

# Visualization of Text Annotations for Corpus Development

Huan He<sup>\*</sup>, Sunyang Fu, Liwei Wang, Andrew Wen, Sijia Liu, Sungrim Moon, Kurt Miller, and Hongfang Liu<sup>+</sup>

Department of AI and Informatics Research, Mayo Clinic, Rochester, MN, USA

\* He.Huan@mayo.edu <sup>+</sup> Liu.Hongfang@mayo.edu

# Background

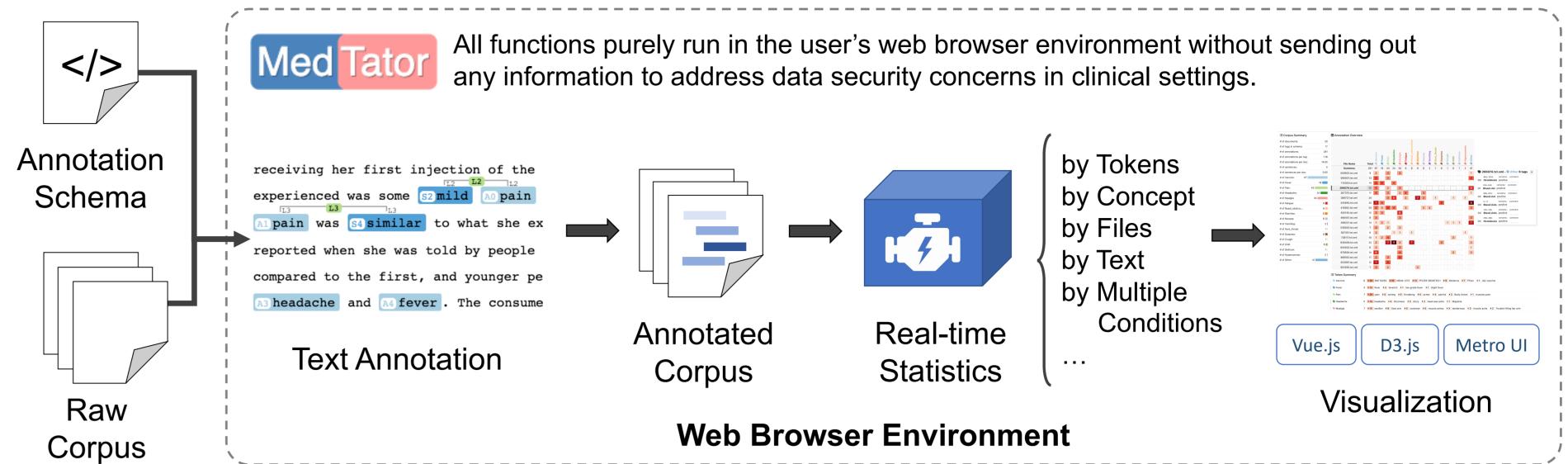
A gold standard annotated corpus is usually indispensable for building high-quality language models for clinical natural language processing.

Existing text annotation tools can provide powerful features to cover various needs for corpus development, but few tools provide real-time visualization to support the needs of annotation analysis during the annotation process.

To address this need, we developed a corpus visualization module in *MedTator*, a serverless annotation tool.

MedTator doesn't require any runtime installation! You can check our online demo at: 间歇着美国

# **System Overview**



https://ohnlp.github.io/MedTator/

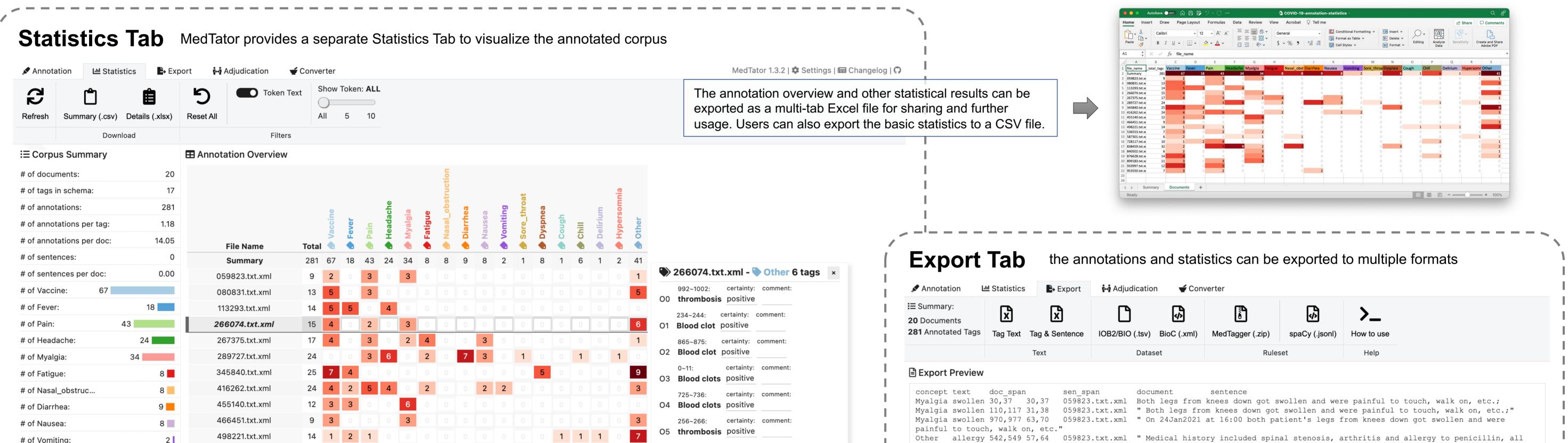
Source codes are available at: https://github.com/OHNLP/MedTator



Users can annotate the corpus by using the Annotation Tab in MedTator, and the text annotations are saved in memory.

The annotations are analyzed in real-time and shown in the Statistics Tab.

### Visualization and Interactivity Designs



The real-time statistics are shared by the Export Tab for exporting the results to other formats.	Vaccine, years) a Vaccine Vaccine, years) a Pain Pain Pain Pain Other Other Other Other male pat hurt so Other pain on right gr	<pre>, Lot number EX923, in hosp at single dose for COVID-19 PFIZER-BIONTECH 274,289 4 , Lot number EX923, in hosp at single dose for COVID-19 painful 47,54 47,54 09 painful 127,134 48,55 09 painful 987,994 80,87 09 to touch, walk on, etc." Blood clot 0,10 0 Blood clots 44,55 1 Blood clot 368,378 22 tient, who received Moderna much, blood clot in his le Blood clot 1019,1029 the left side of his upper roin, which was treated by</pre>	spital, via an unspecified route of 19 immunization." 47,62 059823.txt.xml " A 66-year spital, via an unspecified route of 19 immunization." 059823.txt.xml Both legs from knew 059823.txt.xml " Both legs from knew 059823.txt.xml " On 24Jan2021 at 1 0,10 080831.txt.xml Blood clot 1,12 080831.txt.xml " Blood clot 1,12 080831.txt.xml " A spontan ha's COVID-19 vaccine (mRNA-1273) a left and right lung and blood clots 9 148,158 080831.txt.xml " er body and was hospitalized, the for y Eliquis."	of administration in left arm from ars-old male patient received BNT of administration in left arm from ees down got swollen and were pair inees down got swollen and were pair 16:00 both patient's legs from k in his left lung and right lung ots in his right groin;" meous report was received from a and experienced terrible pain on is in right groin." On 14 Feb 2021, 3 days after var tests revealed blood clot in his	painful to touch, walk on, etc.;" knees down got swollen and were
are shared by the Export Tab for exporting the results to other formats.	Vaccine Vaccine, years) a Pain Pain Pain Pain Pain Other Other Other male pat hurt so Other pain on right gr	PFIZER-BIONTECH 274,289 4 , Lot number EX923, in hosp at single dose for COVID-19 painful 47,54 47,54 09 painful 127,134 48,55 09 painful 987,994 80,87 09 to touch, walk on, etc." Blood clot 0,10 0, Blood clots 44,55 1, Blood clot 368,378 22 tient, who received Moderna much, blood clot in his le Blood clot 1019,1029 the left side of his upper groin, which was treated by	47,62 059823.txt.xml " A 66-year spital, via an unspecified route or 19 immunization." 059823.txt.xml Both legs from knee 059823.txt.xml " Both legs from knee 059823.txt.xml " On 24Jan2021 at 3 0,10 080831.txt.xml Blood clot 1,12 080831.txt.xml " Blood clot 226,236 080831.txt.xml " A spontar na's COVID-19 vaccine (mRNA-1273) a left and right lung and blood clots 9 148,158 080831.txt.xml " a spontar solution and blood clots 9 148,158 080831.txt.xml " a spontar solution and blood clots 9 148,158 080831.txt.xml " a spontar solution and blood	of administration in left arm from ees down got swollen and were pair inees down got swollen and were pair 16:00 both patient's legs from k in his left lung and right lung ots in his right groin;" ineous report was received from a and experienced terrible pain on is in right groin." On 14 Feb 2021, 3 days after var tests revealed blood clot in his	om 19Jan2021 at 20:45 (at age of 66 inful to touch, walk on, etc.; painful to touch, walk on, etc.;" knees down got swollen and were g; a consumer concerning a 78 years old n the left side of his upper body, it accination, patient had a terrible
are shared by the Export Tab for exporting the results to other formats.	years) a Pain Pain Pain painful Other Other Other male pat hurt so Other pain on right gr Other	at single dose for COVID-19 painful 47,54 47,54 09 painful 127,134 48,55 09 painful 987,994 80,87 09 to touch, walk on, etc." Blood clot 0,10 0 Blood clots 44,55 1 Blood clot 368,378 22 tient, who received Moderna much, blood clot in his left Blood clot 1019,1029 the left side of his upper proin, which was treated by	19 immunization." 059823.txt.xml Both legs from knee 059823.txt.xml "Both legs from knee 059823.txt.xml "On 24Jan2021 at 2 0,10 080831.txt.xml Blood clot 1,12 080831.txt.xml "Blood clot 226,236 080831.txt.xml "A spontan ha's COVID-19 vaccine (mRNA-1273) a left and right lung and blood clots 9 148,158 080831.txt.xml " er body and was hospitalized, the for y Eliquis."	ees down got swollen and were pair inees down got swollen and were pair 16:00 both patient's legs from k in his left lung and right lung ots in his right groin;" ineous report was received from a and experienced terrible pain on is in right groin." On 14 Feb 2021, 3 days after van tests revealed blood clot in his	<pre>inful to touch, walk on, etc.; painful to touch, walk on, etc.;" knees down got swollen and were g; a consumer concerning a 78 years old n the left side of his upper body, it accination, patient had a terrible</pre>
Tab for exporting the results to other formats.	Pain Jain Pain painful Other Other male pat hurt so Other pain on right gr Other	painful 47,54 47,54 05 painful 127,134 48,55 05 painful 987,994 80,87 05 to touch, walk on, etc." Blood clot 0,10 0, Blood clots 44,55 1, Blood clot 368,378 22 tient, who received Moderna much, blood clot in his le Blood clot 1019,1029 the left side of his upper roin, which was treated by	059823.txt.xml Both legs from knew 059823.txt.xml "Both legs from knew 059823.txt.xml "On 24Jan2021 at 2 0,10 080831.txt.xml Blood clot 1,12 080831.txt.xml "Blood clot 226,236 080831.txt.xml "A spontage na's COVID-19 vaccine (mRNA-1273) at left and right lung and blood clots 9 148,158 080831.txt.xml " er body and was hospitalized, the start y Eliquis."	inees down got swollen and were particles of the patient's legs from kinds in his left lung and right lung onts in his right groin;" Ineous report was received from a and experienced terrible pain on its in right groin." On 14 Feb 2021, 3 days after variates revealed blood clot in his	painful to touch, walk on, etc.;" knees down got swollen and were g; a consumer concerning a 78 years old n the left side of his upper body, it accination, patient had a terrible
r •1 J&J vaccine muscles pain	Pain painful Other Other male pat hurt so Other pain on right gr Other	painful 987,994 80,87 05 to touch, walk on, etc." Blood clot 0,10 0, Blood clots 44,55 1, Blood clot 368,378 22 tient, who received Moderna much, blood clot in his le Blood clot 1019,1029 the left side of his upper roin, which was treated by	059823.txt.xml " On 24Jan2021 at 2 0,10 080831.txt.xml Blood clot 1,12 080831.txt.xml " Blood clot 226,236 080831.txt.xml " A spontar ha's COVID-19 vaccine (mRNA-1273) a left and right lung and blood clots 9 148,158 080831.txt.xml " er body and was hospitalized, the for y Eliquis."	16:00 both patient's legs from kinds in his left lung and right lung ots in his right groin;" neous report was received from a and experienced terrible pain on is in right groin." On 14 Feb 2021, 3 days after var tests revealed blood clot in his	knees down got swollen and were g; a consumer concerning a 78 years old n the left side of his upper body, it accination, patient had a terrible
r ▶1 J&J vaccine muscles pain 3 tenderness ▼3 muscle ache ▼2 Troubl	Other Other Other male pat hurt so Other pain on right gr Other	Blood clot 0,10 0, Blood clots 44,55 1, Blood clot 368,378 22 tient, who received Moderna much, blood clot in his le Blood clot 1019,1029 the left side of his upper roin, which was treated by	0,10 080831.txt.xml Blood clot 1,12 080831.txt.xml "Blood clot 226,236 080831.txt.xml "A spontan ha's COVID-19 vaccine (mRNA-1273) a left and right lung and blood clots 9 148,158 080831.txt.xml " er body and was hospitalized, the to y Eliquis."	ots in his right groin;" neous report was received from a and experienced terrible pain on is in right groin." On 14 Feb 2021, 3 days after var tests revealed blood clot in his	a consumer concerning a 78 years old n the left side of his upper body, it accination, patient had a terrible
a muscles pain 3 tenderness ▼3 muscle ache ▼2 Troubl	Other male pat hurt so Other pain on right gr Other	Blood clot 368,378 22 tient, who received Moderna much, blood clot in his le Blood clot 1019,1029 the left side of his upper roin, which was treated by	226,236 080831.txt.xml " A spontan ha's COVID-19 vaccine (mRNA-1273) a left and right lung and blood clots 9 148,158 080831.txt.xml " er body and was hospitalized, the to y Eliquis."	neous report was received from a and experienced terrible pain on s in right groin." On 14 Feb 2021, 3 days after var tests revealed blood clot in his	n the left side of his upper body, it accination, patient had a terrible
a muscles pain 3 tenderness ▼3 muscle ache ▼2 Troubl	hurt so Other pain on right gr Other	much, blood clot in his le Blood clot 1019,1029 the left side of his upper roin, which was treated by	left and right lung and blood clot: 9 148,158 080831.txt.xml " er body and was hospitalized, the f y Eliquis."	s in right groin." On 14 Feb 2021, 3 days after va tests revealed blood clot in his	accination, patient had a terrible
a muscles pain 3 tenderness ▼3 muscle ache ▼2 Troubl	pain on right gr Other	the left side of his upper roin, which was treated by	er body and was hospitalized, the t y Eliquis."	tests revealed blood clot in his	
a muscles pain 3 tenderness ▼3 muscle ache ▼2 Troubl	Other			seous report was received from a	
3 tenderness ▼3 muscle ache ▼2 Troubl	Both	1		leous report was received from a	a consumer concerning a 78 years old 🅢
3 tenderness ▼3 muscle ache ▼2 Troubl	Both				
3 tenderness ▼3 muscle ache ▼2 Troubl	Both				
	legs		2/BIO format for	<pre>{"label":"MYALGIA","pattern":"swollen","id": {"label":"MYALGIA","pattern":"sore arm","id" {"label":"MYALGIA","pattern":"trouble liftir</pre>	ledTagger format
	from knees	0		{"label":"MYALGIA","pattern":"muscle aches",	
			ing NLP models	{"label":"OTHER","pattern":"thrombosis","id"	or building rule-
3 > 455140.txt.xml 3 > 840502.txt.xml 2 > 26737		B-Myalgia (e.g.	., fine-tuning	<pre>("label":"OTHER", "pattern": "trouble sleepin; ("label":"OTHER", "pattern": "sepsis", "id":"Ot ("label":"OTHER", "pattern": "sudden death", "i ("label":"OTHER", "pattern": "died", "id":"Other "label":"OTHER", "pattern": "died", "id":"Other"</pre>	ased NLP systems
s of each token are listed, and the u			RT NER models)	<pre>{"label":"OTHER","pattern":"parkinson's dise {"label":"OTHER","pattern":"parkinsons dises {"label":"OTHER","pattern":"stroke history",</pre>	r <b>spaCy</b> based
e name to look back to the selected	Parmur			{"label":"OTHER","pattern":"bacterial pneumo	
in the Annotation Tab for revision.	l couch	-			ulesets
	~				
tion Tab MedTat	tor provides all c	ore functions for te	xt annotation tasks for m	ultiple documents	
🖞 Statistics 🔹 Export 🚧 Adjudication 🛫 C	Converter				MedTator 1.3.2   🏟 Settings   🖬 Changelog   🖓
s 1294 chars			0	No Hint	<i>A</i> <sub>a</sub> i
15 tags Save Sa	ave as Set Labels O	Sentences 👌 Clear	Color Only Show Lines		chema Wiki
on / dtd) 🚣 Annotation File (.xml) Save	re File Label Dis	bisplay Mode Search E	Entity Marks Link Marks	Hint Marks He	lelp
× 20 files I O3 Blood clots in	arm;				
9 <b>o</b> 2 Arm was sore;					
	ort was received fr	om a healthcare profes	sional concerning a 67-year-	old, female patient, who rea	ceived V3 Moderna 's COVID-19
15 <b>•</b> vaccine ( <u>vo</u> mRNA-12	273) and experience	ed MOSore arm / POpain	in arm and OlBlood clot in	her arm/05thrombosis.	
	dical history, as p	provided by the reporte	er included diabetes.		
25 🗢 24 👝 5 Concomitant medica	ations were not inc	luded.			
12 On 15 Feb 2021,	prior to the onset	of the events, the pat	cient received their first of	two planned doses of vimm	RNA-1273 (Lot number: 031M20A)
	r prophylaxis of CO	OVID-19 infection.			Entity Tags:
7 • 7 On 15 Feb 2021,	post vaccination t	the patient experienced	I MISore arm and her conditi	on got worse.	
	n to emergency room	a with 04 Blood clots i	in her arm.		<ul> <li>2 Fever</li> <li>3 Pain</li> </ul>
32 • Action taken with					4 Headache
6		_	_	;	<ul> <li>5 Myalgia</li> <li>• q Fatigue</li> </ul>
11 •			-		
					e Diarrhea
		macron and cemporar as	Sector Detween the use of	the product and the start of	date of t nausea
		Text	Attributes		💊 a Sore_throat
	is opinio	thrombosis	certainty comment		Dyspnea a Cough
0			positive		f Chill
2 🎙 Myalgia	M0 209~217	Sore arm	certainty comment		g Delirium
0			positive		<ul> <li>z Hypersomnia</li> <li>x Other</li> </ul>
3 Nyalgia	M1 645~653	Sore arm	certainty comment positive		
ion 0 🎙 Myalgia	M2 852~860	Sore arm	certainty comment		
0			positive		
0 🗣 Other 🝵	01 234~244	,			
-		Blood clot	certainty comment		
0			certainty comment positive		
0 0 Other		Blood clot Blood clot			
s s	Statistics Export i Adjudication i Save S 266074.txt.xml 1294 chars 15 tags Save S a/ dtd) Annotation File (.xml) Save 20 files All 1 3 Annotation File (.xml) Save 20 files All 1 3 Arm was sore; 13 0 14 0 15 0 14 0 12 0 14 0 15 Feb 2021, 5 he was then take Action taken with 16 0 11 0 11 0 12 0 14 0 10 0 15 Feb 2021, 5 he was then take Action taken with 16 0 17 0 10 0 11 0 11 0 11 0 11 0 11 0 12 0 11 0 12 0 13 0 14 0 10 0 11	Statistics       ► Export       i Adjudication       ✓ Converter         1294 chars 1294 chars 15 tags       Image: Save as Save as Save Save as Save as Save as Save as Save as Save Save as Save as Sa	Statistics       Export       14 Adjudication       Converter         266074.txt.xml 1294 chars 15 tags       Image: Save Save as Set Labels       Document Set Labels       Sentences       Clear         1 (dd)       Annotation File (xml)       Save       Save Save as Set Labels       Sentences       Clear         20 files       Image: Save Save as 14 do       1       Image: Save Save as Set Labels       Sentences       Clear         2 of files       Image: Save Save Save Save Save Save Save Save	Statistics       Export       H Adjudication       Converter         266074.trt.xml 1294 chars       Image: Save Save as 15 tags       Image: Save Save as Save Save Save as Save Save Save Save Save Save Save Save	Statistics       Is tayon       I Adductation       Converter         286074       Annotation File (xm)       Save       Save Save as SatLabels       Document       Search       Color - HD       Show Links       Image: Adductation File (xm)       Image:



This work has been supported by the National Center for Advancing Translational Science of the National Institutes of Health under award number U01TR002062.