

合成プロセスアノテーション作業の概要

1 ルール概要

1.1 対象

アノテーション作業の対象は, "固体電解質" を "固相または液相法" を用いて "合成 (薄膜を除く)" している論文とする.

以下の論文はアノテーション作業の対象外とする.

1. 固体電解質を合成していない論文 (例. シミュレーションに関する論文)
2. 薄膜合成法を用いている論文 (当該論文に使用される表現の一例 "thin film" etc.)
3. 気相法を用いている論文 (当該論文に使用される表現の一例 "chemical vapor deposition:CVD", "physical vapor deposition:PVD", "atomic layer deposition:ALD", "pulsed laser deposition:PLD", "reactive sputtering", "magnetron sputtering", "molecular beam epitaxy:MBE", "metal-organic vapor phase epitaxy:MOVPE", "ion plating", "vacuum evaporation" etc.)

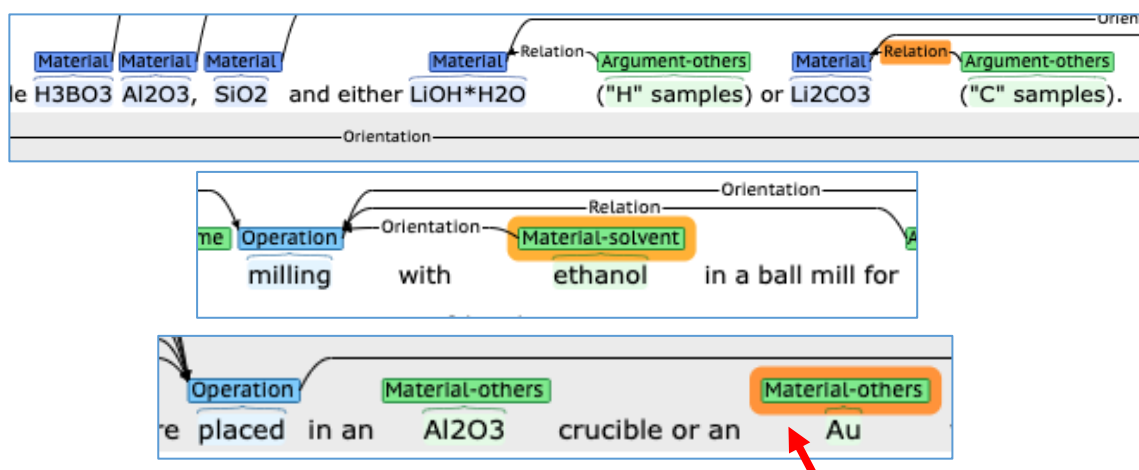
1.2 スパンに対するラベル

テキスト中の単語や, フレーズに対してアノテーションを行うためのラベル.

1.2.1 mat ラベル

1. テキスト中に出現するすべての組成式，または，溶媒に「Material」ラベルを付ける。

「Material」ラベルは「Material-start」（出発材料），「Material-intermedium」（中間材料），「Material-final」（最終材料），「Material-solvent」（溶媒），「Material-others」（その他材料）のラベルをつける



Material が他と関連付かない場合は矢印なしでも良い

図 1 Material ラベル例

※「Material-intermedium」（中間材料）に矢印ラベル「Orientation」を付けないと

合成プロセスが表現できない文の一例: after milling, 80 % of the intermediate material is annealed at 500 K and 20 % was heated at 200 K.

上記のように、特定の割合で分割された中間材料「intermediate material」に対して異なる処理を実施する際は、「milling」→「intermediate material」→「annealed」と、「milling」→「intermediate material」→「heated」のように中間材料を分岐点とする2つの経路をアノテーションする。

2. Start 組成, Target 組成, 溶媒などのいずれかを意味していても、含有元素が分からない単語はアノテーション対象外（例 $\text{Al}(\text{OH})_3$ が Start 組成であり、別の箇所では“the hydroxide”と書かれている場合、 $\text{Al}(\text{OH})_3$ を指すと理解はできるが、アノテーションしない）
3. Start 組成, Target 組成, 溶媒のいずれかを意味する抽象的な単語や指示代名詞はアノテーションの対象外。例を挙げると、Final sample, it, this, the powder, the compound, initial powder, Mother powder など
4. Al-substituted LLZ のように含有元素が分かる説明文付きの単語は mat としてアノテーション

5. 論文中の Abstract と Experimental にある文章がアノテーションの対象

表 1 mat ラベルの例

mat の例
Li ₂ S
P ₂ S ₅
CaYb ₂ S ₄ -2mol%Yb ₂ S ₃ (配合割合を含む組成)
LATP(略式記号で記述された組成)
Acrylamide(単語で記述された組成)
Al-substituted LLZ (説明文込みの組成)
Li _{0.5} La _{0.5} Ti _{1-x} Zr _x O ₅ (変数 x が存在する組成) (x=0,1... の部分は arg にアノテーション)

1.2.2 ope ラベル

1. 処理をアノテーションする為のラベル
2. ope ラベルの例を表 2 に示す.

表 2 ope ラベルの例

ope の例
milling
melting
mixing
heating
sintering
annealing
calcine
pressing
ground

1.2.3 arg ラベル

1. mat と ope と arg の条件や状態などをアノテーションする為のラベル
2. arg は前置詞, 副詞, 形容詞を含んでも良い
3. 前置詞を含む arg の例 At temperature higher than the crystallization temperature など
4. どこでスパン (範囲) を区切ればいいのか迷う場合は作業者自身の裁量でスパンを決定すること。

ただし、スパンをあまり長くなりすぎないように、適度に分割しラベルを付与する

5. arg の単語が持つ役割に応じて付与するラベルが変わる. arg ラベル一覧を表 3 に示す

表 3 arg ラベル一覧

argラベルの種類	役割	例
arg-time	時間条件	10 h, over night
arg-temp	温度条件	80 °, Room temperature(RT), Crystallization temperature
arg-rot	回転速度条件	370 rpm
arg-press	圧力条件	20 MPa, at a pressure of 50 MPa
arg-air	空気条件	under dry argon atmosphere, in primary vacuum, under pure F2 gas stream, in air
arg-others	その他条件	Merck, purity 99 %, bu stirring, molar ratio of 75:25

Argument-equipment ラベル

1. 設備や装置名や型番などをアノテーションする為のラベル.

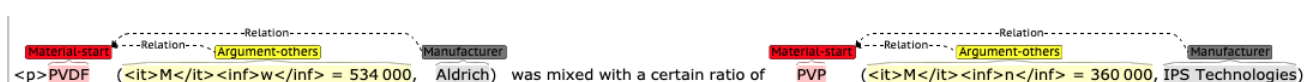
2. NMR analyses were performed at room temperature in deuterated DMSO on a Varian

UnityPlus 500(Equipment) spectrometer at 500 MHz

Argument-manufacturer ラベル

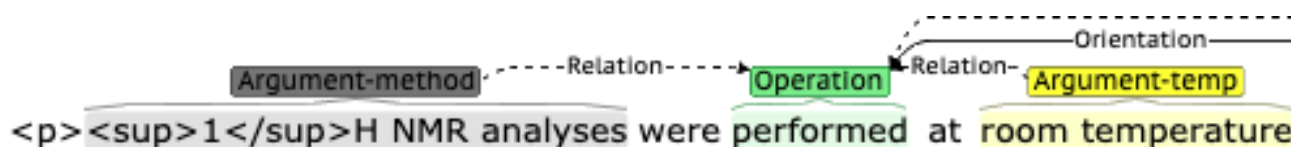
1. 製造会社をアノテーションする為のラベル.

2. mat-start と manufacturer へのスパンラベルと関係ラベル rel の例



Argument-method ラベル

1. 合成や計測などの手法をアノテーションするためのラベル



1.3 関係に対するラベル

2つ以上のフレーズの間の下記で定義される関係が存在する場合は、2つのフレーズの間
に矢印を張る形でアノテーションを行う.

1.3.1 ori ラベル

1. ori ラベルは処理の工程をアノテーションする為のラベル

2. mat - ope(mat から ope に ori すること), ope - ope(ope から ope に ori すること),
ope - mat(ope から mat に ori すること)のアノテーションが可能

1.3.2 rel ラベル

1. rel ラベルは mat と arg あるいは ope と arg の関係をアノテーションする為のラベル
2. mat-arg(mat に arg を rel すること), のアノテーションが可能
3. arg と arg 間の rel ラベル付与は不可とする

1.3.3 abbreviation ラベル

1. 材料名や装置、手法などが略式記号で記述されるときに対応関係を記録するためのラベル。例えば, “Li5La3Ta2O12 (LLTa) was prepared by a solid-state reaction.”のように, 括弧を用いて先頭 material 名称が略されている場合は, Li5La3Ta2O12 と LLTa のそれぞれに Material-final のラベルを付与し, 2 つの単語の間に, abbreviation の関係ラベルを張ることとする.

2. arg-equipment の途中で略されている場合のアノテーション例



3. Abbreviation ラベルは, 略語と正式名称の距離が遠いときにはつけなくても良い。例え

ば、1行目に括弧書きで略字の説明があり、2行目以降その略語が使われている場合、2行目以降にはアノテーションしなくてもよい

1.2.7 特性値のラベル

1. 「Value-conductivity」 「Unit-conductivity」 「Value-activation」 「Unit-activation」

Abstract と Experimental のテキストには、次の図2のように特性値「conductivity」と「activation energy」が記載されていることがある。特性値は、「 $1.09 \times 10^{-3} \text{ Scm}^{-1}$ 」のように、値と単位のペアで記載されているため、特性値が「conductivity」のとき、値には、「Value-conductivity」のラベルを付与し、単位には、「Unit-conductivity」のラベルを付与すること。特性値が「activation energy」のとき、値には、「Value-activation」のラベルを付与し、単位には「Unit-activation」のラベルを付与すること。

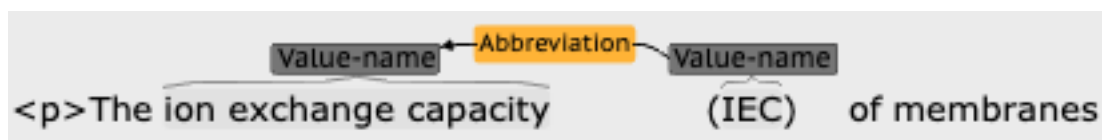
1	Abstract
2	Mechanical milling (MM) has been used to prepare the nanosized Li_{1.4}Al_{0.4}Ti_{1.6}(PO₄)₃ (denoted LATP) glassy powders, which was converted into glass-ceramics through thermal treating at 700–1000°C. The XRD, TEM, FESEM and ac impedance techniques were used to characterize the products.
3	The results showed that completely amorphous products were prepared by MM for 40h, and single-phase LiTi ₂ (PO ₄) ₃ -type structured glass-ceramics were obtained by further heat treatment.
4	The lithium ion conductivity of the glass-ceramics increased with the growth of the crystalline phase and decrease of the grain size.
5	The highest bulk conductivity (σ _b) of 1.09×10⁻³ S cm⁻¹ with an energy of activation as low as 0.28 eV was obtained at room temperature for the specimen treated at 900°C for 6h .
6	The high conductivity, easy fabrication and low cost make the LATP glass-ceramics promising to be used as inorganic solid electrolyte for all-solid-state Li-ion rechargeable batteries.

図2 特性値の記載例

2. 特性値「conductivity」と「activation energy」の測定条件が記載されている場合は、1.2.3で説明した arg ラベルを用いてラベル付けを行うこと。

3. Value-name ラベル

特性値の名称をアノテーションするためのラベル



2 ルール詳細 1(ラベル例外)

2.1 過剰量へのアノテーション

1. Stoichio はアノテーションしない

2. Excess は mat の arg とする

3. 途中の ope から過剰量が追加される場合は, 図 3 のようにアノテーションすれば,

mat-st を途中 (赤字 の ope) で過剰量添加したことが分かる

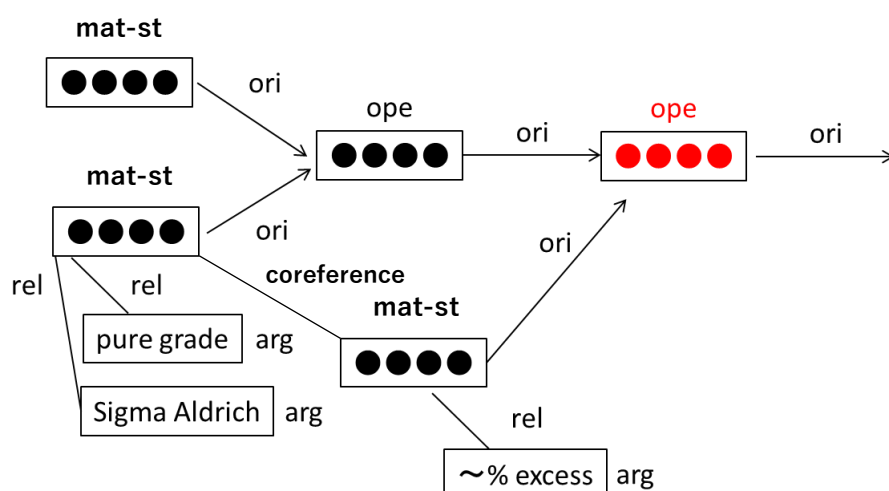


図 3 途中の ope から過剰量が追加される場合のアノテーション

2.3 一連の処理に対するアノテーション

1. 一連の処理全て（複数の ope）に係る arg がある場合, いずれかの ope ひとつに rel する
2. 例 All operation executed in Ar atmosphere など 3. ope 全部に arg を rel してもよい

3 ルール詳細 2(合成手順の合流と分岐)

合成法によっては Target 組成へ至るまでのプロセスが分岐し, 分岐後合流する場合があります.

3.1 ope に係る arg 分岐の場合

図 4 に arg の違いによって, 最終生成物が異なる合成法にアノテーションを行った例を示す.

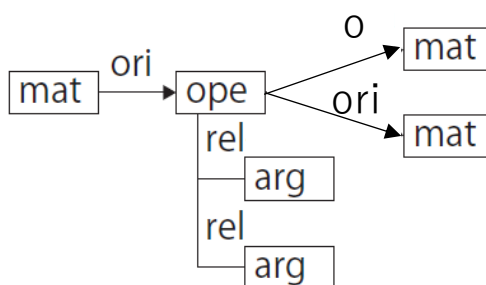


図 4 arg 分岐の場合

3.2 X と同様に Y を合成したと記述されている場合

X と同様に Y を合成したと記述されている場合, Y を合成するときの条件が省略されることがある.

図 5 に A から B を作ったときと同様の作り方で C から D を作ったテキストのアンテーションを示す.

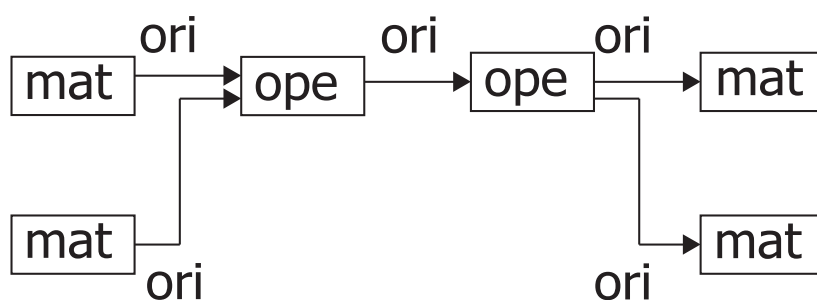


図 5 C から D は A から B と同様に作ったという記述の場合

3.3 Target 組成の状態の違いによる合流

複数の合成法が単一の Target 組成に合流する場合

複数の合成法が glass か glass ceramic の Target 組成へ合流する例を図 6 に示す.

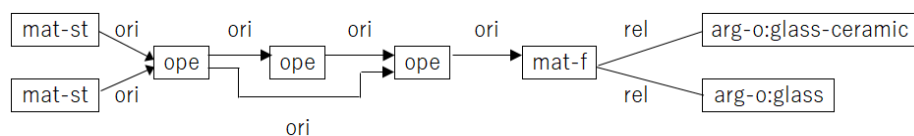


図 6 glass もしくは glass ceramic で分岐する場合

同様に, solid もしくは sol-gel の Target 組成に合流する例を図 7 に示す.

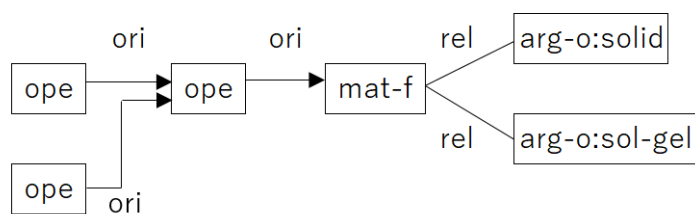


図 7 solid もしくは sol-gel で分岐する場合

4 アノテーションの注意事項

4.1 基本的に組成+括弧で書かれているフレーズの括弧の中身はどんな情報であっても

arg-*のアノテーションをおこなう。括弧の中身はどんな情報であれ、組成に関連付く可能性が高いため。

例：Li₂CO₃ (A.R.) → Li₂CO₃ に m-st, A.R. に a-o をつける

例：heated under flowing nitrogen gas (50 ml min⁻¹, ultrahigh purity, Airgas).

→ heated に ope, flowing nitrogen gas · 50 ml min⁻¹ · ultrahigh purity · Airgas にそれぞれ a-o をつける。

4.2 arg-arg 間の rel はルールの範囲外

一部をアノテーションしない、もしくは近くの ope に arg を関連付けるなどで対応する。

判断が難しい例：

Water was purified by ELIX system (Millipore).

→ a-o の purified by ELIX system にメーカー名 Millipore が関連付いているため、arg-arg と付けることも考えられるが arg-arg はルールに無いため、Millipore にはラベルを付けずに、purified by ELIX system だけを次の mat に関連付ける。

4.3 テキスト内の2つ以上の表現が同じものを参照する時（共参照）は，下のように coreference のアノテーションを行う．

例：train/1-s2.0-S0013468602007491-main

(C₂H₅)₄NClO₄ (TEAP) was obtained by metathesis of (C₂H₅)₄NBr (Reachim) with HClO₄ (Aldrich) in water(mat-so).

The products were recrystallised twice from water(mat-so) and dried under reduced pressure at 100 degC for 24 h.

Water(mat-st) was purified by ELIX system (Millipore).

The procedure of sol-gel electrolyte preparation was based on that described in Ref. [5].

Three milliliters of the hydrolysed sol was prepared by mixing TMOS with water in 1:2 molar ratio (2.4 ml TMOS and 0.6 ml H₂O(mat-st)).

4.4 動詞と副詞・形容詞を分けてアノテーション

後段の処理では，Material は名詞，Operation は動詞，Argument は形容詞あるいは副詞と仮定して抽出するので，できる限りで上の仮定で，ラベル付与する．

*“... was dehydrated by heating at 110 degC overnight in primary vacuum.” の文のように，dehydrated と heating のどちらに operation をつけていいか迷う場合は，

((“dehydrated”, Operation), (“by heating”, Argument-others)) のように動詞と前置詞+
動名詞を分けてラベル付けする.

* “The powders were then uniaxially pressed into pellets …” の文のように,

((“uniaxially”, Argument-others), (“pressed”, Operation))か, (“uniaxially pressed”,
Operation)か迷う場合

この場合は, (“uniaxially”, Argument-others), (“pressed”, Operation)のように動詞と副
詞 (あるいは形容詞) は分けてラベル付けする.

* “cold-press”のように動詞と副詞・形容詞がくっついて, 1 フレーズになっている場
合は, 1 フレーズに 2 つのラベルをつけるのではなく, “cold-press”を Operation と
してラベル付けする.

4.5 溶媒として頻出する distilled water, deionized water, deionized (DI) water, DI water, など
は, それぞれまとめて m-so としてラベル付けする.

ただし, (水以外も含めて) 蒸留が操作の一部である場合は distilled(ope)または
distilled(arg-o)として独立させてラベル付けする.

4.6 必要に応じて sol, gel, powder, aq, glass, glass ceramic など, 材料の状態を表すフレー
ズを arg-o として mat に関連付ける.

PDF から TXT へ変換するまでの手順

1.1 PDF からテキスト抽出

論文 PDF の「Experimental」またはそれに相当する章から『合成プロセス（“固体電解質”を“固相または液相法”を用いて合成するもの（薄膜を除く））』に関するテキストを抽出する。 ※「電極の作成」や「複合材料の調整」は合成プロセスの対象外とする。

テキスト抽出は PDF ファイルからコピー＆ペーストでテキストエディタやメモ帳などに貼り付ける方法で行う。

1.2 テキストの修正

貼り付けた文章テキストに対して、以下の変換ルールに従ってテキストの修正を行う。

文字変換のルール表

ルール	変換前	変換後
組成式において上付き文字， 下付き文字は半角数字で表記	H2O	H2O
値における上付き文字， 下付き文字の表現	10^8 , 10^{-8} , T_x	10^8 , 10^{-8} , T_x
組成式からスペースを削除	(NH4)2C2O4	(NH4)2C2O4
センタードットは*(アスタリスク)に変換	Co(COOH)2 · 4H2O	Co(COOH)2*4H2O
°CはdegCに変換	100 °C	100 degC
値と単位の間にはスペースを挿入	40min, 80degC	40 min, 80 degC
値が1つのフレーズとして認識できる場合は，スペースを削除	T > 800 degC	T>800 degC
改行の際に発生するハイフネーション及びスペースを削除	oxida- <改行> tive	oxidative
2バイト以上の文字は1バイト文字に変換	ω , α	ω , α
商標マークなどの削除	®, ©, ™	変換無し
プラスマイナス記号	±	+-
不等号記号	≥, ≤	>=, <=
各行末の改行コードは削除	各行末に改行コード	改行コードの削除 (以下の例を参照)
単語と単語の間のスペースが消失した場合はスペースを挿入	anddried	and dried
文末のピリオドで改行	改行無し	改行する (以下の例を参照)

《変換前》

LAGP solid electrolyte was synthesized using a conventional
solidstate reaction method. Li₂CO₃ (99%, Alfa Aesar), Al₂O₃
(AR, Nanjing Chemical Reagent Co., P. R. China), . . .

行末の改行コード

文末のピリオド

《変換後》

LAGP solid electrolyte was synthesized using a conventional solidstate reaction method.
Li₂CO₃ (99%, Alfa Aesar), Al₂O₃ (AR, Nanjing Chemical Reagent Co., P. R. China), . . .

brat を用いて TXT へアノテーションする手順

1.1 テキストアノテーションツール”brat”を開く

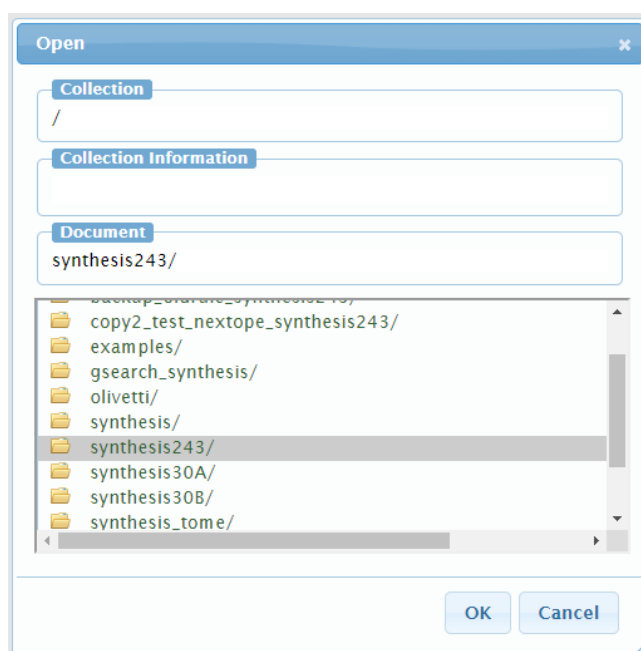
ブラウザで URL(管理者より通知)に接続し、画面右上の brat ボタンをクリックし、以下の ID と PW でログインする.

ID: (管理者より通知)

PW: (管理者より通知)

1.2 brat 上の synthesis フォルダを開く

マウスカーソルを左上部に移動させ、Collection タブを選択し、作業フォルダ「synthesis243」を選択し、OK ボタンをクリックし、作業フォルダを開く.



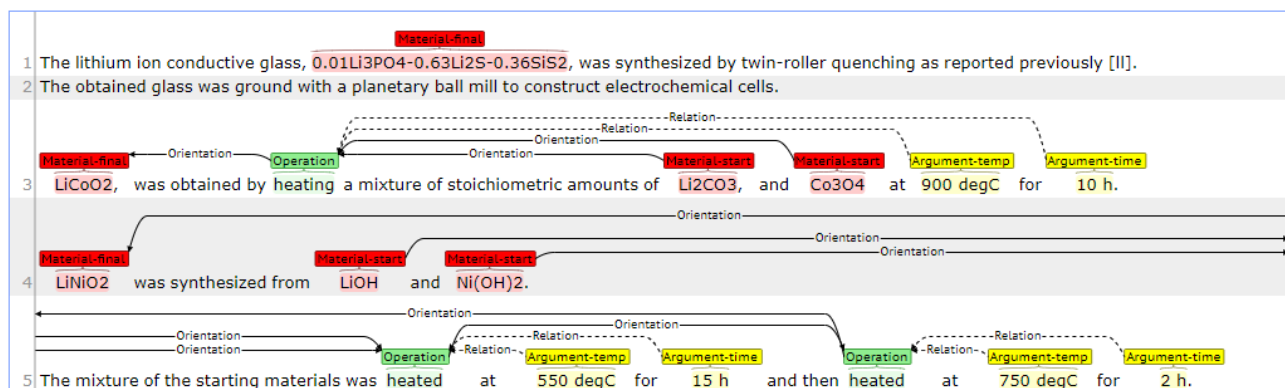
1.3 brat 上でアノテーションを行う

単語のアノテーション

単語スパンをドラッグ&ドロップで選択. ラベルを選択するポップアップウィンドウが出現するので、ラベル名を選択.

関係のアノテーション

あらかじめタグ付けした単語スパンを選択し、ドラッグ&ドロップすると、矢印が出現する. 矢印の先を対象の単語スパンに置くと、ポップアップウィンドウが出現するので、ラベル名を選択. 一部、アノテーションが許可されていない関係が存在. 下記に矢印連結可否ルールを示す.



brat でのアノテーション例

矢印の「元」	→	矢印の「先」	ラベルの種類	複数の「元」から矢印を受けて良いか	複数の「元」に矢印を発して良いか
Material	→	Material	Coreference	○	○
Material	→	Operation	Orientation	○	○
Material	→	Argument	不可	—	—
Operation	→	Material	Orientation	○	○
Operation	→	Operation	Orientation	○	○
Operation	→	Argument	不可	—	—
Argument	→	Material	Relation	○	○
Argument	→	Operation	Relation	○	○
Argument	→	Argument	不可	—	—

矢印連結可否ルール