01:21:22.155 --> 01:21:23.755 the results for the models
NOTE Confidence: 0.96610016
01:21:23.755 --> 01:21:24.875 to load to start the
NOTE Confidence: 0.96610016
01:21:24.875 --> 01:21:25.934 training and evaluation.
NOTE Confidence: 0.9943743
01:21:27.770 --> 01:21:28.250 And,
NOTE Confidence: 0.8964094
01:21:28.650 --> 01:21:29.550 this is the annotation
NOTE Confidence: 0.9282996
01:21:30.010 --> 01:21:31.689 using Blue that Yuja has
NOTE Confidence: 0.9282996
01:21:31.689 --> 01:21:33.050 mentioned before, so I'm just
NOTE Confidence: 0.9282996

01:21:33.050 --> 01:21:35.130 gonna skip this. And here
NOTE Confidence: 0.9282996

01:21:35.130 --> 01:21:36.810 is the annotated results look
NOTE Confidence: 0.9282996

01:21:36.810 --> 01:21:38.769 like. Usually, we would have
NOTE Confidence: 0.9282996

01:21:38.769 --> 01:21:40.489 a JSON file for each
NOTE Confidence: 0.9282996

01:21:40.489 --> 01:21:42.670 input sample, and each
NOTE Confidence: 0.9364001

01:21:43.045 --> 01:21:44.564 each sample looks like this.
NOTE Confidence: 0.9364001

01:21:44.804 --> 01:21:46.564 So it will record file
NOTE Confidence: 0.9364001

01:21:46.564 --> 01:21:48.664 name and the original sentence
NOTE Confidence: 0.9364001

01:21:48.724 --> 01:21:50.885 and also the entities and
NOTE Confidence: 0.9364001

01:21:50.885 --> 01:21:52.564 the positions of entities that
NOTE Confidence: 0.9364001

01:21:52.564 --> 01:21:53.045 we,
NOTE Confidence: 0.97788703

01:21:54.005 --> 01:21:54.505 marked.
NOTE Confidence: 0.9883942

01:21:56.040 --> 01:21:57.420 And here is the prompt.
NOTE Confidence: 0.9883942

01:21:57.560 --> 01:21:59.580 I'm just gonna skip this.
NOTE Confidence: 0.9883942

01:21:59.800 --> 01:22:00.300 And
NOTE Confidence: 0.9955813

01:22:00.920 --> 01:22:02.600 after we get all the

NOTE Confidence: 0.9955813
01:22:02.600 --> 01:22:04.120 files and the prompt, we
NOTE Confidence: 0.9955813
01:22:04.120 --> 01:22:05.240 need to process the data
NOTE Confidence: 0.9955813
01:22:05.240 --> 01:22:07.000 based on different prompts. For
NOTE Confidence: 0.9955813
01:22:07.000 --> 01:22:07.500 example,
NOTE Confidence: 0.9780505
01:22:08.135 --> 01:22:09.995 our task is to, like,
NOTE Confidence: 0.9780505
01:22:10.055 --> 01:22:12.215 annotate all the text in
NOTE Confidence: 0.9780505
01:22:12.215 --> 01:22:14.555 the original sentence with, HTML
NOTE Confidence: 0.936676
01:22:14.935 --> 01:22:16.775 tag. So we might need
NOTE Confidence: 0.936676
01:22:16.775 --> 01:22:18.535 to process the input as
NOTE Confidence: 0.936676
01:22:18.535 --> 01:22:20.295 the original sentence, and the
NOTE Confidence: 0.936676
01:22:20.295 --> 01:22:22.295 output and the target output
NOTE Confidence: 0.936676
01:22:22.295 --> 01:22:22.775 would be,
NOTE Confidence: 0.9617362
01:22:24.030 --> 01:22:25.889 the same sentence, but with
NOTE Confidence: 0.9617362
01:22:25.949 --> 01:22:26.929 all the entities
NOTE Confidence: 0.9393909
01:22:27.469 --> 01:22:29.010 wrapped with this, like,
NOTE Confidence: 0.8737771

01:22:30.429 --> 01:22:31.969 a language front or language
NOTE Confidence: 0.8737771

01:22:32.030 --> 01:22:32.989 sum, those tags,
NOTE Confidence: 0.96883994

01:22:33.869 --> 01:22:35.250 to wrap it with HTML
NOTE Confidence: 0.96883994

01:22:35.469 --> 01:22:35.969 tag.
NOTE Confidence: 0.9640541

01:22:36.995 --> 01:22:38.515 And this is also the
NOTE Confidence: 0.9640541

01:22:38.515 --> 01:22:40.675 data preparation for the models
NOTE Confidence: 0.9640541

01:22:40.675 --> 01:22:42.035 to load. I will prepare
NOTE Confidence: 0.9640541

01:22:42.354 --> 01:22:43.795 I will I will show
NOTE Confidence: 0.9640541

01:22:43.795 --> 01:22:45.235 the code, later so you
NOTE Confidence: 0.9640541

01:22:45.235 --> 01:22:46.515 can just directly try the
NOTE Confidence: 0.9640541

01:22:46.515 --> 01:22:48.295 code to to process it.
NOTE Confidence: 0.99248844

01:22:49.080 --> 01:22:51.080 So after we prepare the
NOTE Confidence: 0.99248844

01:22:51.080 --> 01:22:52.460 data, we now finally
NOTE Confidence: 0.91558474

01:22:53.960 --> 01:22:55.400 can start the fine tuning
NOTE Confidence: 0.91558474

01:22:55.400 --> 01:22:55.900 process.
NOTE Confidence: 0.9909002

01:22:57.800 --> 01:22:59.180 So fine tuning

NOTE Confidence: 0.9938337
01:22:59.560 --> 01:23:00.780 process is actually
NOTE Confidence: 0.959562
01:23:01.240 --> 01:23:03.020 a process to adjust the
NOTE Confidence: 0.9244277
01:23:03.354 --> 01:23:04.795 weights of the large language
NOTE Confidence: 0.9244277
01:23:04.795 --> 01:23:06.875 models to make it adapt
NOTE Confidence: 0.9244277
01:23:06.875 --> 01:23:09.695 to our task specific data.
NOTE Confidence: 0.99752694
01:23:10.074 --> 01:23:12.074 So, usually, we need to,
NOTE Confidence: 0.99752694
01:23:12.074 --> 01:23:12.574 like,
NOTE Confidence: 0.9793553
01:23:12.875 --> 01:23:14.235 adjust all the weights of
NOTE Confidence: 0.9793553
01:23:14.235 --> 01:23:15.750 the model, but like, we
NOTE Confidence: 0.9793553
01:23:15.750 --> 01:23:16.950 know that for large language
NOTE Confidence: 0.9793553
01:23:16.950 --> 01:23:17.910 model, there are a lot
NOTE Confidence: 0.9793553
01:23:17.910 --> 01:23:19.290 of parameters. So,
NOTE Confidence: 0.9554377
01:23:19.990 --> 01:23:21.270 the full fine tuning would
NOTE Confidence: 0.9554377
01:23:21.270 --> 01:23:22.570 be very, like,
NOTE Confidence: 0.91050994
01:23:23.110 --> 01:23:23.610 computational
NOTE Confidence: 0.94671893

01:23:24.390 --> 01:23:25.830 cost cost will be very
NOTE Confidence: 0.94671893

01:23:25.830 --> 01:23:27.910 high. So instead, we would
NOTE Confidence: 0.94671893

01:23:27.910 --> 01:23:28.410 use,
NOTE Confidence: 0.9892634

01:23:28.790 --> 01:23:29.290 like,
NOTE Confidence: 0.906847

01:23:29.595 --> 01:23:31.614 widely used way, LoRa, to
NOTE Confidence: 0.906847

01:23:31.835 --> 01:23:33.435 do the fine tuning. LoRa
NOTE Confidence: 0.906847

01:23:33.435 --> 01:23:34.095 is actually,
NOTE Confidence: 0.99097186

01:23:34.635 --> 01:23:36.015 a low rank adaptation
NOTE Confidence: 0.9307817

01:23:36.475 --> 01:23:38.575 that use, two, like, different,
NOTE Confidence: 0.99631524

01:23:39.435 --> 01:23:41.755 small vectors here in the
NOTE Confidence: 0.99631524

01:23:41.755 --> 01:23:42.255 green
NOTE Confidence: 0.9600558

01:23:42.555 --> 01:23:43.055 green
NOTE Confidence: 0.97866243

01:23:44.270 --> 01:23:45.170 in the green,
NOTE Confidence: 0.95588195

01:23:46.430 --> 01:23:47.570 in the green part.
NOTE Confidence: 0.91337526

01:23:48.350 --> 01:23:50.110 Instead of fine tuning the
NOTE Confidence: 0.91337526

01:23:50.110 --> 01:23:52.030 entire large network model, we

NOTE Confidence: 0.91337526
01:23:52.030 --> 01:23:53.550 we only need to adjust
NOTE Confidence: 0.91337526
01:23:53.550 --> 01:23:55.310 the parameters in this, like,
NOTE Confidence: 0.91337526
01:23:55.310 --> 01:23:56.130 small vectors.
NOTE Confidence: 0.9623676
01:23:56.535 --> 01:23:58.695 So, compared with full fine
NOTE Confidence: 0.9623676
01:23:58.695 --> 01:24:00.135 tuning, it is much more
NOTE Confidence: 0.9623676
01:24:00.135 --> 01:24:01.814 faster, and we only need
NOTE Confidence: 0.9623676
01:24:01.814 --> 01:24:03.114 minimal training resources.
NOTE Confidence: 0.94789445
01:24:03.735 --> 01:24:05.335 And, but it it needs
NOTE Confidence: 0.94789445
01:24:05.335 --> 01:24:06.475 high quality datasets.
NOTE Confidence: 0.998497
01:24:07.015 --> 01:24:08.215 So we need to, like,
NOTE Confidence: 0.998497
01:24:08.215 --> 01:24:09.675 define the task and
NOTE Confidence: 0.9629548
01:24:10.270 --> 01:24:12.449 develop the annotation guideline carefully.
NOTE Confidence: 0.9421113
01:24:13.389 --> 01:24:14.909 And it also has some
NOTE Confidence: 0.9421113
01:24:14.909 --> 01:24:16.369 high risk of over fifty.
NOTE Confidence: 0.9669823
01:24:16.989 --> 01:24:18.290 So for the resources,
NOTE Confidence: 0.9804156

01:24:18.750 --> 01:24:20.750 for eight eight billion model,
NOTE Confidence: 0.9804156

01:24:20.750 --> 01:24:21.949 it might need one a
NOTE Confidence: 0.9804156

01:24:21.949 --> 01:24:22.989 one hundred or h one
NOTE Confidence: 0.9804156

01:24:22.989 --> 01:24:24.670 hundred GPU. But for seventy
NOTE Confidence: 0.9804156

01:24:24.670 --> 01:24:26.085 billion models, we might need
NOTE Confidence: 0.9804156

01:24:26.085 --> 01:24:27.705 two h one hundred GPUs
NOTE Confidence: 0.9804156

01:24:27.844 --> 01:24:28.905 to do the fine tuning.
NOTE Confidence: 0.9763425

01:24:30.165 --> 01:24:32.325 So, here is some environment
NOTE Confidence: 0.9763425

01:24:32.325 --> 01:24:34.005 setup. I also provided the
NOTE Confidence: 0.9763425

01:24:34.005 --> 01:24:35.145 code in the,
NOTE Confidence: 0.9485906

01:24:35.685 --> 01:24:37.205 at last so you can
NOTE Confidence: 0.9485906

01:24:37.205 --> 01:24:38.665 check it for more details.
NOTE Confidence: 0.9989551

01:24:40.080 --> 01:24:41.360 And if we want to
NOTE Confidence: 0.9989551

01:24:41.360 --> 01:24:43.280 do the fine tuning, we
NOTE Confidence: 0.9989551

01:24:43.280 --> 01:24:44.340 need to, like,
NOTE Confidence: 0.7491257

01:24:44.800 --> 01:24:45.300 modify

NOTE Confidence: 0.960806
01:24:45.680 --> 01:24:47.600 the config file in the
NOTE Confidence: 0.960806
01:24:47.600 --> 01:24:48.820 code that I provided.
NOTE Confidence: 0.97095335
01:24:49.920 --> 01:24:51.439 The first one is to,
NOTE Confidence: 0.97095335
01:24:51.439 --> 01:24:53.040 like, indicate where is the
NOTE Confidence: 0.97095335
01:24:53.040 --> 01:24:53.540 model
NOTE Confidence: 0.93771785
01:24:53.925 --> 01:24:55.445 stored in the Camino or
NOTE Confidence: 0.93771785
01:24:55.445 --> 01:24:56.505 CHP environment.
NOTE Confidence: 0.9992549
01:24:56.885 --> 01:24:57.385 So
NOTE Confidence: 0.98880905
01:24:57.765 --> 01:24:59.045 in the folder, it should
NOTE Confidence: 0.98880905
01:24:59.045 --> 01:25:00.325 look like this. It has,
NOTE Confidence: 0.98880905
01:25:00.645 --> 01:25:01.625 many, like,
NOTE Confidence: 0.98070705
01:25:02.085 --> 01:25:03.365 weights of the model, some
NOTE Confidence: 0.98070705
01:25:03.365 --> 01:25:04.425 details of the model.
NOTE Confidence: 0.93422216
01:25:04.885 --> 01:25:06.645 And for the data, we
NOTE Confidence: 0.93422216
01:25:06.645 --> 01:25:08.005 also need to provide the
NOTE Confidence: 0.93422216

01:25:08.005 --> 01:25:08.910 the pass of the data
NOTE Confidence: 0.93422216

01:25:08.910 --> 01:25:09.729 that we
NOTE Confidence: 0.94478625

01:25:10.270 --> 01:25:11.950 processed before to tell the
NOTE Confidence: 0.94478625

01:25:11.950 --> 01:25:13.250 model where is the data.
NOTE Confidence: 0.9989673

01:25:13.710 --> 01:25:14.210 And
NOTE Confidence: 0.96566683

01:25:14.510 --> 01:25:15.790 beside the model and the
NOTE Confidence: 0.96566683

01:25:15.790 --> 01:25:16.990 data, we also need to
NOTE Confidence: 0.96566683

01:25:16.990 --> 01:25:19.010 set up some other, like,
NOTE Confidence: 0.9954912

01:25:19.470 --> 01:25:21.390 configs. For example, the most
NOTE Confidence: 0.9954912

01:25:21.390 --> 01:25:22.750 important one might be the
NOTE Confidence: 0.9954912

01:25:22.750 --> 01:25:24.614 learning rate. You could, like,
NOTE Confidence: 0.9954912

01:25:24.614 --> 01:25:26.375 adjust the learning rate based
NOTE Confidence: 0.9954912

01:25:26.375 --> 01:25:26.875 on
NOTE Confidence: 0.99796

01:25:27.255 --> 01:25:29.434 the evaluation results of the,
NOTE Confidence: 0.7313857

01:25:29.814 --> 01:25:30.715 trained model.
NOTE Confidence: 0.9385293

01:25:32.375 --> 01:25:34.534 And finally, we need we

NOTE Confidence: 0.9385293
01:25:34.534 --> 01:25:36.314 can we can start the
NOTE Confidence: 0.9385293
01:25:36.375 --> 01:25:37.835 model's fine tuning.
NOTE Confidence: 0.97857505
01:25:38.420 --> 01:25:39.619 The fine tuning process is
NOTE Confidence: 0.97857505
01:25:39.619 --> 01:25:41.320 actually very easy. Once we,
NOTE Confidence: 0.95994025
01:25:41.860 --> 01:25:43.460 finish the config file, we
NOTE Confidence: 0.95994025
01:25:43.460 --> 01:25:45.139 can just start the fine
NOTE Confidence: 0.95994025
01:25:45.139 --> 01:25:46.580 tuning with only one line
NOTE Confidence: 0.95994025
01:25:46.580 --> 01:25:47.320 of a command.
NOTE Confidence: 0.9659059
01:25:47.780 --> 01:25:49.460 So after fine tuning, we
NOTE Confidence: 0.9659059
01:25:49.460 --> 01:25:51.400 will get adapt adaptor
NOTE Confidence: 0.62211555
01:25:52.314 --> 01:25:52.814 adapter
NOTE Confidence: 0.9915429
01:25:53.675 --> 01:25:55.435 parameters, which is a very
NOTE Confidence: 0.9915429
01:25:55.435 --> 01:25:56.175 small file.
NOTE Confidence: 0.9712057
01:25:56.555 --> 01:25:57.995 So after we get this
NOTE Confidence: 0.9712057
01:25:57.995 --> 01:25:59.675 LoRa adapter, we need to
NOTE Confidence: 0.9712057

01:25:59.675 --> 01:26:01.275 combine the adapter with the
NOTE Confidence: 0.9712057

01:26:01.275 --> 01:26:03.375 original backbone model to formalize
NOTE Confidence: 0.9980898

01:26:03.755 --> 01:26:05.775 your own customized model.
NOTE Confidence: 0.9953414

01:26:06.770 --> 01:26:08.550 Once we get our customized
NOTE Confidence: 0.9953414

01:26:08.610 --> 01:26:09.910 model, we need to,
NOTE Confidence: 0.9539719

01:26:10.290 --> 01:26:11.729 test the model to see
NOTE Confidence: 0.9539719

01:26:11.729 --> 01:26:13.590 if the performance actually gained
NOTE Confidence: 0.9539719

01:26:13.650 --> 01:26:15.670 compared with the backbone model.
NOTE Confidence: 0.9539719

01:26:15.729 --> 01:26:16.930 So we need to do
NOTE Confidence: 0.9539719

01:26:16.930 --> 01:26:18.290 the inference on the test
NOTE Confidence: 0.9539719

01:26:18.290 --> 01:26:18.790 data.
NOTE Confidence: 0.938043

01:26:19.395 --> 01:26:20.914 And here is also some
NOTE Confidence: 0.938043

01:26:20.914 --> 01:26:22.514 example how to setting up
NOTE Confidence: 0.938043

01:26:22.514 --> 01:26:23.335 all the environment.
NOTE Confidence: 0.92958814

01:26:23.795 --> 01:26:24.934 And this is the example
NOTE Confidence: 0.92958814

01:26:24.994 --> 01:26:26.914 of how to, like, do

NOTE Confidence: 0.92958814
01:26:26.914 --> 01:26:28.354 the inference on the test
NOTE Confidence: 0.92958814
01:26:28.354 --> 01:26:29.474 data data to get to
NOTE Confidence: 0.92958814
01:26:29.474 --> 01:26:30.054 the results.
NOTE Confidence: 0.9676111
01:26:32.460 --> 01:26:33.660 This is also to set
NOTE Confidence: 0.9676111
01:26:33.660 --> 01:26:35.040 up all the inference,
NOTE Confidence: 0.609944
01:26:36.300 --> 01:26:37.040 from fix.
NOTE Confidence: 0.94219714
01:26:37.820 --> 01:26:39.260 For example, the max max
NOTE Confidence: 0.94219714
01:26:39.260 --> 01:26:41.680 token indicate how long you'd
NOTE Confidence: 0.94219714
01:26:41.740 --> 01:26:43.760 expect the model to output,
NOTE Confidence: 0.94219714
01:26:43.820 --> 01:26:46.265 and the stop token EOS
NOTE Confidence: 0.98197305
01:26:46.725 --> 01:26:48.745 means once the model, like,
NOTE Confidence: 0.81300694
01:26:50.165 --> 01:26:51.685 generated the US to token,
NOTE Confidence: 0.81300694
01:26:51.685 --> 01:26:52.745 it would finish
NOTE Confidence: 0.9980465
01:26:53.285 --> 01:26:54.025 the generation
NOTE Confidence: 0.976923
01:26:54.485 --> 01:26:56.245 instead of generate, you know,
NOTE Confidence: 0.976923

01:26:56.245 --> 01:26:57.385 five hundred tokens.
NOTE Confidence: 0.9607987

01:26:58.600 --> 01:27:00.540 So once we get the
NOTE Confidence: 0.9607987

01:27:00.600 --> 01:27:02.940 inference results, we could evaluate
NOTE Confidence: 0.9607987

01:27:03.000 --> 01:27:03.580 the performance
NOTE Confidence: 0.9320425

01:27:03.880 --> 01:27:05.560 of the model and compare
NOTE Confidence: 0.9320425

01:27:05.560 --> 01:27:06.840 it with the performance of
NOTE Confidence: 0.9320425

01:27:06.840 --> 01:27:07.979 the backbone model.
NOTE Confidence: 0.93886817

01:27:08.280 --> 01:27:09.820 And here is some evaluation
NOTE Confidence: 0.93886817

01:27:09.960 --> 01:27:11.235 metric that you just have
NOTE Confidence: 0.93886817

01:27:11.235 --> 01:27:12.915 the introduced before, so I'm
NOTE Confidence: 0.93886817

01:27:12.915 --> 01:27:14.295 just gonna skip this.
NOTE Confidence: 0.99826854

01:27:14.675 --> 01:27:16.034 And I also provide some
NOTE Confidence: 0.99826854

01:27:16.034 --> 01:27:16.534 scripts
NOTE Confidence: 0.9991893

01:27:16.835 --> 01:27:17.574 for the,
NOTE Confidence: 0.99006045

01:27:17.955 --> 01:27:19.635 like, evaluation. You can also
NOTE Confidence: 0.99006045

01:27:19.635 --> 01:27:20.675 refer to the code that

NOTE Confidence: 0.99006045
01:27:20.675 --> 01:27:22.295 I provided for more details.
NOTE Confidence: 0.9579332
01:27:23.155 --> 01:27:24.205 And here is the,
NOTE Confidence: 0.96793836
01:27:25.430 --> 01:27:26.950 fine tune results that we
NOTE Confidence: 0.96793836
01:27:26.950 --> 01:27:29.590 have once after we do
NOTE Confidence: 0.96793836
01:27:29.590 --> 01:27:30.790 the fine tuning with eight
NOTE Confidence: 0.96793836
01:27:30.790 --> 01:27:32.810 hundred samples that we annotated
NOTE Confidence: 0.96793836
01:27:32.870 --> 01:27:33.370 before.
NOTE Confidence: 0.7531381
01:27:33.990 --> 01:27:35.030 So the fine tune the
NOTE Confidence: 0.7531381
01:27:35.030 --> 01:27:36.170 means that,
NOTE Confidence: 0.9470556
01:27:36.470 --> 01:27:38.310 we use the three seventy
NOTE Confidence: 0.9470556
01:27:38.310 --> 01:27:40.435 billion instructor model as a
NOTE Confidence: 0.9470556
01:27:40.435 --> 01:27:41.555 backbone model to do the
NOTE Confidence: 0.9470556
01:27:41.555 --> 01:27:42.295 fine tune.
NOTE Confidence: 0.9900349
01:27:42.595 --> 01:27:44.115 So compared with the backbone
NOTE Confidence: 0.9900349
01:27:44.115 --> 01:27:45.575 model, we see that for
NOTE Confidence: 0.8697076

01:27:45.955 --> 01:27:47.795 every tech language, fluent, stem,
NOTE Confidence: 0.8697076

01:27:47.795 --> 01:27:49.155 no, and other, all the
NOTE Confidence: 0.8697076

01:27:49.155 --> 01:27:50.595 performance of all the f
NOTE Confidence: 0.8697076

01:27:50.595 --> 01:27:52.615 one score actually in improved.
NOTE Confidence: 0.8697076

01:27:52.675 --> 01:27:53.175 So
NOTE Confidence: 0.9956323

01:27:53.720 --> 01:27:54.840 in this case, we can
NOTE Confidence: 0.9956323

01:27:54.840 --> 01:27:56.060 say that we have
NOTE Confidence: 0.7887035

01:27:56.439 --> 01:27:58.220 a effective, like, fine tuned
NOTE Confidence: 0.7887035

01:27:58.360 --> 01:27:59.420 customized model.
NOTE Confidence: 0.9522619

01:28:00.520 --> 01:28:02.280 And once we find that
NOTE Confidence: 0.9522619

01:28:02.280 --> 01:28:04.280 the customized model performed dropped,
NOTE Confidence: 0.9522619

01:28:04.280 --> 01:28:05.640 we need to go back
NOTE Confidence: 0.9522619

01:28:05.640 --> 01:28:07.400 to the training process to
NOTE Confidence: 0.9522619

01:28:07.400 --> 01:28:09.045 retrain the model to see,
NOTE Confidence: 0.9522619

01:28:09.445 --> 01:28:09.945 to,
NOTE Confidence: 0.9937003

01:28:10.245 --> 01:28:10.745 like,

NOTE Confidence: 0.9373411
01:28:11.685 --> 01:28:12.185 iteratively
NOTE Confidence: 0.9890705
01:28:12.564 --> 01:28:14.005 check the performance to see
NOTE Confidence: 0.9890705
01:28:14.005 --> 01:28:15.445 if there is any,
NOTE Confidence: 0.9625634
01:28:15.765 --> 01:28:16.265 gain.
NOTE Confidence: 0.96200556
01:28:17.205 --> 01:28:18.645 And here is the code
NOTE Confidence: 0.96200556
01:28:18.645 --> 01:28:21.045 data that we provided for,
NOTE Confidence: 0.96200556
01:28:21.045 --> 01:28:22.025 like, more details.
NOTE Confidence: 0.9710739
01:28:22.810 --> 01:28:24.570 And if you have any,
NOTE Confidence: 0.9710739
01:28:24.570 --> 01:28:26.010 like, interest, you can leave
NOTE Confidence: 0.9710739
01:28:26.090 --> 01:28:27.770 if, have any question, you
NOTE Confidence: 0.9710739
01:28:27.770 --> 01:28:29.689 can leave comments or directly
NOTE Confidence: 0.9710739
01:28:29.689 --> 01:28:30.750 send me an email.
NOTE Confidence: 0.9922016
01:28:31.290 --> 01:28:32.570 Oh, okay. Thank you so
NOTE Confidence: 0.9922016
01:28:32.570 --> 01:28:33.070 much.