

Formulating Functional Terminology: Manual annotation Guideline for Body Function

Epidemiology and Biostatistics Section:

Rehabilitation Medicine Department

National Institutes of Health (NIH) Clinical Center

October 2021

Table of Contents

1	Introduction.....	4
2	Data Collection.....	4
3	Body Function Annotation Schema.....	6
3.1	<i>Definition of Terms in the Body Function Schema</i>	8
4	Body function annotation	9
4.1	General Considerations.....	9
4.1.1	Gate Developer.....	9
4.1.2	Optical character recognition (OCR).....	10
4.1.3	Double/overlapping annotations	11
4.1.4	Minimum information needed	11
4.1.5	Annotate Objective information only	11
4.1.6	Do not annotate body structure information.....	12
4.1.7	Punctuation	12
4.2	Annotation of Entities.....	12
4.2.1	Body Function Entity	12
4.2.2	Body Function Type Sub-entity.....	14
4.2.3	Body Location Sub-Entity	18
4.2.4	Qualifier Sub-Entity	19
4.2.5	Body Function Context Entity.....	23
4.2.6	Possible Body Function Entity	25
5	Sample Difficult cases.....	26
5.1	Double annotation when Rom and Strength form the same body function mention	26
5.2	Extraocular movements derivatives	27
5.3	ROM vs. Strength inference.	28
5.4	Negation annotations.	28
6	Software.....	28
6.1	The Gate Developer Main Window	29
6.2	Loading and viewing documents.....	29
6.3	Loading a Schema Type.....	31
6.4	Creating and editing annotations.....	33
6.5	Examples of Annotating Different Mentions.....	34
6.6	Saving GATE software.....	37

6.7	Exporting Annotation Data	37
6.8	Exiting GATE Software	38
7	References.....	38
8	Appendixes.....	38

1 Introduction

The focus of the Body Function (BF) annotation task is to develop a natural language processing (NLP) tool that will capture strength, range of motion (ROM), and reflex information in the medical record. In developing such a tool, we hope to be able to augment other existing NLP algorithms that are currently applied, or will be applied, to capture information about function in the medical record.

In short, the goals of this project are to:

- Develop an NLP tool that will capture strength, ROM, or reflexes in the medical record.
- Develop principles and algorithms that will augment other NLP tools and aid in other

To achieve this tasks we utilized the Systemized Nomenclature of Medicine (SNOMED), a highly structured database containing standardized language and definitions for clinical terms and/or concepts (<https://www.nlm.nih.gov/healthit/snomedct/index.html>).

SNOMED codes form the basis for our BF annotations to 1) ensure consistent application of definitions, and 2) ground our annotations in a wider, conceptual framework of terminology.

The following SNOMED codes were employed in this annotation task:

1. **Strength (SCTID: 44432004)**
2. **ROM (SCTID: 364564000)**
3. **Reflexes (SCTID: 87572000)**

2 Data Collection

For annotating information related to Body Function, data were gathered from two different sources. The first dataset was provided by the Biomedical Translational Research Information System (BTRIS) at the NIH. A team of clinicians and researchers discussed which type of medical documentation would likely yield Body Function information. From these discussions, five medical record types were selected, then sampled from within the NIH CC's Department of Rehabilitation Medicine for documents containing BF information. These five types of medical documents were:

1. History and Physical Exams (HPEs)
2. Physiatry (PHYS)

3. Occupational Therapy (OT)
4. Physical Therapy Initial Assessment (PTIA)
5. Physical Therapy Assessment and Discharge (PTAD)

The process of completing the manual annotations on BTRIS data is outlined below in figure 2.

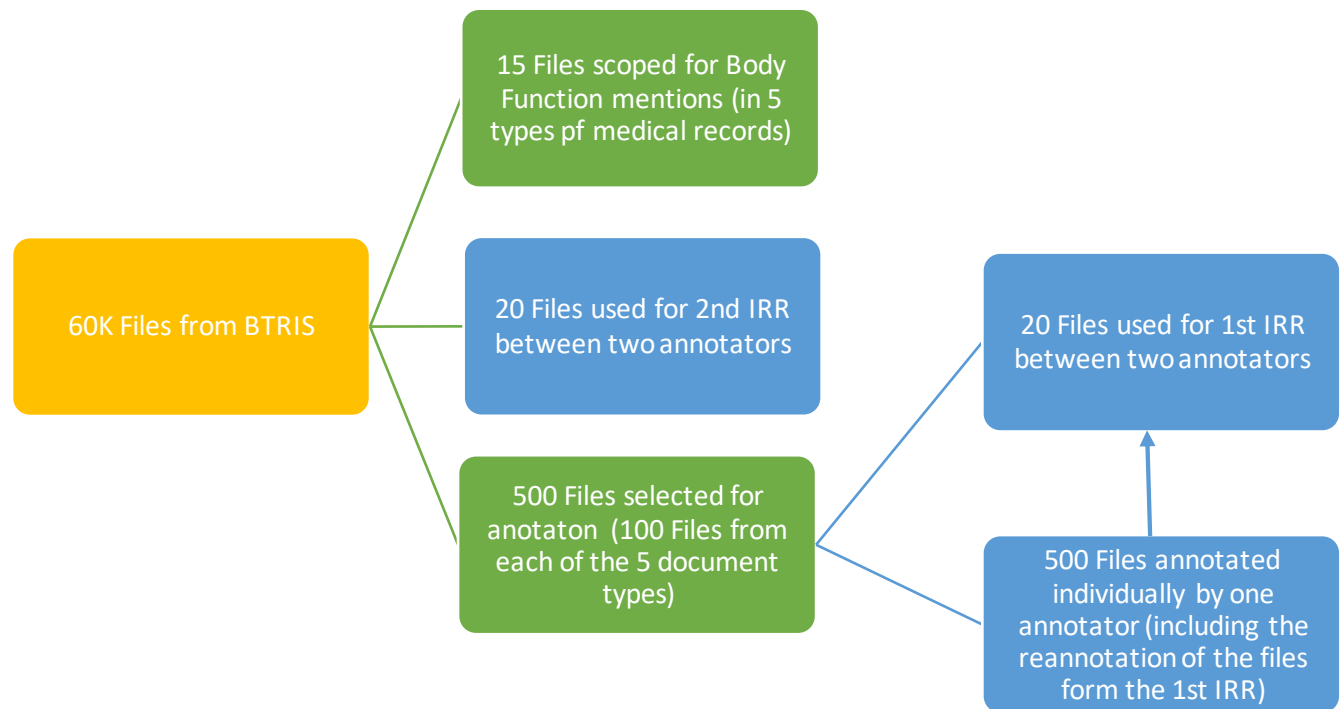


Figure 1. BTRIS data annotations

The second dataset came from the Social Security Administration (SSA) 2019 Adult Consultative Examination (CE) Reports, Medical Evidence Record (MER) and to a minimum extend from other data sources provided by the claimant for disability determination. A subset of this 2019 SSA dataset was scoped for relevant body function mentions using a lexicon of 2885 terms.

The process of completing the manual annotations on SSA data is outlined below in figure 3.

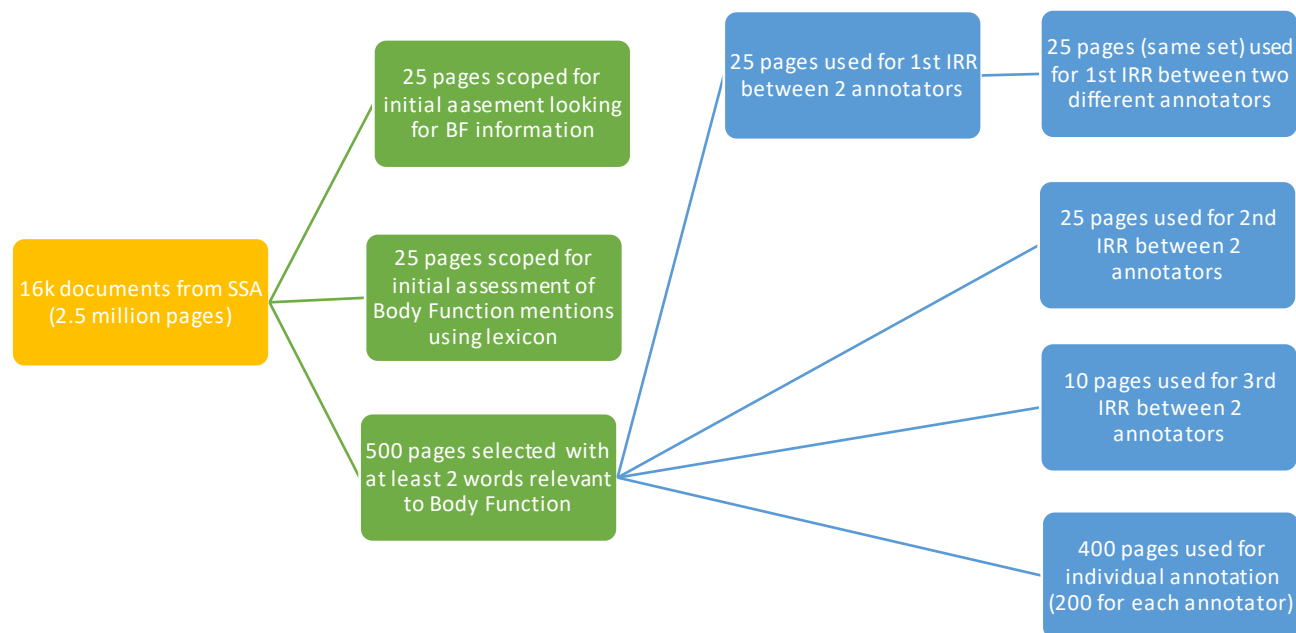


Figure 2. SSA data annotation

3 Body Function Annotation Schema

An annotation schema to label Body function related information in clinical notes was developed. The Body function Annotation Schema (see Figure 4) displays the hierarchical structure of the Entities, Sub-entities, Attributes, and Values. Entities are the highest level of the schema (Body Function, Body Function Context, Possible Body Function), followed by Sub-entities [Body Function Type, Body Location, Qualifier], Attributes [see light green boxes], and Values [see purple-colored boxes]. The schema does not force mandatory assignment of all Attributes and Values which is why the annotator should be careful in making sure the right information is selected.

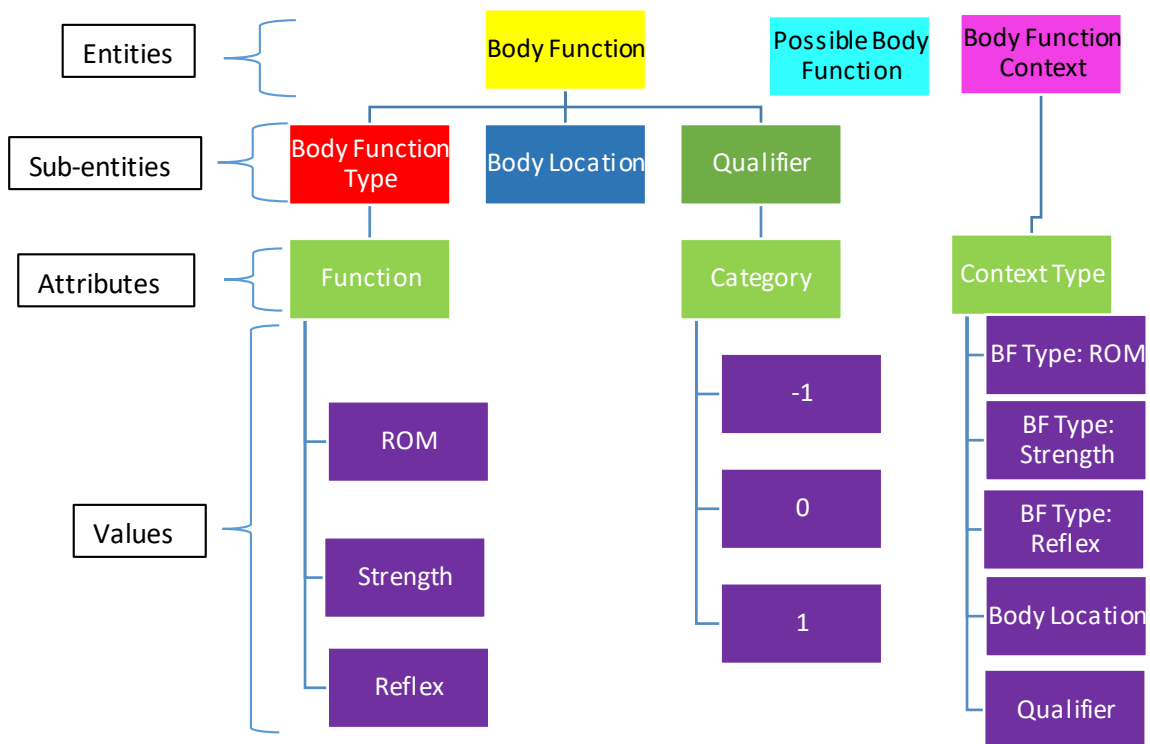


Figure 3: Body Function Annotation schema

Below in figure 4 is an example of how clinical free text was annotated using the Body Function schema.

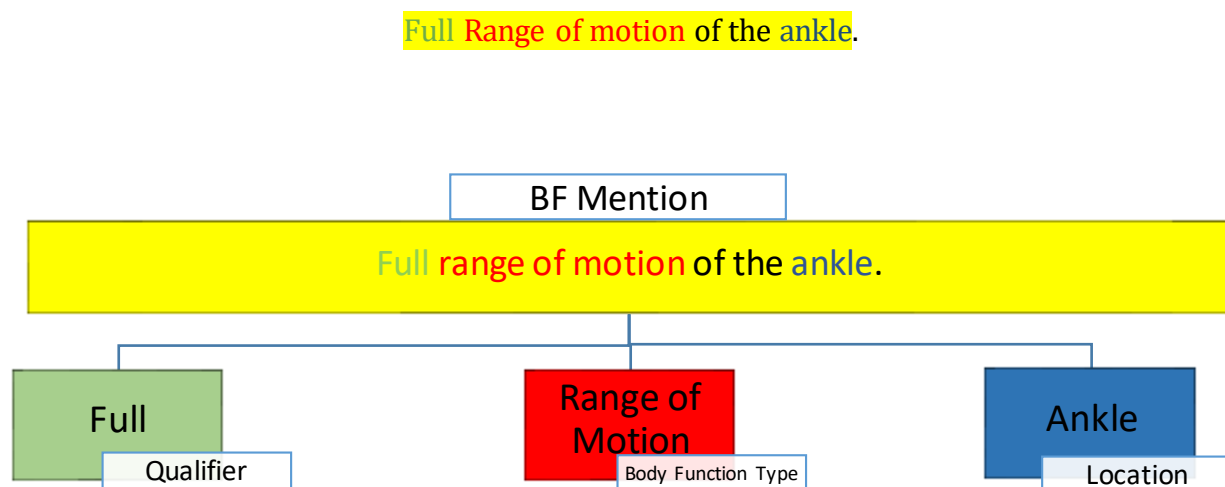


Figure 4: Body Function mention, sub-entities and values are demonstrated here

3.1 *Definition of Terms in the Body Function Schema*

Mention/Sub-mention: A textual statement that represents an instance of an underlying entity/sub-entity. It is a contiguous span of tokens that makes up a phrase, a simple sentence, or a complex sentence.

Body Function mention: A self-contained, well-defined description of information that includes range of motion, strength, and/or reflex, along with its qualifiers/location.

Body Function Type Sub-mention: information about the mention of range of motion, strength, or reflex.

Body Location Sub-mention: Information about the location in the body related to a range of motion, strength, or reflex mention.

Qualifier Sub-mention: Objective information related to range of motion (degrees and/or percentages of range of motion), strength (standard MMT scoring), or reflex (hypoactive/absent, normal, hyperactive).

Body Function Context mention: Information about the semantic context of the section (which can contain one or more body function mentions) that follows.

Possible Body Function mention: Information that appears to have body function information in it but is difficult to determine the meaning of because it is formatted inconsistently with the data presented or appears OCR'd poorly.

4 Body function annotation

4.1 General Considerations

4.1.1 Gate Developer

The GATE Developer software program was used for Body Function annotation with the following color-coded scheme: the entire Body Function mention is highlighted in yellow, the Body Function Type mention text is red, the Body Location mention text is blue, the Qualifier mention text is green, Body Function Context mentions are highlighted in violet, and Possible Body Function mentions are highlighted in turquoise.

Table 1. The color-coding scheme for Body Function mentions.

Color scheme in GATE	Color schema in Document
Body Function	Yellow highlight
Body Function Type	Red text
Body Location	Blue text
Qualifier	Green text
Body Function Context	Violet highlight
Possible Body Function	Turquoise Highlight

The sequence to annotate Body Function is to: (1) first capture the Body Function entity, (2) followed by its Sub-entities (3) followed by Attributes, (4) followed by the Values. Mentions should be annotated with the least amount of text in mind. Pronouns, such as he, she, patient, etc. are not included if they can be avoided. Redactions are not annotated unless relevant data cannot be otherwise captured, i.e., often found mid mention. Attributes classify and provide information about the specific sub-entities (Body Function Type, Body Location,

Qualifier) that fall under the Body Function Entity. Attributes are assigned Values that provide more detailed information regarding specific type and amount of body function present.

4.1.2 Optical character recognition (OCR)

OCR errors do occur in clinical documents for annotation. As much as possible, if the text is understandable, it is annotated. Below is a table with list of common OCR errors and how to accommodate in annotations.

Table 2: OCR errors and solutions.

Type of OCR error:	Example of annotation	How to handle
Substitution of original characters for similar ones:	Strength \$/5, 5t/5, 5-15	Annotate as if no OCR error present if sure about information. In this case assume the qualifiers for the first two examples are 5/5 and the third one in 5-/5
Misspelling	Ioskeletal: Normal strength, normal ROM	Annotate as Musculoskeletal BF Context Type: location given subject matter expert can assume as such
Misspelling	Nonnal motor function in all four extremties	Annotate assuming "nonnal" indicates "normal"

In addition, sometimes OCR makes the document impossible to annotate even though there is relevant Body Function information, in these cases we annotate the information under the "Possible Body Function" entity. Please refer to that section for more information.

4.1.3 Double/overlapping annotations

Annotations of different sub-entities may overlap. However, the same entity is not permitted to overlap with itself. Also, the Body Function entity (that contains all sub-entities) never overlaps with another Body Function entity.

Example 1: Ankle dorsiflexion ROM Bilaterally is 15 degrees

In this example, the annotation for the Location entity should be “ankle bilaterally”. There is no way to connect these two words without including the words “dorsiflexion ROM” so the Location entity in this case is annotated as “ankle dorsiflexion ROM bilaterally” and we understand that the Location entity overlaps with the Body Function Type entity.

4.1.4 Minimum information needed

At a minimum, Body Function information should contain a Body Function, Body Function Type, and Qualifier entities. Location entity can be left out if not available.

4.1.5 Annotate Objective information only

- Objective information is captured in our use case for facilitation of SSA disability adjudication to present objective findings regarding body function. It is based on standardized documentation of ROM, Strength and Reflex values. Please click on the respective links to access the documentation.
- Do not annotate subjective or planning information is annotated (i.e., HPI, PMH, goals, Plan).
- Do not annotate treatment information considered in body function work, it doesn't contain qualifiers to allow objective documentation of BF.
- Do not annotate treatment plans, as they do not necessarily speak to an individual's ability to perform certain tasks from an objective measurement perspective (described with the purpose of treating a demonstrated limitation in function).

4.1.6 Do not annotate body structure information

No body structure information is included in annotations. Some of these mentions could have trigger words related to body function but need to be excluded because we cannot infer measurements of strength, ROM, or reflexes.

Example 1: Does not have knee joint flexion contractures

Example 2: LUMBAR SPINE No scoliosis, asymmetry or abnormal curvature noted

These examples would NOT be annotated as they do not infer measurement of BF of strength, ROM or reflexes and are likely a postural/structural observation.

4.1.7 Punctuation

Do not include periods and/or commas in the annotation, the goal is to have the absolute shortest span of text.

4.2 Annotation of Entities

4.2.1 Body Function Entity

The Body Function Entity is defined as self-contained, well-defined information containing mentions of Range of motion, strength, and/or reflexes, along with the location of these descriptions, and accompanying qualifiers (i.e., degrees of ROM, Strength scale, Reflex scale).

Example 1: Shoulder ROM is 120 degrees

Example 2: Manual Muscle test of the quadriceps is 4/5 on the L and 5/5 on the R

4.2.1.1 Attributes of the Body Function Entity

The Body Function entity has NO attributes related to it

4.2.1.2 Special Cases and Considerations of Body Function Entity

4.2.1.2.1 Mentions referring to body function that are too broad should not be annotated

Some phrases containing the word “weak” or “weakness” (that sometimes are annotated because they refer to the strength of the patient) are not annotated as Body Function entities as we cannot ascertain that this is considered a deficit in muscle strength given lack of adjacent text to further inform us of specific Body Function Type “strength”.

Example 1: Generalized weakness

Example 2: Weakness throughout

Example 3: Weak and pale

Example 4: Nauseous and weak

4.2.1.2.2 Annotation of Headers and Titles

Headers/Titles are annotated as part of the Body function mention when:

1. There is a sentence related to Body Function immediately following the header/title, and,
2. There is only one sentence related to Body Function in the entire section related to that header/title.

Both conditions must be met for annotating a title/header as part of the Body Function entity.

Example 1: UPPEREXTREMITIES: Full ROM of shoulders.

Example 2: LUE: strength 5/5. The patient also shows weakness in right lower extremity

In the first example the only information following the title is directly related to body function so the whole span of text is annotated. In the second example, the information directly following the title is related to Body function but the rest of the information in the paragraph is not, so we can exclude it from the Body Function mention. This distinction of

WHEN to annotate the title/header as part of the Body Function mention or not is important when comparing this kind of annotation to annotating titles/header as part of the Body Function Context as it helps the annotators be more consistent without losing relevant Body function information. Please refer to the *Body Function Context entity* section of this guideline for more information.

4.2.2 Body Function Type Sub-entity

Body Function Type is defined as the mention of range of motion, strength, and/or reflexes.

4.2.2.1 Attributes and values of the Body Function Type Sub-Entity

Body Function Type has only one attribute named “Function”, and four values named “ROM”, “Strength”, “Reflex”, and “Ambiguous”.

4.2.2.1.1 ROM

Range of motion, abbreviated as ROM, is a term to identify available motion at a joint or body part (i.e., shoulder, trunk, forearm), and is measured in degrees or percentages in this work. Various types of ROM are denoted as: passive range of motion (PROM), active range of motion (AROM), resisted range of motion (RROM) and gross ROM. Range of motion is also often described as joint motion in terms such as flexion, extension, rotation, supination, pronation, ulnar deviation, dorsiflexion etc. and can be further identified by their qualifier in degrees which inform us of ROM body function type.

Example 1: Full ROM of shoulders, elbows, forearms, and wrists bilaterally

Example 2: Bilateral Feet: Dorsiflexion 20, plantar flexion 40, inversion 30', eversion 20

Example 3: CERVICAL SPINE: Flexion 30 degrees, extension 25 degrees, lateral flexion 35 degrees bilaterally

4.2.2.1.1.1 Special Cases and Considerations ROM

1. Adjacent text spans of the same entity: annotate separately. Sometimes, the notes provide “redundant” information.

Example 1: Ankle DF ROM is full

In this example, “DF” and “ROM” are both terms that describe ROM. Even though “DF” is more specific, we will capture “ROM” as well as a separate Body Function Type mention with a value of ROM.

2. Words that need to be annotated together

Some Range of Motion mentions need to be annotated together to better define the concept of a ROM mention. The following spans of text are annotated together as ONE Body Function Type: ROM mention.

Example 1: Lateral flexion

Example 2: Forward flexion

Example 3: Radial deviation

Example 4: Ulnar Deviation

4.2.2.1.2 *Strength*

Defined as one’s ability to resist force using a standardized manual muscle test scale of 0-5/5 or 0-10/10 to quantify the amount of strength available to a tester’s resistance. Testing in this format (Manual Muscle Testing) considers the subject’s (body part) position and its relationship to gravity as well as amount of manually applied resistance the subject can withstand and is graded accordingly. Strength can be referred to both by action, such as “flexion” or “abduction”, or the muscle being isolated such as “anterior deltoid”.

Example 1: Strength 5/5 in proximal and distal muscles

NOTE: Please have in mind that we do not annotated coordination, atrophy, and tone as part of Strength.

4.2.2.1.2.1 Special cases for Strength

1. When strength is not otherwise mentioned but the action is graded on a 0-5 scale or 0-10 scale, or the term “resistance” or position related to gravity is contained in the mention, we can assume from the qualifier it is referring to strength as opposed to Range of Motion. In addition, there are a couple of considerations to have in mind when annotating this kind of information.

Example 1: Dorsiflexion 5/5

Example 2: 4/5 grip, 4/5 small finger abduction, 4/5 wrist extension, 4/5 wrist flexion.

In the above two examples we can see the qualifiers are an indication of strength. To keep consistency and correctly identify these annotations, they will be annotated only as **Body Function Type: Strength**, and NOT double annotated as both Strength and ROM.

Example 3: Anterior deltoid 3/5 (location and type)

Example 4: 4/5 biceps, 4/5, Triceps, 4/5 deltoid.

In cases like examples 3 and 4, where because of the qualifiers we know the information being annotated is related to strength, but we find mentions of locations instead of strength (e.g., biceps, triceps, deltoid, we will annotate these mentions as both **Body Function Type: Strength**, as well as **Location**.

2. Annotating “Weak” and “weakness”

The word “weak” and its derivatives are only annotated if they are understood as synonyms of strength (e.g., mentioned next to a body part) and not as broad statements like we indicated in the *Body Function entity in 4.2.1.2.1*. In cases like the examples below a annotate “weakness” as **Body Function Type: strength**.

Example 1: Shoulder weakness

Example 2: weak left hip

NOTE: Have in mind that in these cases the word “weak” and its derivatives are also annotated as **Qualifier: -1** as explained in the *Qualifier section* of this guideline.

4.2.2.1.3 Reflex

A reflex is an automatic and often inborn response to a stimulus that typically involves a nerve impulse passing inward from a receptor to the spinal cord and then passing outward from the spinal cord to an effector (such as a muscle or gland) without reaching the level of consciousness and often without passing to the brain (Merriam-Webster, n.d.)

Example 1: Achilles tendon reflex 2+ b/l

Example 2: DTRs are 2+ and symmetric at the knees and ankles

4.2.2.1.3.1 Special cases for Reflex

1. **Babinski, Knee jerk, Achilles reflexes:** When we find mentions of Babinski sign, Knee jerk, or Achilles reflexes, annotate them as reflex mentions even if the word “reflex” is not there.

Example 1: Babinski: down going

Example 2: Knee jerk: Normal

Example 3: Achilles: Normal

2. **Pupillary reflex:** When we find mentions of pupillary reflex annotate them as reflex mentions even if the word “reflex” is not there.

Example: pupils were equal and reactive to light and accommodation

3. **Hofmann:** When we find mentions of Hoffman reflex annotate them as reflex mentions even if the word “reflex” is not there.

Example: Hoffman: negative

4. **PERRL:** This abbreviation means “Pupils equally round and reactive to light” which in turn means that there is a Normal reflex response in the eye. Therefore, annotate this abbreviation as Body Function Type: Reflex. (Also annotate the whole abbreviation as a Qualifier mention with a Value of 1, and a location mention).

Example: Eyes: **PERRL**

4.2.3 Body Location Sub-Entity

Location is considered where the body function is located such as “shoulder”, “right wrist”, “bilaterally”.

Example 1: **Left shoulder strength 4/5**

4.2.3.1 Attributes of the Body Location Sub-Entity

The location Sub-Entity has NO Attributes or values.

4.2.3.2 Special Cases and Considerations of Body Location Sub-Entity

1. Laterality should be included WITH a body location where possible to further inform us of the body parts involved. Moreover, where mentions contain locations that are distant within the mention, we will capture the whole mention to maintain a laterality with a location:

Example 1: **Bicep strength is 4/5 on the R and 2/5 on the L**

The entire mention would be considered the location as “Bicep” is the initial location followed by the “R” and “L” and should be captured together for a more granular definition of location.

Example 2: **Shoulder strength is 4/5 bilaterally**

Same as the previous example, the entire span of text is considered a Location mention to further inform of the body parts involved.

Example 3: **Biceps, triceps and brachioradialis reflexes are 2 and bilaterally**

Here the entire span from “Biceps” to “bilaterally” would be annotated as body location instead of “Biceps”, “triceps”, “brachioradialis” annotated independently of each other as separate locations to allow further definition of “biceps, triceps and brachioradialis as “bilaterally”.

Example 4: Shoulder flexion on the right is 120 degrees on the L is 130 degrees

In this example the Location mention stops at “L” because we always try to annotate the shortest amount of text within a particular mention.

2. Inference of locations can be annotated as such according to the context of the BF mention.

Example: Decreased grip and pinch strength

In this example we can clearly determine that “grip” and “pinch” refer to the hand, therefore can be annotated as Location mentions.

4.2.4 Qualifier Sub-Entity

The Qualifier sub-entity refers to numerical values associated with each of the Body Function types (ROM, strength, reflex). In the case of ROM, the Qualifier is determined from the degrees of movement of a part of the body; For Strength, the Qualifier is determined according to a 0-5/5 or 0-10/10 scale; and for Reflexes, the Qualifier is determined according to a standard scale for reflexes. A more detailed explanation on how to select the correct value for the “Category” attribute of this Sub-entity is described in the next paragraph.

4.2.4.1 Attributes of the Qualifier Sub-Entity

Qualifier has an attribute of “Category” with values of -1 (Impaired), 0 (Ambiguous), and 1 (Normal) that are assigned according to the standardized scales related to *ROM, Strength, or Reflex* referenced here.

4.2.4.1.1 ROM Qualifiers

The numerical values associated with ROM are compared to normative values by motion/joint as outlined in American Academy of Orthopedic Surgeons (AAOS, 1965). This document contains tables with the normal range of motions in different joints and it is this information that we should consider. When no degree is noted (e.g., 50%, partial, decreased) then the qualifier goes into the “impaired” category.

Example 1: Decreased cervical ROM (-1)

Example 2: Full range of motion in all extremities (1)

NOTE: Please have in mind that if the document we are annotating explicitly defines the norm for ROM we must compare the qualifier to that norm and not to the normative values outlined in American Academy of Orthopedic Surgeons (AAOS, 1965).

Example 1: Shoulder: Flexion 180/180, Extension 50/50, Abduction 180/180

4.2.4.1.2 Strength qualifiers

The numerical values associated with Strength are compared to normative values as outlined by Kendall (Kendall et al., 1983). Objectively measured in rehabilitation practices, standard MMT scoring uses 5/5 as “normal” strength or 10/10 on a 10-pt. scale. Scores below 5/5, assessed as outlined above, are given a -1 or “below level” qualifier, scores equal to 5/5 or 10/10 are given a 1 value, and if no determination can be made, a value of 0 is given.

Example 1: 5/5 Hip flexors (1)

Example 2: RUE strength good (0)

Example 3: Hamstrings 4/5 on the right (-1)

4.2.4.1.3 Reflex qualifiers

The numerical values associated with Strength are compared to normative values as outlined by Walker (Walker, K. H., 1990). Reflexes are objectively measured on a 0-4 scale. 2+ is determined in clinical language as a normal deep tendon reflex and so it is assigned a value of 1 (same applies for “intact” reflexes). Below 2+ and above 2+ are considered hypoactive and hyperactive responses respectively so these annotations are assigned a value of -1 (same applies for “absent” reflexes).

Example 1: DTRs are 2+ in Upper and Lower extremities (1)

Example 2: Reflexes symmetric in both upper extremities (0)

Example 3: Decreased reflexes in right lower extremity (-1)

4.2.4.1.4 Additional Qualifiers

The following mentions are also annotated as Qualifiers with their respective values:

Able to (value 1)

Without difficulty (Value 1)

Physiologic (Value 1)

Adequate (Value 1)

Equal (Value 0 or 1)

Flexion **vu degrees** (Value 0 since we know they are talking about a Qualifier mention but cannot determine the value)

Unable to (Value -1)

Weak/Weakness (Value -1)

With difficulty (Value -1)

Keep in mind that these Qualifiers should also be annotated only as part of a Body Function mention:

Example: Able to extend the hand

4.2.4.2 Special Cases and Considerations of Qualifier Sub-Entity

1. When a Body Function mention contains two or more Qualifiers, we can still annotate them as individual Qualifiers.

Example1: Actively, pt. able to flex RUE to 175 degrees

2. There are rare cases where a qualifier does not apply to the previously mentioned guidelines. In cases like this we can still capture the Qualifier mentions with a value of 0 (ambiguous)

Example 1: Shoulder rotation to T/L

In this case the qualifier is not defined in degrees of ROM but to a location, and therefore we cannot align it to our references.

3. +/- Signs will be captured to indicate a positive or negative finding for limitation in Body Function (hence it is annotated with a value of 0) and are relevant to qualifier levels such as reflexes.

Example 1: Achilles reflex: +

4. The following table shows a summary of the most common words used as qualifiers and the value assigned to them when annotating them as such.

Table 3. Summary of the most common words used as qualifiers

-1 (below-level function)	0 (ambiguous function)	+1 (at-level function)
Asymmetric/asymmetries	Good	Intact
Hyperactive	Improved/Improvement	Negative (ex. Musculoskeletal: <i>Negative</i> ; not in reference to degrees)
Hypoactive	Within functional limits/WFL	Full
Hypermobility	Greater than anti-gravity (we don't know if resistance was applied to determine if at level)	5/5 (in reference to MMT)

Hypomobile	Good functional	10/10 (in reference to MMT)
Weakness/weak	Functional	Within normal limits/WNL
Limited/limitations	Strong	Full resistance
Impaired/impairment	Symmetric	2+ (in reference to reflexes)
decrease/decreased	+/-	Symmetric
Reduced		Unaffected
Non-functional		Physiologic
Less than anti-gravity		Preserved
Trace		Equal, round, and reactive to light
Unable		Able to
Tight/tightness		
Stiff/stiffness		
Clonus		
Slight resistance		
Slightly against gravity (implies they cannot move against gravity with full resistance completely and therefore would be below level)		

4.2.5 Body Function Context Entity

Body Function Context is a label to be used to set the semantic context of the section (which contains one or more body function mentions) that follows. It is not itself a body function mention.

Example: UPPER EXTREMITIES: Full ROM of shoulders bilaterally. Full ROM of elbows, forearms, wrists, and fingers bilaterally.

4.2.5.1 Attributes and values of the Body Function Context Entity

The Body Function Context entity has one attribute of “Context Type” and five values: BF type ROM, BF type Strength, BF type Reflex, Body Location, and Qualifier. Following are examples on annotations of Body Function Context with its different values. The same definitions of ROM, Strength, Reflex, and Location we use to annotate the Body Function Types can guide our decision making.

4.2.5.1.1 Context Type with value BF type ROM

Example: AROM: Rot R: 72, Rot L: 64

4.2.5.1.2 Context Type with value BF type Strength

Example: Strength: 4/5 grip, 4/5 small finger abduction

4.2.5.1.3 Context Type with value BF type Reflex

Example: Reflexes: 2+ Achilles, 2+ knee jerk

4.2.5.1.4 Context Type with value Body Location

Example: THORACIC AND LUMBAR SPINES: Flexion with support 70 degrees, extension 10 degrees

4.2.5.2 Special Cases and Considerations of Body Function Context Entity

1. Annotate a header/title as a Body Function Context when there is ONE relevant Body Function information related to this header/title but there is irrelevant information in between these two.

Example 1: Back: Tenderness to palpation: upper T-spine, ROM: limited

Example 2: UPPER EXTREMITIES: No joint inflammation, effusion, or instability. Strength 5/5 in proximal and distal muscles

2. Annotate a header/title as a Body Function Context when there are multiple Body Function mentions related to this header/title.

Example 1: Motor: 4/5 grip, 4/5 small finger abduction, 4/5 wrist extension

Example 2: THORACIC AND LUMBAR SPINES: Flexion with support 70 degrees, extension 10 degrees

3. In some instances, the context is not determined by a title /header but by the closest text related to the Body Function mention we need to annotate.

Example 1: The lower back exam shows good ROM with normal flexion and extension

NOTE: Please compare the guidelines for annotation of context with the annotation of header/titles that are considered part of the Body Function entity by referring to *Annotation of Headers and Titles*

4.2.6 Possible Body Function Entity

Documentation that appears to have body function information in it but is formatted inconsistently with the data presented or appears OCR'd poorly such as the examples below.

Example 1: Muscle testing:

Hip flexion

Knee extension

Knee flexion

Ankle Dorsiflexion

Ankle Plantarflexion

Left

4/5

4/5

4/5

4/5

3/5

Right

4/5

4/5

4/5

5/5

4/5

This is likely body function either in tabular format or OCR'd with different formatting so we cannot assume which strength entity type goes with which qualifier with certainty.

Example 2: Shoulder

**** Measured in Degrees ****

AROM PROM MMT

Right Left Right Left Right Left

Shoulder - Flexion 3+/5

Shoulder - Extension 4/5

Shoulder - Abduction 3+/5

Shoulder - Int Rot 90 4/5

Shoulder - Ext Rot 90 4/5

This is also likely body function in tabular format so we cannot assume which strength entity type goes with which qualifier with certainty.

4.2.6.1 Attributes of the Possible Body Function Entity

There are NO attributes for the Possible Body Function Entity

5 Sample Difficult cases

5.1 Double annotation when Rom and Strength form the same body function mention

Please annotate the following Body Function Type mentions TWICE as both ROM and Strength since the information in the text refers to both.

Example1: Making a fist.

In this example the action of making a fist requires both having range of motion of the fingers to create a fist AND some degree of strength to move the fingers. Therefore, we understand the whole span of text can be annotated twice for Body Function Type as both ROM and Strength.

5.2 Extraocular movements derivatives

Extraocular movements are a function related with range of motion and strength that can be represented in the text in different ways. The following are examples of these representations and an explanation on how capture them.

Example 1: Extraocular movements are intact

In this case we can make a clear distinction of different sub-entities and annotate “extraocular” as a Body location and “intact” as a Qualifier (with a value of 1). The word “movements” will be annotated twice, once as Body Function Type with a value of “ROM”, and again as Body Function Type with a value of “Strength” because the movements represent both.

Example 2: EOM intact

In this scenario we can still annotate “intact” as a Qualifier (with a value of 1) but the abbreviation “EOM” that stands for “Extraocular movements” is impossible to separate. In this case we will annotate the entire span of text “EOM” as Body Location, Body Function Type with a value of “ROM”, and Body Function Type with a value of “Strength”, for a total of 3 overlapping annotations.

Example 3: EOMI

In this case, like the previous example, we cannot separate the abbreviation into its components so the entire span of text will be annotated 4 times, as Body Location, Body

Function Type with a value of “ROM”, Body Function Type with a value of “Strength”, and lastly as Qualifier with a value of 1.

5.3 ROM vs. Strength inference.

Flexion vs. flexors, extension vs extensors, rotation vs rotators, inversion vs. invertors, eversion vs. evertors, and adduction vs adductors are examples of ROM vs Strength inference. When encountering words like flexors or adductors, for example, we refer to strength because it is the representation of a group of muscles. Examples like flexion and extension on the other hand will refer to range of motion.

Example 1: Weak hip abductors. (Body Function Type: Strength)

Example 2: Limited hip abduction (Body Function Type: ROM and Strength)

5.4 Negation annotations.

There are instances where the negation of a Body function sub-mention cannot be separated in order to capture the true value of the mention.

Example 1: Negative for weakness.

In this example the “negative” actually provides the information to capture the whole span of text as Qualifier +1.

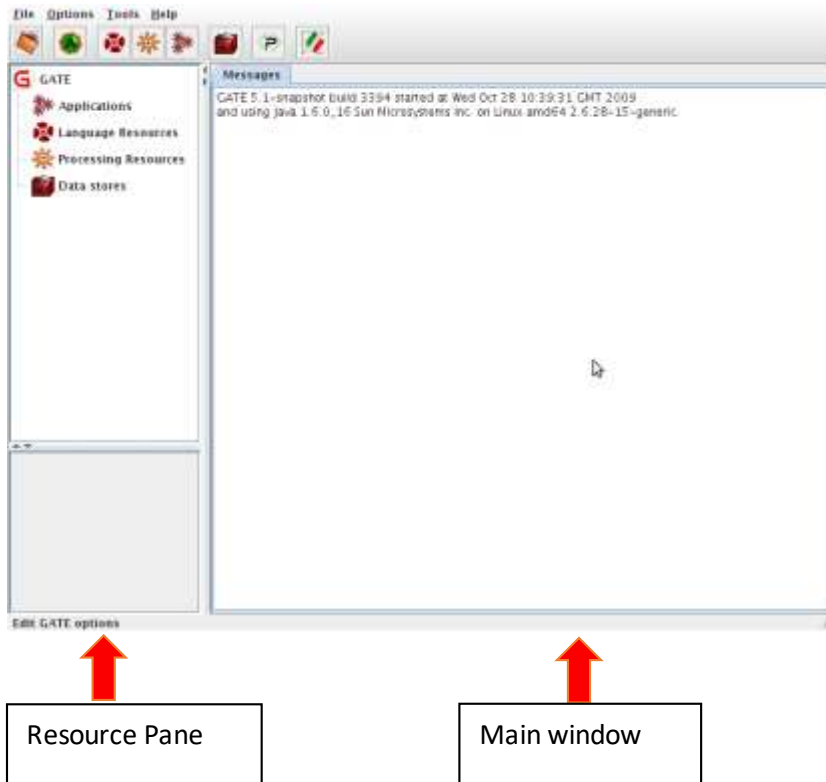
Example 2: Negative for dizziness or focal weakness

6 Software

Gate Developer annotation tool is used for annotation. Detailed instructions regarding uploading annotation schema and documents and exporting annotated documents are provided in the following sections.

6.1 The Gate Developer Main Window

When the user launches GATE software, a main window of GATE developer will be displayed as in the picture below.

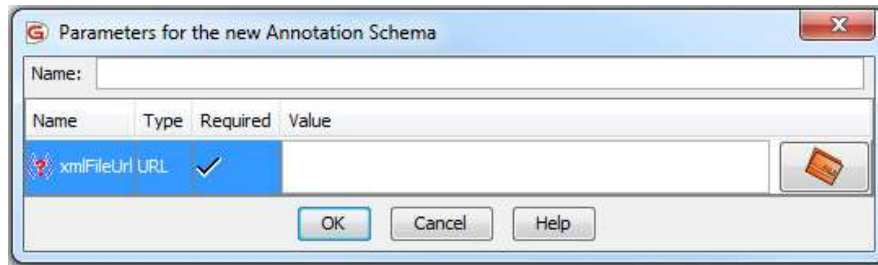


For this tutorial, we focus on the necessary steps to annotate medical files pertinent to this task; for a more detailed explanation of all the components of GATE developer, please refer to the user guide: <https://gate.ac.uk/sale/tao/split.html>.

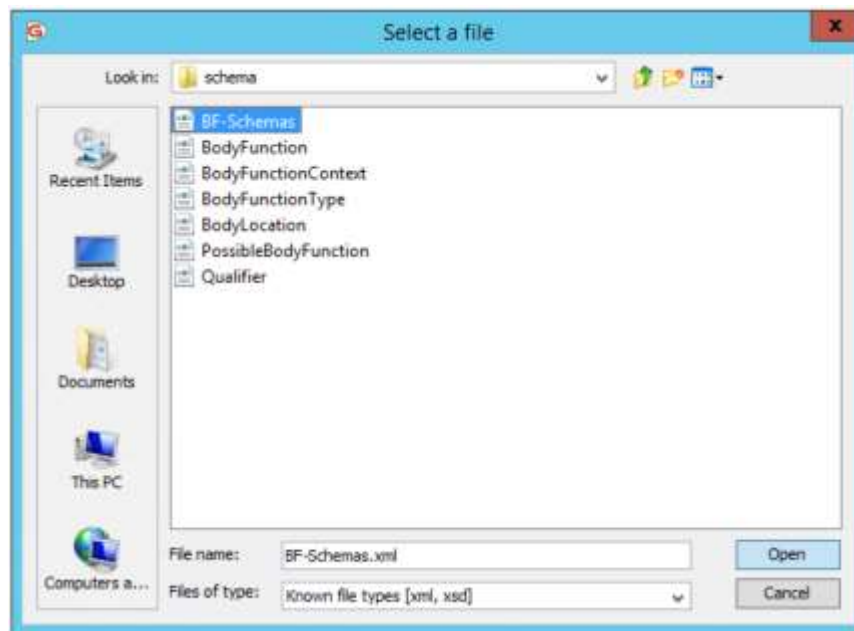
6.2 Loading and viewing documents

On the left side of the screen, we can see the resources pane that contains “Language Resources” and “Data stores”, both of which we will use for this task.

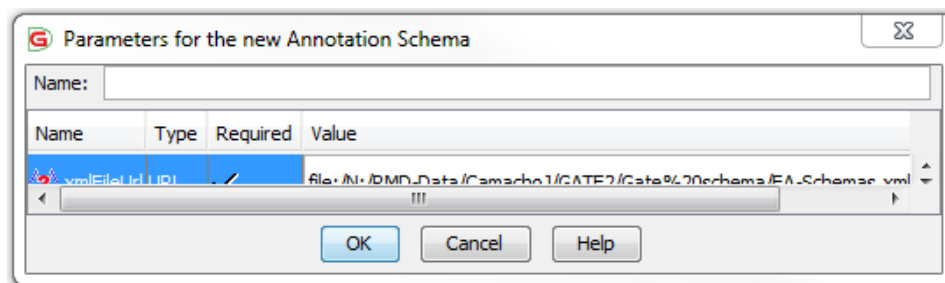
- Right click on “Language Resources” in the resource pane, select “new” and “annotation schema”. You will then see the window in the picture below. Select the “folder” icon to access the schema files.



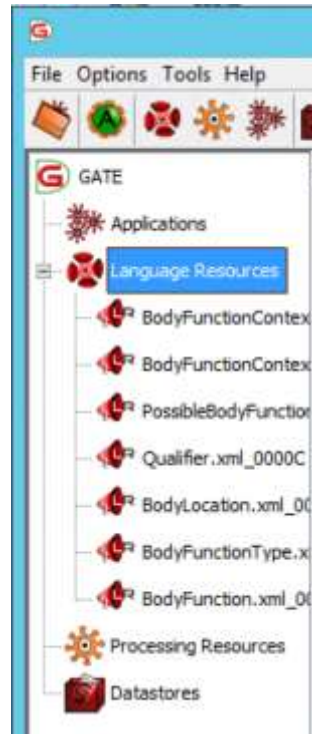
- Navigate to the folder with the schema you want to select and choose “BF-schemas” then click the “Open” button like in the picture below.



- Click OK to load the Schema.

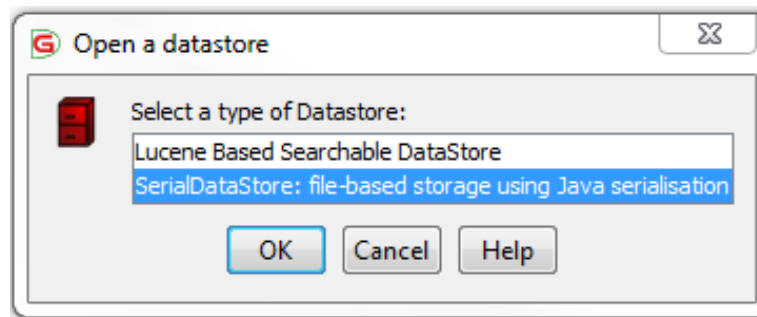


- Once the BF-schemas file is loaded you should be able to see the schemas in the Language Resources pane like in the picture below.



6.3 Loading a Schema Type

- Now it is time to load the files we will annotate. To do this, right click on “Datastores” in the resource pane and select “open Datastore”. You should be able to see a window like the picture below where you will select the option “SerialDataStore” and click the OK button.



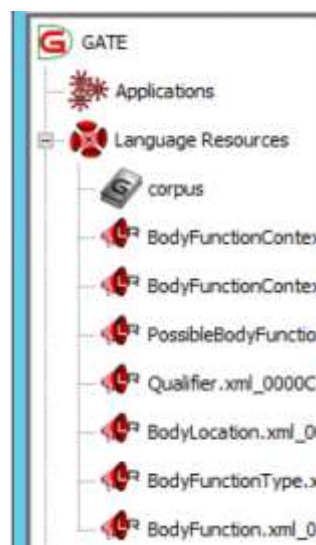
- Select the Datastore you are annotating and click the “Open” button. This will load the Datastore, and you will be able to see it on the Sources pane under “Datastores”. In this case, we selected a Datastore named “Jcm”, and it should look like this:



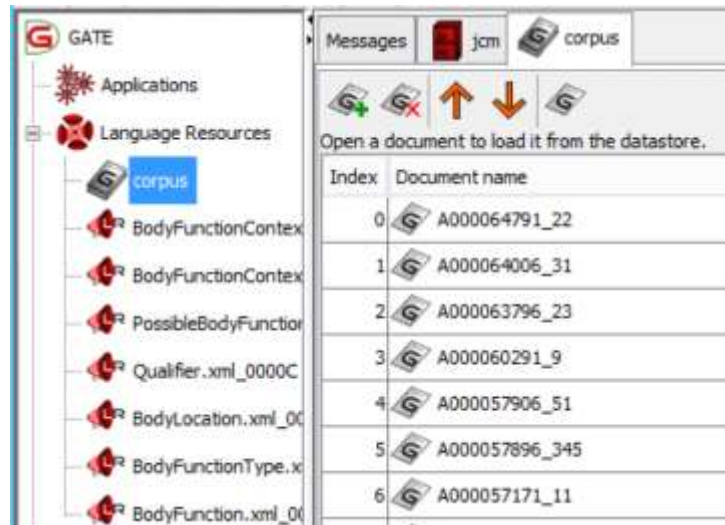
- Double click on the Datastore you just loaded (e.g., “Jcm”), and you should be able to see the folders “GATE Document” and “GATE Serial Corpus” in the main window on the right like in the picture below.



- You can load the entire corpus by double clicking the “GATE Serial Corpus” folder and then double clicking on the desired file.
- The “GATE Document” folder option contains all the individual files in case you would like to annotate a specific file.
- Selecting an option in the “GATE Serial Corpus” folder in the main window (e.g., “Corpus”) will load the entire corpus into the “Language Resources” pane.



- By double clicking this corpus in the “Language Resources” pane, you will be able to see all the files from that corpus in the main window like in the picture below. Double click on the files you want to open.

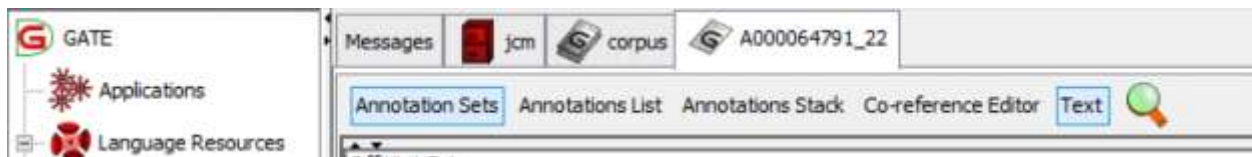


6.4 Creating and editing annotations

a) Creating a schema type:

The following steps will need to be done every time we select a new file.

- Once you have selected the file you want to annotate, make sure you select the option “Annotation Sets” in the upper-left corner of the main window. Upon selecting the “Annotation Sets” tab, another pane will open on the right side of the main window. This is the markup pane.



- At the bottom of the markup pane, type “BF” and click “New” to create a schema type like in the example below. Then you will see the schema type “BF” at the top of the markup

pane. Make sure you select the “BF” schema type before you start your annotations; by doing so, everything you annotate will be within this schema type.

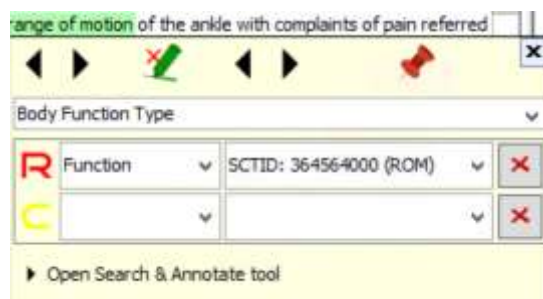


b) Annotation of different color-coded mentions:

- To create an annotation, select the span of text you want to annotate and place the cursor on this selection. Select “Create new annotation” and you will see a window like the one in the example below.

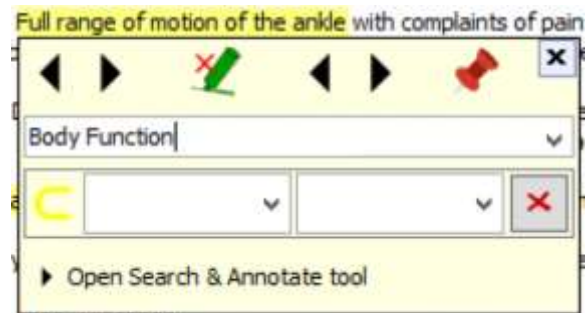


- In this window select the entity to annotate from the drop-down menu (“Body Function Type” in this example) and more options will be displayed for this particular entity (if it has any). Select accordingly.



6.5 Examples of Annotating Different Mentions

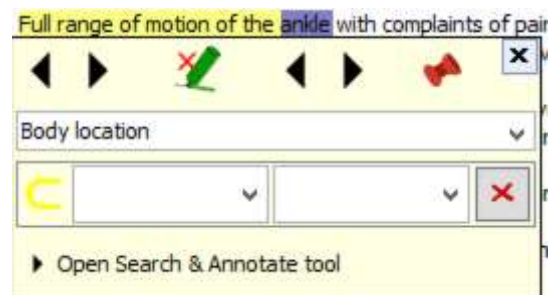
Example 1: Capturing a Body Function mention.



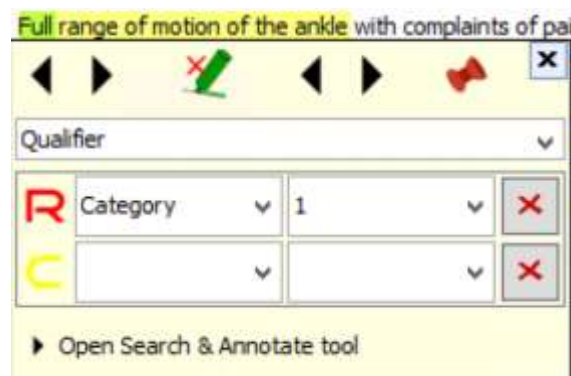
Example 2: Capturing Body Function Type mention.



Example 3: Capturing Body Location mention.



Example 4: Capturing a Qualifier mention.



Example 5: Capturing a Body Function Context mention.

DTR: 2+ biceps, triceps, knees and anides tdiateral. Toes down.



Body Function Context 

 Context Type 

BF Type: Reflexes 







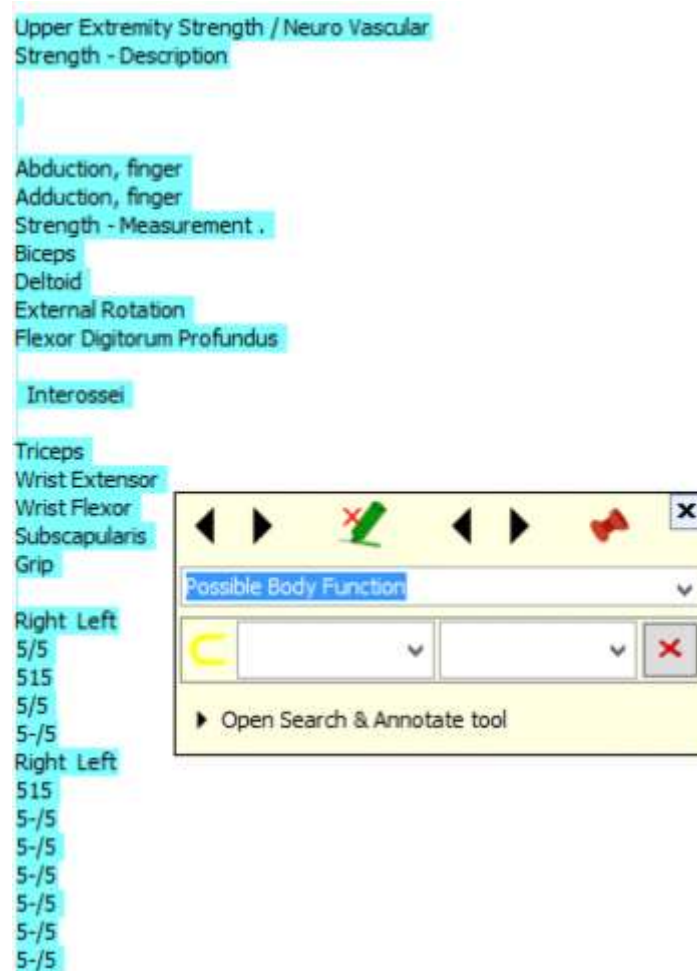




 Open Search & Annotate tool

Example 6: Capturing Possible Body Function mention.

Please describe patient's ability to perform activities of daily living such as eating, dressing.



6.6 Saving GATE software

- Save your progress before exiting GATE: Right click on the name of the file you are annotating (it could be at the top of the main window or in the resources pane) and select the option "Save to its Datastore." We recommend you do this for every file after finishing with the annotations and before continuing to the next one.

6.7 Exporting Annotation Data

- You can also export annotated documents by right clicking on the name of the corpus or file and selecting "Save as: GATE XML" followed by the selection of the folder where you want to save it.

6.8 Exiting GATE Software

- Exit the software by selecting “File” in the upper-left corner and then select “exit GATE”. It is important that you always exit this way after saving your documents since it is possible that information will be lost if you just click on the “X” button in the upper right corner.

7 References

- American Academy of Orthopaedic Surgeons & American Academy of Orthopaedic Surgeons. (1965). Joint Motion (12th ed.). Churchill Livingstone.
- Kendall, F. P., McCreary, E. K., Crosby, R. W., Abeloff, D. K., Gregerman, M. B., Loechel, W. E., & Krause, C. C., Jr. (1983). *Muscles: Testing and Function, 3rd Edition* (3rd ed.). Williams & Wilkins.
- Merriam-Webster. (n.d.). Reflex. In Merriam-Webster.com dictionary. Retrieved July 7, 2022, from <https://www.merriam-webster.com/dictionary/reflex>
- Walker, K. H. (1990) Chapter 72, Deep Tendon Reflexes, in Walker, K. H, Hall, D. W., & Hurst, W. J. (Eds) *Clinical Methods: The History, Physical and Laboratory Examinations* (3rd edition., pp. 365-368).

8 Appendixes

Appendix A: Common Acronyms and Abbreviations observed in our annotation work.