

Linee guida per l'annotazione delle 5W su notizie di furto

VERSIONE 2.1

14-02-2023

Giovanni Bonisoli*, Maria Pia di Buono^,
Gennaro Nolano^, Laura Po*, Federica Rollo*

*Università degli studi di Modena e Reggio Emilia

^Università di Napoli L'Orientale

1. Obiettivo	3
2. Domande di competenza	3
3. Cosa annotare	3
3.1. Quali documenti annotare	3
3.2. Span	5
3.3. Sinonimi e termini specifici	5
4. Annotazioni multiple	6
4.1 Stessa etichetta su span diversi	6
4.2 Stesso span con etichette diverse	7
5. Relazioni	7
6. Oggetto rubato (OBJ)	7
7. Autore e Gruppo di autori	9
7.1. Autore (AUT)	9
7.2. Gruppo Autori (AUTG)	12
8. Vittima, Gruppo di vittime e Parte Lesa	13
8.1. Vittima (VIC)	13
8.2. Gruppo Vittime (VICG)	15
8.3. Parte lesa (PAR)	15
9. Luogo (LOC)	16
10. Tempo	18
11. Come	18
Appendice A - Docciano	19

1. Obiettivo

Queste linee guida definiscono le regole da seguire per l'annotazione di testi di articoli di giornale riguardanti eventi criminali al fine di evidenziare gli elementi che caratterizzano l'evento descritto.

In particolare si fa riferimento alle notizie contenute nel dataset ITALIAN CRIME NEWS.¹

Il documento mira a fornire informazioni sul compito di annotazione delle 5W+1H: WHAT, WHERE, WHO, WHEN, WHY, HOW.

Nel seguito si descriveranno specifici concetti per identificare le 5W+1H su notizie che trattano di eventi di tipo furto.

Tra i tool disponibili per l'annotazione è stato scelto Doccano,² di conseguenza nel seguito vengono presentati alcuni esempi di annotazione attraverso tale tool.

2. Domande di competenza

Per l'identificazione di eventi nel testo delle notizie e per la valutazione della somiglianza di due notizie, è necessario identificare, per le tipologie di eventi FURTO:

- **OGGETTO RUBATO** (cosa è stato rubato)
- **AUTORE** e/o **GRUPPO AUTORI** (colui o il gruppo che ha compiuto il furto)
- **VITTIMA**, **GRUPPO VITTIME** e/o **PARTE LESA** (la persona o l'organizzazione lesa)
- **LUOGO** (dove ha avuto luogo il furto)

La risposta a queste domande deve essere sempre una parte del testo della notizia e può essere contenuta su più porzioni di testo, denominati "span".

La risposta a una o più delle domande sopra riportate può non essere presente nella notizia.

Si suggerisce di leggere prima tutto il testo della notizia, e poi procedere all'annotazione valutando con precisione gli elementi del testo da annotare.

3. Cosa annotare

3.1. Quali documenti annotare

Se un evento di furto avviene in un singolo luogo, per mano dello/degli stesso/i autore/i e in un singolo momento allora è definito come evento singolo di furto perché corrisponde a una singola azione di furto.

Devono essere annotati solo eventi singoli.

¹ <https://paperswithcode.com/dataset/italian-crime-news>

² <https://github.com/doccano/doccano>

In alcuni casi è possibile che oltre a descrivere l'evento si faccia riferimento ad altri eventi già accaduti (eventi di contorno) che hanno qualcosa in comune con l'evento descritto. Se i riferimenti agli eventi di contorno si limitano a un paio di frasi, escludere tali riferimenti dall'annotazione e limitarsi all'annotazione dell'evento principale.

Se il furto descritto nella notizia è un tentativo di furto, occorre annotare tutti gli elementi del furto presenti e inserire "tentativo di furto" come **commento a livello di documento** in Doccano (vedi Appendice A - Doccano).

Una notizia potrebbe riportare informazioni su eventi avvenuti dopo il furto quali:

- il ritrovamento di un oggetto rubato,
- la vendita della refurtiva (ricettazione).

In questi casi annotare soltanto gli elementi che descrivono il furto avvenuto e non il ritrovamento o la ricettazione (non è necessario aggiungere un commento).

Non devono essere annotate notizie in cui vengono descritte:

- altre tipologie di crimini, diverse dal furto;
- altre tipologie di notizie, come report statistici sui crimini;
- eventi multipli.

Per tali notizie non va eseguita l'annotazione ma deve essere aggiunto un commento a livello di documento:

- "tipologia diversa: ..." (indicare la vera tipologia, es. rapina, aggressione, frode);
- "eventi multipli";
- altro.

Cosa sono gli eventi multipli:

- tanti eventi scollegati descritti nella stessa notizia: più autori e/o più vittime e/o più reati e/o più luoghi eventualmente sparsi nel tempo
- tanti eventi collegati descritti nella stessa notizia: più reati (ovvero più azioni di furto) che possono avere in comune stesso autore, oppure stesso oggetto rubato, oppure stesso luogo. In questi casi almeno uno degli elementi: luogo, autore, momento varia, pertanto non è più considerabile come un singolo evento.

Esempio di Evento multiplo (eventi scollegati)

Era poco dopo l'ora di cena quando è stata rubata un'auto Ford X nel parcheggio del centro commerciale. Da mesi la zona è al centro di diversi furti di auto e bici. Due settimane fa il furto di una Fiat Panda proprio di prima mattina... Gli abitanti della zona sono preoccupati... I controlli nella zona si intensificano... Già un mese fa c'erano state diverse lamentele e tre denunce da abitanti del quartiere: il furto di un furgone, lo scippo di una signora in bicicletta e un appartamento messo a soqquadro in pochi minuti.

Esempio di Evento multiplo (eventi collegati): singolo luogo, stesso autore ma momenti diversi, le azioni di furto sono due, una successiva all'altra.

Ieri al centro commerciale, un uomo ha rubato un cellulare a una signora 30enne e subito dopo un portafoglio e una borsa a una modenese ottantenne.

Esempio di Evento singolo: singolo luogo, stesso autore stesso momento, due vittime, un'unica azione di furto.

Il ladro entrato nell'appartamento ha sottratto un computer, una tv e 200 euro. Ne hanno denunciato il furto i proprietari Mario Rossi e Anna Verdi.

Esempio di Evento singolo: singolo luogo, stesso autore stesso momento, con due oggetti rubati e due vittime, un'unica azione di furto

Il ladro entrato nell'appartamento ha sottratto un computer appartenente a una studentessa spagnola e un iPhone appartenente alla proprietaria di casa, Luisa Rossi.

3.2. Span

Con il termine “span” si indica una porzione di testo che può includere una o più parole.

Lo span da annotare deve riferirsi solo all'entità e non deve contenere articoli determinativi/indeterminativi o preposizioni, né spazi o punteggiatura all'inizio o alla fine.

Esempi (a destra della freccia viene sottolineato ciò che deve essere annotato):

- Una modenese → Una modenese
- A Modena → A Modena
- Il furto è stato commesso a San Cesario sul Panaro → Il furto è stato commesso a San Cesario sul Panaro
- Modenese di origini tunisine di quarant'anni → Modenese di origini tunisine di quarant'anni

Alcune **entità sono costituite da più parole**, in questi casi lo span da annotare è l'insieme di tutte le parole.

Esempi:

- carrozzeria di Ferrari e Giovanardi
- pizzeria “Re Pipin”
- centro commerciale Borgogioioso
- Playstation 4
- computer Apple
- medicinali per animali
- villetta a schiera

3.3. Sinonimi e termini specifici

Nel caso in cui all'interno della notizia siano presenti dei sinonimi che si riferiscono alla stessa entità è preferibile annotare la prima occorrenza.

La casa è stata messa a soqquadro [...] Rossella ha denunciato il furto della sua costosa
•LOC
macchina fotografica che i ladri hanno rubato nella sua abitazione.

Figura 3.3.1: Si annota “casa” poiché compare prima di “abitazione”.

Nel caso in cui siano presenti termini di cui uno più specifico dell'altro, annotare sempre il termine più specifico.

I ladri sono entrati in casa e hanno rotto la cassaforte [...] I proprietari sono sconcertati
perchè si tratta già del secondo furto nella villetta a schiera.
•LOC

Figura 3.3.2: Si annota “villetta a schiera” perché più specifico rispetto a “casa”.

In tutti i casi, sono da evitare annotazioni di termini generici:

- bottino, refurtiva, maltolto, oggetti, merce
- ladro, malvivente, malfattore, aggressore
- vittima, malcapitato/a, proprietario/a, titolare, signora/i/e.

Esempi: (viene sottolineato ciò che deve essere annotato)

“il ladro è stato riconosciuto essere un uomo di 50anni”

“il ladro è stato riconosciuto essere di 50anni”

“il bottino è stato trafugato da due ladri”

“il bottino è stato trafugato da due ladre”

“le vittime sono 3 signore”

“le vittime sono 3 signore ottantenni”

“le vittime sono 3 signore si tratta di donne ottantenni”

4. Annotazioni multiple

4.1 Stessa etichetta su span diversi

In questo caso, per annotazioni multiple si intende l'assegnazione della stessa etichetta a più span.

Le annotazioni multiple si possono verificare in due casi:

1- più span rispondono alla stessa domanda (es. cosa è stato rubato? un portafoglio, una borsetta):

2- più caratteristiche sono fornite per una entità (es. chi ha commesso il furto? un modenese residente a Reggio Emilia, di 46 anni)

Le annotazioni multiple acquisiscono significati diversi:

- **oggetto rubato**, si hanno più annotazioni ciascuna riferita a un diverso oggetto rubato e/o alla quantità (si veda sezione 6);
- **luogo del furto**, si hanno più annotazioni le quali si riferiscono tutte al luogo dove è avvenuto il furto con una granularità differente, es. riferendosi al luogo generico (es. appartamento, abitazione, supermercato, nome del negozio es. “coop Grandemilia”) e/o all'area geografica (via, città etc.) (si veda sezione 9);

- **autore/vittima**, si hanno più annotazioni ciascuna riferita a un diverso autore/vittima e/o alle informazioni aggiuntive di ciascun autore/vittima (si veda sezione 7.1).

Alcune annotazioni multiple sono collegate da relazioni.

4.2 Stesso span con etichette diverse

In questo caso, uno stesso span può essere annotato con etichette diverse perchè ha più ruoli nell'evento descritto nella notizia (es. pizzeria Marechiaro -> sia PARTE LESA, che LUOGO).

5. Relazioni

Le relazioni servono a collegare le entità con ulteriori informazioni che ne specificano le caratteristiche (vedi Appendice A - Doccano). Solo per alcune entità è possibile specificare una relazione:

- **autore/vittima**: le annotazioni che si riferiscono allo stesso soggetto sono collegate tra loro da relazioni;
- **oggetto rubato**: le annotazioni che si riferiscono alla quantità dell'oggetto sono collegate all'oggetto da relazione.

Attenzione: le relazioni hanno un verso, è importante che questo venga inserito secondo le indicazioni specificate per ciascuna categoria.

6. Oggetto rubato (OBJ)

	Ammesse/non ammesse	Uso corretto
annotazioni multiple	ammesse	fanno riferimento a oggetti diversi
relazioni	ammesse	solo tra oggetto e quantità

L'oggetto rubato è l'entità che è stata sottratta durante il furto. In presenza di più termini, i termini specifici sono da preferire (es. meglio *diamanti* e *oro* rispetto a *gioielli*), questo anche se i termini specifici non descrivono completamente tutta la refurtiva (problema della granularità delle informazioni).

Nel caso in cui siano presenti due termini che indicano la stessa entità, scegliere il termine più specifico, se ciò non è possibile, annotare la prima occorrenza.

I ladri hanno svaligiato il negozio portando via diamanti, collane, orologi e altri preziosi.
 •OBJ •OBJ •OBJ

Figura 6.1: Annotazione degli oggetti rubati: “preziosi” non viene annotato perchè gli altri span sono più specifici.

Hanno portato via Otto, un lagotto di un anno e mezzo, un cane da tartufo dall'ottimo fiuto.
•OBJ

Figura 6.2: Tra gli span “lagotto” e “cane da tartufo” non è possibile decidere quale dei due è più specifico, quindi viene annotato “lagotto” perché appare prima nel testo.

Non devono essere incluse nello span le caratteristiche dell'oggetto, quali il valore, il materiale, il colore (ad esempio, in anelli d'oro viene annotato soltanto anelli, in 250 euro in contanti viene annotato 250 euro) a meno che non denotino una specializzazione (quali la marca, il modello, ecc.). In quest'ultimo caso, lo span identifica una entità specifica in riferimento alla categoria generale (es. computer Apple, tv a plasma, medicinali per animali) (Figura 6.3).

Anche il computer era sparito, come la stampante e una piccola tv al plasma.
•OBJ •OBJ •OBJ

Figura 6.3: Annotazione di una entità specifica: tv al plasma

Se il termine specializzato si trova in **prossimità** del termine generale, questi devono essere annotati insieme, ad esempio “vettura Volvo VX 60”.

Se invece gli span sono discontinui e non c'è prossimità si annota il termine più specifico (vedi Figura 6.4).

Ieri è stata rubata una vettura, poi ritrovata nel fosso, era una Volvo VX 60.
•OBJ

Figura 6.4: Lo span “vettura” non viene annotato.

Se sono indicati nel testo più oggetti rubati, è necessario inserire una annotazione per ciascun oggetto.

Devono essere annotate con **annotazioni multiple** le entità uniche che rappresentano più oggetti rubati (Figura 6.5).

È stata rubata una borsa contenente un cellulare e un portafoglio.
•OBJ •OBJ •OBJ

Figura 6.5: Annotazione multipla di tre oggetti: *borsa, cellulare, portafoglio*

Devono essere annotate con **annotazioni multiple** e **collegate da relazioni** le caratteristiche dell'oggetto quando siamo in presenza di **aggettivi numerali** che denotano la quantità specifica dell'oggetto rubato, ad esempio *diciassette computer* (Figura 6.6).

Attenzione: Le relazioni hanno un verso. E' necessario che questo sia rispettato creando dei collegamenti a partire dall'entità (es. *computer Apple*) verso la quantità (es. *diciassette*).

Non devono essere annotati riferimenti indefiniti come “alcune”, “decine”.

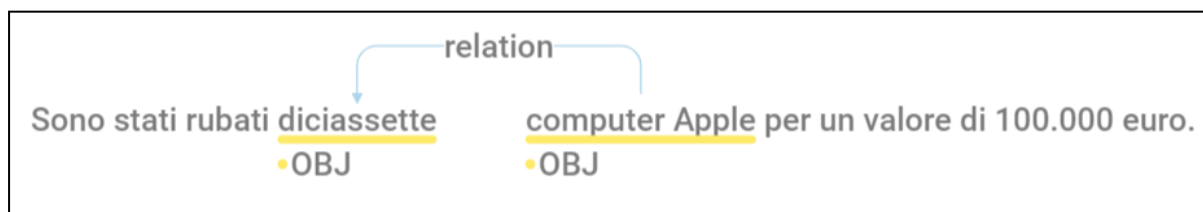


Figura 6.6: La relazione collega il termine che denota l’oggetto rubato (*computer Apple*) con la quantità (*diciassette*).

7. Autore e Gruppo di autori

L’entità che compie il furto può essere descritta come uno o più esseri viventi e/o un gruppo. L’annotazione rispecchia questa differenza individuando:

- **AUT** (essere vivente): es. *donna*, *38enne*, *modenese*, “*nome e cognome*” etc.,
- **AUTG** (gruppo di essere viventi): es. *tre uomini romagnoli*, *5 nordafricani*

7.1. Autore (AUT)

	Ammesse/non ammesse	Uso corretto
annotazioni multiple	ammesse	fanno riferimento ad autori diversi o a caratteristiche di un autore
relazioni	ammesse	tra l’identificatore che denota l’autore e una informazione più specifica

AUT indica la singola entità che commette il furto, un furto può essere commesso da più autori descritti singolarmente.

L’annotazione di span separati viene utilizzata per far riferimento ad autori distinti (Figura 7.1.1).

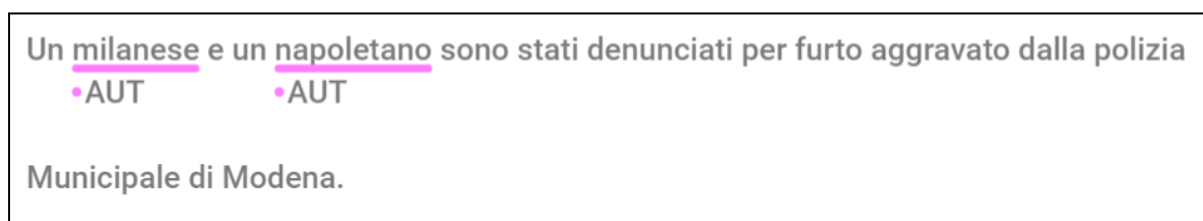


Figura 7.1.1: Esempio di due autori descritti come entità distinte

Nel testo possono essere presenti più informazioni relative all’entità che ha commesso il furto. Per annotare correttamente l’autore e le informazioni relative si deve adottare una procedura in tre step.

Primo step:

Per un singolo autore, prima vengono annotati gli span che lo descrivono, ovvero il **nome proprio** o le **iniziali** e/o le **informazioni socio-demografiche** di riferimento (Figura 7.1.2), quali età, razza, etnia, residenza, abitante/nativo, sesso, occupazione, status giuridico (es. incensurato, pregiudicato, gravato da precedenti). Non devono essere annotate altre caratteristiche o condizioni o ruoli (es. autore, delinquente, biondo, giovane, anziano, in stato di fermo, in libertà vigilata, evaso, marito, moglie ...).

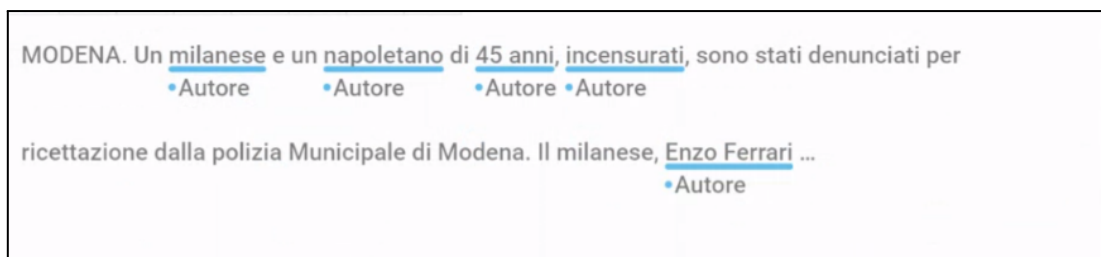


Figura 7.1.2: Primo step - in Doccano, si annotano le informazioni socio-demografiche degli autori

Altri esempi (a destra della freccia sono sottolineati gli span da annotare):

- un giovane di Carpi minorenni → un giovane di Carpi minorenni ("Carpi" può indicare che il minorenne abita a Carpi oppure è nativo di Carpi, quindi viene annotato)
- un giovane residente a Carpi minorenni → un giovane residente a Carpi minorenni
- un minorenne che vive a Carpi → un minorenne che vive a Carpi
- un minorenne originario di Napoli che vive a Carpi → un minorenne originario di Napoli che vive a Carpi ("lavora a Carpi", "studia a Carpi" non vanno annotati)

Eccezione:

- un giovane di Carpi → un giovane di Carpi (in questo caso "giovane" viene annotato perché è l'unico nome che indica il soggetto e permette di creare il collegamento con l'informazione socio-demografica "Carpi")

Attenzione: non vanno annotate informazioni non certe, (ad es. "le erano parsi magrebini" → non annotare; "cittadini dell'Est, probabilmente rumeni" → non annotare perché cittadini dell'Est non è un'informazione socio-demografica, e rumeni non è un'informazione certa).

Secondo step:

Tra i diversi span che sono annotati per singolo autore, occorre scegliere uno span che denoti il soggetto chiamato **identificatore** dell'autore all'interno del testo.

L'identificatore sarà il nome proprio o le iniziali, se presenti, (es., *Enzo Ferrari*), oppure la prima occorrenza che denota l'autore all'interno del testo.

Se uno span viene scelto come identificatore di un autore, non può essere usato come identificatore per altri autori.

Se ci sono più autori ciascuno dovrà avere un identificatore diverso.

Esempio:

“Il furto è stato commesso da un uomo, napoletano 45enne residente a Carpi che ha agito assieme a un complice, un altro uomo milanese residente a Modena”

Autore1 è descritto da 4 span: *uomo*, *napoletano*, *45enne*, *residente a Carpi*

→ **identificatore**: uomo

Autore2 è descritto da 3 span: *uomo*, *milanese*, *residente a Modena*

→ **identificatore**: milanese

Terzo step:

L'identificatore deve essere collegato attraverso una **relazione** con ciascuna altra informazione che lo descrive (Figura 7.1.3).

Attenzione: Le relazioni hanno un verso. E' necessario che questo sia rispettato creando dei collegamenti, ciascuno a partire dall'identificatore (es. *uomo*) verso le altre caratteristiche del soggetto (es. *39enne*, *residente a Carpi*).

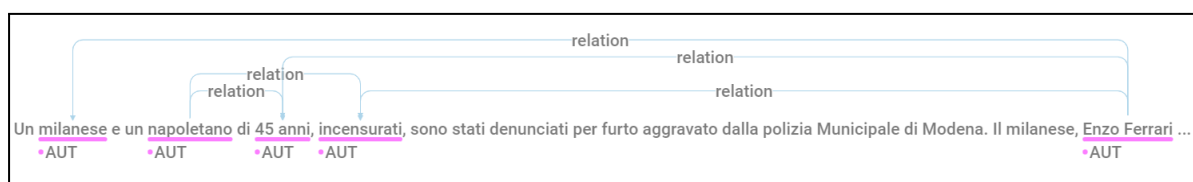


Figura 7.1.3: Terzo step - in Doccano, si inseriscono le relazioni che collegano ciascun autore con le informazioni socio-demografiche

Se sono presenti indicazioni che riguardano autori non identificabili, per i quali non vengono riportati nè nome nè informazioni socio-demografiche, tali autori non vanno annotati.

Esempio:

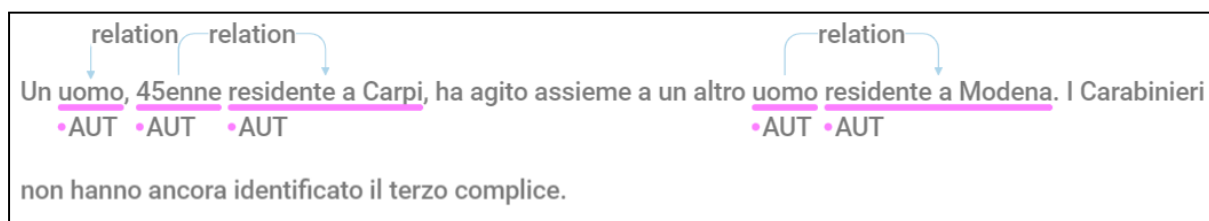


Figura 7.1.4: Non sono presenti informazioni utili sul terzo complice, quindi non viene annotato

CASTELFRANCO. Ieri notte, i carabinieri della Compagnia di Modena, hanno arrestato
in flagranza di tentato furto in appartamento, un cittadino <u>rumeno</u> , <u>classe 1999</u> . Il
•Autore •Autore

Figura 7.1.5: L'autore viene annotato indicando la nazionalità e l'età.

7.2. Gruppo Autori (AUTG)

	Ammesse/non ammesse	Uso corretto
annotazioni multiple	ammesse	fanno riferimento al gruppo di autori
relazioni	non ammesse	

Se il furto è compiuto da un gruppo di autori nel testo potrebbero esserci riferimenti che descrivono tale gruppo (es. una banda di 5 ladri marocchini → una banda di 5 ladri marocchini).

Come per gli autori singoli, anche per un gruppo, devono essere annotati gli span che lo descrivono ovvero le **informazioni socio-demografiche** di riferimento quali età, razza, etnia, residenza, abitante/nativo, sesso, occupazione, status giuridico (es. *incensurato*, *pregiudicato*) che saranno riferite al gruppo che ha compiuto il furto e non a singoli autori o sottogruppi.

Attenzione: non si usano le relazioni

Non devono essere annotati con etichetta AUTG eventuali descrizioni di sottogruppi (es. “altri *due complici* non sono stati identificati”).

Non vanno annotati i termini generali, quali *malviventi*, *banda*, *ladri*, ma soltanto il numero dei componenti e/o caratteristiche socio-demografiche.

Il furto è stato compiuto da una banda di <u>5</u> ladri <u>marocchini</u> , di cui due sono stati fermati.
•AUTG•AUTG

Figura 7.2.1: Annotazione del gruppo di autori, notare che “ladri” non viene annotato.

Una banda di ladri ha compiuto il saccheggio ieri.
--

Figura 7.2.2: Non viene annotato nessuno span perché “banda” è un termine generico.

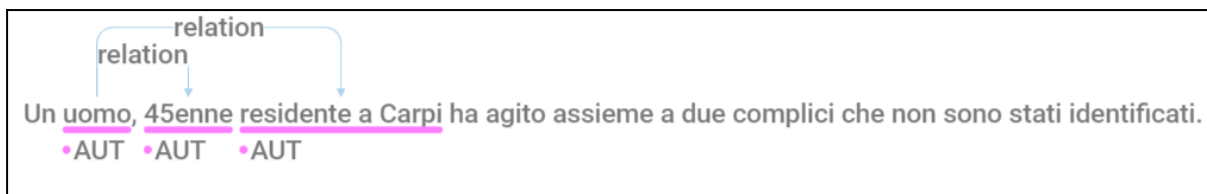


Figura 7.2.3: Lo span “due complici” non viene annotato perché generico.

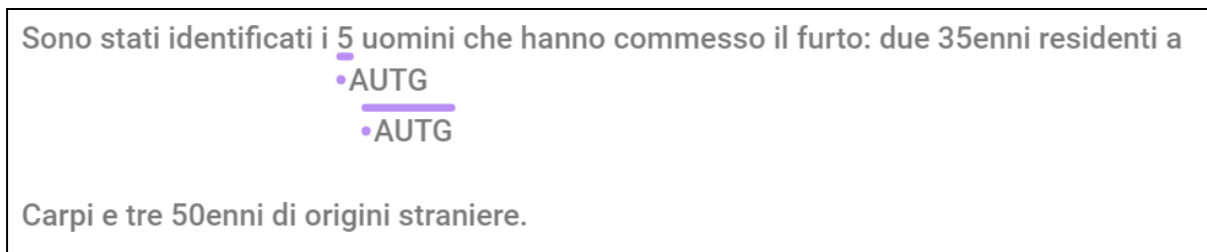


Figura 7.2.4: Sono annotati due span differenti per indicare il gruppo di ladri, non è inserita nessuna relazione.

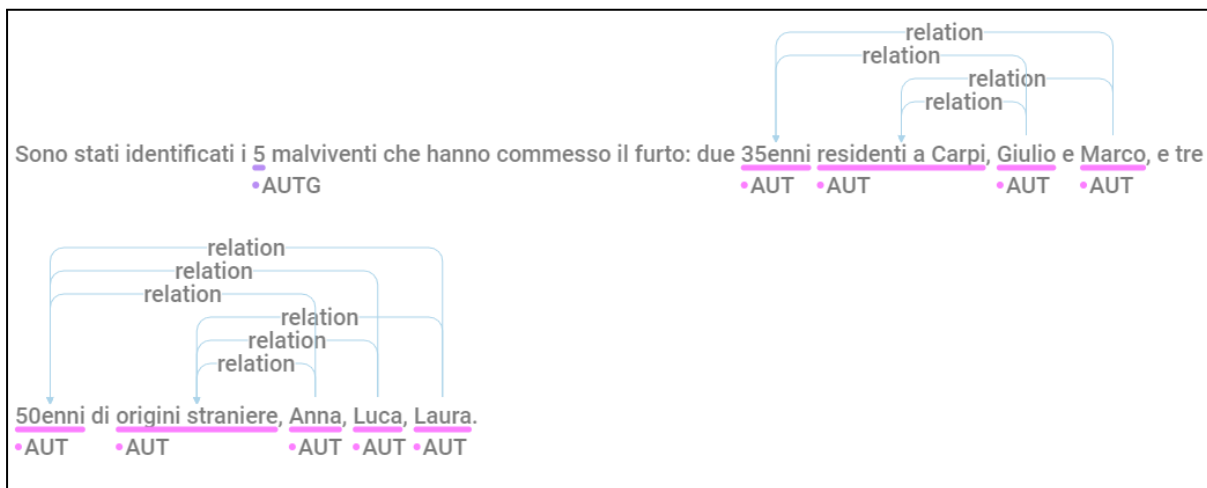


Figura 7.2.5: In questa notizia si annota sia il gruppo di autori sia gli autori singolarmente poiché è possibile identificare ogni autore e vengono fornite diverse informazioni per ognuno di loro. Il gruppo di autori è identificato dallo span “5”, mentre “malviventi” non viene annotato perché generico.

8. Vittima, Gruppo di vittime e Parte Lesa

L'entità che subisce il furto può essere un essere vivente o una organizzazione.

L'annotazione rispecchia questa differenza individuando:

- VITTIMA (essere vivente): es. *38enne, modenese, “nome e cognome”, etc.*
- PARTE LESA (organizzazione): es. *coop I Gelsi, ospedale, reparto di Pediatria, negozio di gioielli, “ragione sociale dell'attività”, etc.*

8.1. Vittima (VIC)

	Ammesse/non ammesse	Uso corretto
annotazioni multiple	ammesse	fanno riferimento a vittime diverse
relazioni	ammesse	tra l'identificatore della vittima e una informazione più specifica

L'annotazione della vittima segue le regole dell'annotazione dell'autore.

Se nella notizia sono indicate una o più vittime del furto, come già detto per l'autore le informazioni che si riferiscono allo stesso soggetto sono collegate tra loro da relazioni.

Si segue il processo di annotazione in 3 step presentato in Sezione 7.1.

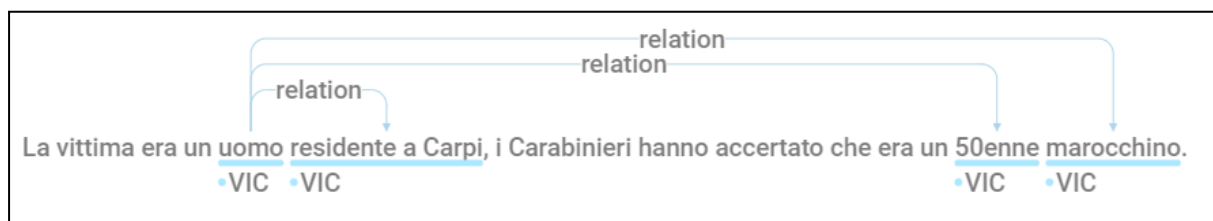


Figura 8.1.1: Annotazione di una vittima con più riferimenti.

Se un furto è stato compiuto ai danni di un negozio, non è corretto annotare il proprietario come vittima, occorre invece annotare il negozio come parte lesa (vedi Sezione 8.3).

Se in un furto ci sono più vittime, occorre interrogarsi se si tratti o meno di un multi evento. Se l'evento avviene in un singolo luogo per mano dello stesso autore/i in un singolo momento temporale è un unico evento di furto, se invece il luogo varia, o gli autori del furto variano, oppure i momenti temporali sono successivi ci troviamo nel caso di un multievento (si vedano esempi in Sezione 3.1).

Attenzione: le relazioni tra etichette diverse non sono ammesse, se in una notizia gli oggetti rubati sono riferibili alle vittime singolarmente vittime e oggetti si annotano separatamente e non si inseriscono relazioni tra oggetto rubato e vittima.

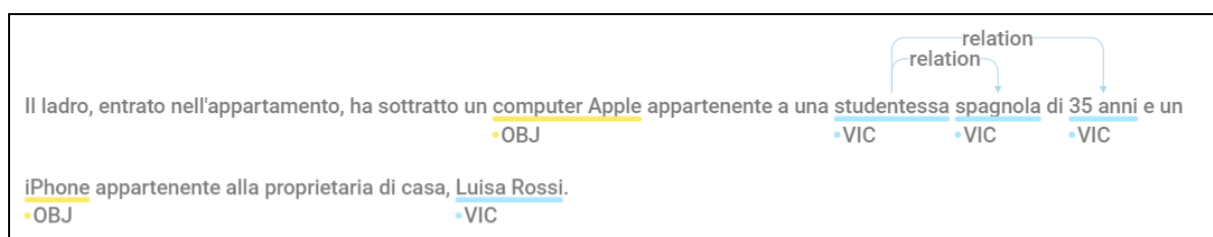


Figura 8.1.2: Annotazione di evento singolo con due vittime e più oggetti rubati (relazioni dirette tra oggetto e vittima non sono ammesse).

8.2. Gruppo Vittime (VICG)

	Ammesse/non ammesse	Uso corretto
annotazioni multiple	ammesse	fanno riferimento al gruppo di vittime
relazioni	non ammesse	

Se il furto è subito da un gruppo di vittime, nel testo potrebbero esserci riferimenti che descrivono tale gruppo (es. 3 signore ottantenni → 3 signore ottantenni).

Come per le vittime singole, anche per un gruppo, devono essere annotati gli span che lo descrivono ovvero le **informazioni socio-demografiche** di riferimento quali età, razza, etnia, residenza, abitante/nativo, sesso, occupazione, status giuridico che saranno riferite al gruppo che ha compiuto il furto e **non a singoli autori o sottogruppi**.

Attenzione: non si usano le relazioni

Non devono essere annotati con etichetta VICG eventuali descrizioni di sottogruppi (es. sono stati derubati la proprietaria, Luisa Rossi, e altri due inquilini → sono stati derubati la proprietaria, Luisa Rossi, e altri due inquilini).

Il ladro è entrato nell'appartamento abitato da 3 studentesse e ha sottratto un computer e una tv.
•LOC •VICG •OBJ •OBJ
•VICG

Figura 8.2.1: Viene annotato il gruppo di vittime in due span separati.

Ieri al centro commerciale, un uomo ha rubato un cellulare, un portafoglio e una borsa. Le vittime sono due signore ottantenni.
•LOC •AUT •OBJ •OBJ •OBJ •VICG •VICG

Figura 8.2.2: Viene annotato il gruppo di vittime in due span separati.

Non vanno annotati i termini generali come *vittime*, *persone*, *malcapitati*, *proprietari*, ma si annotano il numero dei componenti e le informazioni socio-demografiche.

8.3. Parte lesa (PAR)

	Ammesse/non ammesse	Uso corretto
annotazioni multiple	non ammesse	
relazioni	non ammesse	

La parte lesa si riferisce a organizzazioni o aziende che sono vittime di furto.

Per vittime e parte lesa valgono le regole appena riportate per gli altri soggetti.

Se in una notizia viene indicata la parte lesa del furto, non va annotata la vittima.

Sono entrati dalla porta sul retro e hanno svuotato il frigorifero pieno di salumi locali. Mario Rossi, il titolare della
•OBJ
pizzeria Marechiaro conta i danni del furto ...
•PAR

Figura 8.3.1: Annotazione della parte lesa, notare che “Mario Rossi” non viene annotato.

Se un evento presenta la parte lesa, invece della vittima, è probabile che questa coincida con il luogo in cui avviene il furto. In tal caso il relativo span avrà una doppia annotazione, come luogo e come parte lesa.

Hanno fatto sparire anche i computer per la didattica rubati dalle scuole elementari Giovanni XXIII.
•Luogo
•Parte_leso

Figura 8.3.2: Annotazione della parte lesa che coincide con il luogo del furto.

Se il termine specializzato si trova in **prossimità** del termine generale, questi devono essere annotati insieme, ad esempio “pizzeria Re Pipin”.

Se invece gli span sono discontinui e non c'è prossimità si annota il termine più specifico, ad esempio “è stata derubata una pizzeria. Il proprietario di Re Pipin dichiara che...”

9. Luogo (LOC)

	Ammesse/non ammesse	Uso corretto
annotazioni multiple	ammesse	tutti gli span si riferiscono al luogo del furto e forniscono informazioni su luogo generico e area geografica
relazioni	non ammesse	

Il luogo dove avviene un furto è annotato con più span che denotano il luogo generico (es. appartamento, abitazione, azienda di logistica, supermercato, nome del negozio - es. “coop Grandemilia”, “centro commerciale Borgogioioso”) e/o l'area geografica (via, città, etc.).

Il luogo generico permette di identificare la tipologia di luogo dove è avvenuto il furto, mentre l'area geografica permette la geolocalizzazione.

Se l'area geografica è indicata con diversi livelli di specificità, es. “via della Chimica, 11”, “centro storico”, “Carpi”, scegliere le informazioni più specifiche, che permettono di geo-localizzare il luogo del furto, es. “via della Chimica, 11” e “Carpi”.

Se un luogo specifico non c'è, ma ci sono informazioni generali, allora annotare le zone generali “centro storico”, “zona artigianale”, “quartiere Madonnina”.

Attenzione: In alcuni casi il testo della notizia comincia con l'indicazione di una area geografica/nome di città (“data” in gergo giornalistico), generalmente in maiuscolo, che è la città del corrispondente. Tale luogo non va confuso con il luogo in cui avviene il furto, che invece viene riportato nel testo della notizia. Non va mai annotata la città del corrispondente.

MODENA. Oltre 5mila euro di bottino tra salumi, carne e formaggi alla macelleria da Mario di Baggiovara. Un colpo
 •OBJ •OBJ •OBJ •PAR •LOC
 •LOC
 perpetrato da una banda che ben sapeva come muoversi e soprattutto cosa rubare per massimizzare il furto in un
 esercizio commerciale che si preparava all'ultimo giorno dell'anno, uno di quelli in cui l'affluenza di clienti è maggiore.

Figura 9.1 - Notizia che riportata “Modena” come prima parola, che indica la città del corrispondente, diversa dalla città dove avviene il furto che è riportata nel testo della notizia “Baggiovara”.

l'aspirazione di quasi 5mila litri di gasolio. È avvenuto tutto dopo la mezzanotte tra via
 •LOC
 Della Chimica e via Del Lavoro, strade parallele della zona artigianale di Fiorano. Al civico
 •LOC •LOC
 16 c'è il quartier generale della Gelin Trasporti, importante azienda di logistica
 •PAR •LOC
 •LOC
 internazionale che può vantare circa 300 camion di cui almeno una sessantina sono

Figura 9.2 - Sono annotati 5 riferimenti al luogo del furto. Lo span “civico 16” non è annotato perchè non c'è un riferimento alla via, quindi non è un'annotazione utile alla geo-localizzazione.

Non c'è pace per una maranellese che abita a Torre Maina in una casa rurale appena fuori
 •LOC •LOC
 dalla frazione. Il furto è stato consumato l'altro giorno dentro la casa in via Montina.
 •LOC

Figura 9.3: Sono annotati 3 riferimenti al luogo: 1 luogo generico (“casa rurale”) e 2 riferimenti all’area geografica (“Torre Maina” e “via Montina”).

Alcuni luoghi sono costituiti da span di più parole.

Se il termine specializzato si trova in **prossimità** del termine generale, questi devono essere annotati insieme, ad esempio “*centro commerciale Borgogioioso*”.

Se invece gli span sono **discontinui** e non c’è prossimità si annotano tutti gli span separatamente.

Il furto è avvenuto al centro commerciale. I malviventi sono entrati dalla porta laterale del
•LOC
Borgogioioso.
•LOC

Figura 9.4: Poiché nel testo lo span “centro commerciale” appare distaccato da “Borgogioioso”, i due span vengono annotati separatamente.

Attenzione: non confondere il luogo in cui avviene il furto con il luogo dove viene ritrovata la refurtiva

Ieri una Volante della Polizia è riuscita a rintracciare in via Rainusso, a Modena, una Fiat
•OBJ
Uno rubata una decina di giorni fa. L’autovettura, alla quale i ladri hanno divelto il

Figura 9.5: “Via Rainusso” e “Modena” sono riferimenti al luogo in cui è stata ritrovata l’auto rubata, quindi non vengono annotati.

10. Tempo

Si è scelto di non annotare i riferimenti temporali.

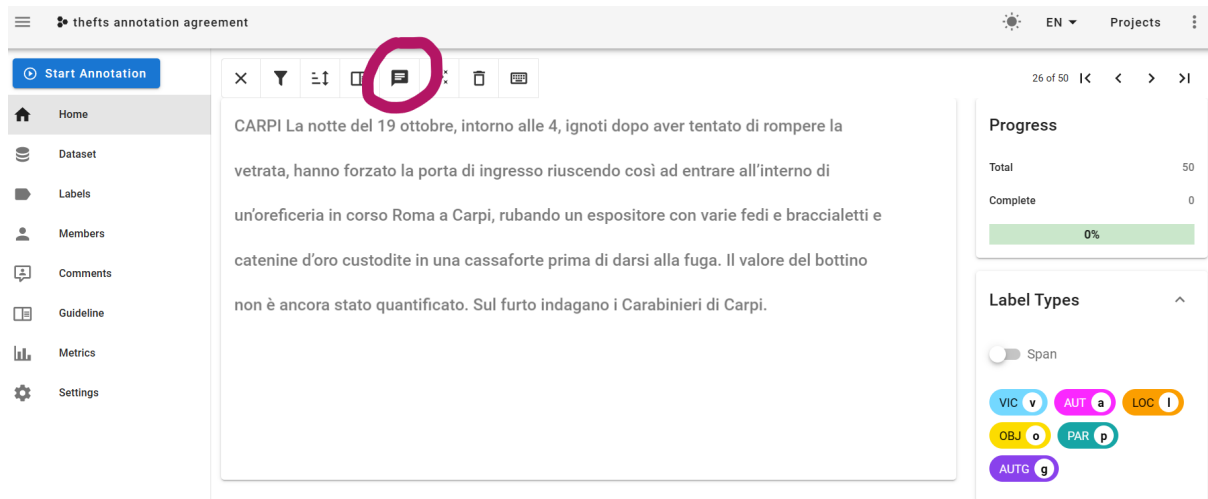
11. Come

Si è scelto di non annotare i riferimenti a come è avvenuto il furto.

Appendice A - Doccano

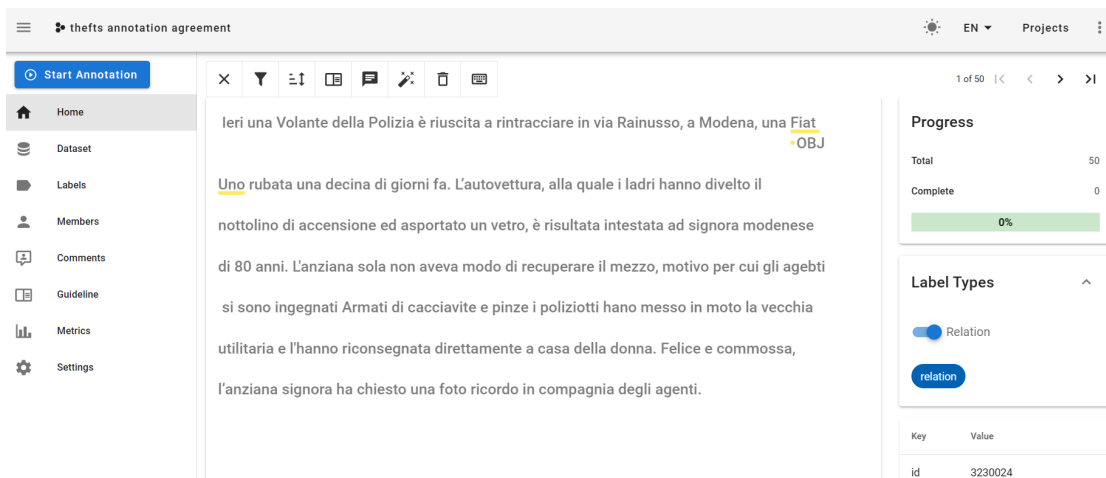
Come inserire i commenti:

nella schermata di annotazione del documento, cliccare sull'icona indicata nella figura sottostante, digitare il commento e cliccare su “Send” e dopo su “Close”.



Come aggiungere le relazioni:

- 1) nella schermata di annotazione del documento, controllare nella barra a destra che il flag “Relation” sia attivo



- 2) annotare tutti gli span con le opportune etichette

thefts annotation agreement

Start Annotation

Home

Dataset

Labels

Members

Comments

Guideline

Metrics

Settings

1 of 50

leri una Volante della Polizia è riuscita a rintracciare in via Rainusso, a Modena, una Fiat OBJ

Uno rubata una decina di giorni fa. L'autovettura, alla quale i ladri hanno divelto il

nottolino di accensione ed asportato un vetro, è risultata intestata ad signora modenese VIC VIC

di 80 anni. L'anziana sola non aveva modo di recuperare il mezzo, motivo per cui gli agebti VIC

si sono ingegnati Armati di cacciavite e pinze i poliziotti hano messo in moto la vecchia

utilitaria e l'hanno riconsegnata direttamente a casa della donna. Felice e commossa,

l'anziana signora ha chiesto una foto ricordo in compagnia degli agenti.

Progress

Total 50

Complete 0

0%

Label Types

Relation

relation

Key	Value
id	3230024

- 3) cliccare prima sull'etichetta dello span di partenza della relazione, poi sull'etichetta dello span di arrivo, apparirà un menù a tendina con i nomi di tutte le possibili relazioni, quindi cliccare sul nome della relazione che si vuole aggiungere.

thefts annotation agreement

Start Annotation

Home

Dataset

Labels

Members

Comments

Guideline

Metrics

Settings

1 of 50

leri una Volante della Polizia è riuscita a rintracciare in via Rainusso, a Modena, una Fiat OBJ

Uno rubata una decina di giorni fa. L'autovettura, alla quale i ladri hanno divelto il

nottolino di accensione ed asportato un vetro, è risultata intestata ad signora modenese relation VIC VIC

di 80 anni. L'anziana sola non aveva modo di recuperare il mezzo, motivo per cui gli agebti VIC

si sono ingegnati Armati di cacciavite e pinze i poliziotti hano messo in moto la vecchia

utilitaria e l'hanno riconsegnata direttamente a casa della donna. Felice e commossa,

l'anziana signora ha chiesto una foto ricordo in compagnia degli agenti.

Progress

Total 50

Complete 0

0%

Label Types

Relation

relation

Key	Value
id	3230024