

Bibliography

- [1] ALTUN, Y.; TSOCHANTARIDIS, I. ; HOFMANN, T. **Hidden Markov support vector machines**. In: PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON MACHINE LEARNING, 2003.
- [2] BICK, E. **Functional aspects in portuguese ner**. In: PROCEEDINGS OF THE 7TH INTERNATIONAL CONFERENCE ON COMPUTATIONAL PROCESSING OF THE PORTUGUESE LANGUAGE, p. 80–89, 2006.
- [3] BRILL, E. **Transformation-based error-driven learning and natural language processing: a case study in part-of-speech tagging**. In: COMPUTATIONAL LINGUISTICS, p. 543–565, 1995.
- [4] CIARAMITA, M.; ALTUN, Y. **Broad-coverage sense disambiguation and information extraction with a supersense sequence tagger**. In: PROCEEDINGS OF THE 2006 CONFERENCE ON EMPIRICAL METHODS IN NATURAL LANGUAGE PROCESSING, EMNLP'06, p. 594–602, 2006.
- [5] CLERGERIE, E. V. L.; SAGOT, B.; STERN, R.; DENIS, P. ; RECOURCÉ, G. **Extracting and visualizing quotations from news wires**. In: PROCEEDINGS OF LTC 2009, p. 522–532, 2009.
- [6] COLLINS, M. **Discriminative training methods for hidden Markov models: theory and experiments with perceptron algorithms**. In: PROCEEDINGS OF THE ACL-02 CONFERENCE ON EMPIRICAL METHODS IN NATURAL LANGUAGE PROCESSING, p. 1–8, 2002.
- [7] CRAMMER, K.; SINGER, Y. **Ultraconservative online algorithms for multiclass problems**. volume 3, p. 951–991, 2003.
- [8] CUEVAS, R. R. M.; PARABONI, I. **A machine learning approach to portuguese pronoun resolution**. In: PROCEEDINGS OF THE 11TH IBERO-AMERICAN CONFERENCE ON ARTIFICIAL INTELLIGENCE, p. 262–271, 2008.

- [9] DOS SANTOS, C. N.; MILIDIÚ, R. L. **Entropy guided transformation learning**. In: FOUNDATIONS OF COMPUTATIONAL INTELLIGENCE (1), p. 159–184. 2009.
- [10] DOS SANTOS, C. N.; MILIDIÚ, R. L.; CRESTANA, C. E. M. ; FERNANDES, E. R. **Etl ensembles for chunking, ner and srl**. In: ELEVENTH INTERNATIONAL CONFERENCE ON COMPUTATIONAL LINGUISTICS AND INTELLIGENT TEXT PROCESSING, volume 6008 of *Lecture Notes in Computer Science*, p. 100–112. Springer Berlin, 2010.
- [11] DOS SANTOS, C. N.; MILIDIÚ, R. L. ; RENTERIA, R. P. **Portuguese part-of-speech tagging using entropy guided transformation learning**. In: PROCEEDINGS OF THE 8TH INTERNATIONAL CONFERENCE ON COMPUTATIONAL PROCESSING OF THE PORTUGUESE LANGUAGE, p. 143–152, 2008.
- [12] DOS SANTOS, C. N.; OLIVEIRA, C. **Constrained atomic term: Widening the reach of rule templates in transformation based learning**. In: PORTUGUESE CONFERENCE ON ARTIFICIAL INTELLIGENCE, p. 622–633, 2005.
- [13] FERNANDES, E. R.; BREFELD, U. Learning from partially annotated sequences, journal.
- [14] FERNANDES, E. R.; DOS SANTOS, C. N. ; MILIDIÚ, R. L. **A machine learning approach to portuguese clause identification**. In: COMPUTATIONAL PROCESSING OF THE PORTUGUESE LANGUAGE, volume 6001 of *Lecture Notes in Computer Science*, p. 55–64. Springer Berlin / Heidelberg, 2010.
- [15] FERNANDES, E. R.; MILIDIÚ, R. L. **Entropy-guided feature generation for structured learning of portuguese dependency parsing**. In: PROCEEDINGS OF THE 10TH INTERNATIONAL CONFERENCE ON COMPUTATIONAL PROCESSING OF THE PORTUGUESE LANGUAGE, volume 7243 of *Lecture Notes in Computer Science*, p. 146–156. Springer, 2012.
- [16] FERNANDES, E. R.; PIRES, B. A.; DOS SANTOS, C. N. ; MILIDIÚ, R. L. **Clause identification using entropy guided transformation learning**. In: PROCEEDINGS OF THE 7TH BRAZILIAN SYMPOSIUM IN INFORMATION AND HUMAN LANGUAGE TECHNOLOGY, p. 117–124, 2009.

- [17] FINGER, M. **Técnicas de otimização da precisão empregadas no etiquetador tycho brahe**. In: PROCEEDINGS OF THE 5TH INTERNATIONAL CONFERENCE ON COMPUTATIONAL PROCESSING OF THE PORTUGUESE LANGUAGE, p. 141–154, 2000.
- [18] FLORIAN, R. **Named entity recognition as a house of cards: Classifier stacking**. In: PROCEEDINGS OF THE CONFERENCE ON COMPUTATIONAL NATURAL LANGUAGE LEARNING, p. 175–178, 2002.
- [19] FREITAS, M. C.; DUARTE, J. C.; DOS SANTOS, C. N.; MILIDIÚ, R. L.; RENTERIA, R. P. ; QUENTAL, V. **A machine learning approach to the identification of appositives**. In: PROCEEDINGS OF IBERO-AMERICAN ARTIFICIAL INTELLIGENCE CONFERENCE, p. 309–318, 2006.
- [20] HIGGINS, D. **A transformation-based approach to argument labeling**. In: HLT-NAACL 2004 WORKSHOP: 8TH CONFERENCE ON COMPUTATIONAL NATURAL LANGUAGE LEARNING, p. 114–117, 2004.
- [21] LIANG, J.; DHILLON, N. ; KOPERSKI, K. **A large-scale system for annotating and querying quotations in news feeds**. In: PROCEEDINGS OF THE 3RD INTERNATIONAL SEMANTIC SEARCH WORKSHOP, p. 7:1–7:5, 2010.
- [22] LIANG, P.; BOUCHARD-CÔTÉ, A.; KLEIN, D. ; TASKAR, B. **An end-to-end discriminative approach to machine translation**. In: IN PROCEEDINGS OF THE JOINT INTERNATIONAL CONFERENCE ON COMPUTATIONAL LINGUISTICS AND ASSOCIATION OF COMPUTATIONAL LINGUISTICS, p. 761–768, 2006.
- [23] MANGU, L.; BRILL, E. **Automatic rule acquisition for spelling correction**. In: PROCEEDINGS OF THE 14TH INTERNATIONAL CONFERENCE ON MACHINE LEARNING, p. 187–194, 1997.
- [24] MEGYESI, B. **Shallow parsing with pos taggers and linguistic features**. In: JOURNAL OF MACHINE LEARNING RESEARCH, p. 2:639–668, 2002.
- [25] MILIDIÚ, R. L.; CRESTANA, C. E. M. ; DOS SANTOS, C. N. **A token classification approach to dependency parsing**. In: PROCEEDINGS OF THE 7TH BRAZILIAN SYMPOSIUM IN INFORMATION AND HUMAN LANGUAGE TECHNOLOGY, p. 80–88, 2009.

- [26] MILIDIÚ, R. L.; DOS SANTOS, C. N. ; DUARTE, J. C. **Journal of the Brazilian Computer Society**. Portuguese corpus-based learning using etl, journal, p. 17–27, 2008.
- [27] MILIDIÚ, R. L.; DOS SANTOS, C. N.; DUARTE, J. C. ; RENTERRIA, R. P. **Semi-supervised learning for portuguese noun phrase extraction**. In: PROCEEDINGS OF THE 7TH INTERNATIONAL CONFERENCE ON COMPUTATIONAL PROCESSING OF THE PORTUGUESE LANGUAGE, p. 200–203, 2006.
- [28] MILIDIÚ, R. L.; DUARTE, J. C. ; CAVALCANTE, R. **Machine learning algorithms for portuguese named entity recognition**. In: INTELIGÊNCIA ARTIFICIAL, REVISTA IBEROAMERICANA DE INTELIGÊNCIA ARTIFICIAL, p. 67–75, 2007.
- [29] MILIDIÚ, R. L.; SANTOS, C. N. ; DUARTE, J. C. **Phrase chunking using entropy guided transformation learning**. In: PROCEEDINGS THE 46TH ANNUAL MEETING OF THE ACL, p. 647–655, 2008.
- [30] POULIQUEN, B.; STEINBERGER, R. ; BEST, C. **Automatic detection of quotations in multilingual news**. In: PROCEEDINGS OF THE INTERNATIONAL CONFERENCE RECENT ADVANCES IN NATURAL LANGUAGE PROCESSING, p. 25–32, 2007.
- [31] RAMSHAW, L.; MARCUS, M. **Text chunking using transformation-based learning**. In: PROCEEDINGS OF THE 3RD ACL WORKSHOP ON VERY LARGE CORPORA, p. 82–94, 1995.
- [32] ROSENBLATT, F. **Psychological Review**. The perceptron: a probabilistic model for information storage and organization in the brain, journal, v.65, p. 386–408, 1958.
- [33] SARMENTO, L. **Siemês: a named entity recognizer for portuguese**. In: PROCEEDINGS OF THE 7TH INTERNATIONAL CONFERENCE ON COMPUTATIONAL PROCESSING OF THE PORTUGUESE LANGUAGE, p. 90–99, 2006.
- [34] SARMENTO, L.; NUNES, S. **Automatic extraction of quotes and topics from news feeds**. In: 4TH DOCTORAL SYMPOSIUM ON INFORMATICS ENGINEERING, 2009.
- [35] SARMENTO, L.; PINTO, A. S. ; CABRAL, L. **Repentino: a wide-scope gazetteer for entity recognition in portuguese**. In: PRO-

- CEEDINGS OF THE 7TH INTERNATIONAL CONFERENCE ON COMPUTATIONAL PROCESSING OF THE PORTUGUESE LANGUAGE, p. 31–40, 2006.
- [36] SOLORIO, T.; LÓPEZ, A. L. **Learning named entity recognition in portuguese from spanish**. In: PROCEEDINGS OF COMPUTATIONAL LINGUISTICS AND INTELLIGENT TEXT PROCESSING, p. 762–768, 2005.
- [37] SOUZA, J. G.; GONÇALVES, P. N. ; VIEIRA, R. **Learning coreference resolution for portuguese texts**. In: PROCEEDINGS OF THE 8TH INTERNATIONAL CONFERENCE ON COMPUTATIONAL PROCESSING OF THE PORTUGUESE LANGUAGE, p. 153–162, 2008.
- [38] TARJAN, R. E. **Networks**. Finding optimum branchings, journal, v.7, p. 25–25, 1977.
- [39] TASKAR, B.; GUESTIN, C. ; KOLLER, D. **Max–margin Markov networks**. In: ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS, 2004.
- [40] TASKAR, B.; KLEIN, D.; COLLINS, M.; KOLLER, D. ; MANNING, C. **Max–margin parsing**. In: PROCEEDINGS OF THE CONFERENCE ON EMPIRICAL METHODS IN NATURAL LANGUAGE PROCESSING, 2004.
- [41] ZHOU, G.; SU, J. **Named entity recognition using an hmm-based chunk tagger**. In: PROCEEDINGS THE 40TH ANNUAL MEETING OF THE ACL, p. 647–655, 2002.

7 Glossary

| | |
|--------|-------------------------------------------------------------|
| BLS | Baseline system |
| DT | Decision tree |
| ETL | Entropy guided transformation learning |
| IG | Information gain |
| ML | Machine Learning |
| NE | Named entity |
| NLP | Natural Language Processing |
| POS | Part-of-speech |
| SP-WIS | Structured perceptron based on weighted interval scheduling |
| TBL | Transformation based learning |
| WIS | Weighted interval scheduling |

A Annotation Guidelines

During the annotation process, we have produced guidelines for named entity annotation, coreference annotation and association between quotation and author annotation. In the next sections, we describe the objective of each annotation and present the produced guidelines.

A.0.1 Named Entity Annotation

We annotate two kinds of named entities (NE): person and organization. In Figure A.1, we present an illustrative example of named entity annotation in a text. Named entities are in **bold**.

O ministro do Meio Ambiente, **Carlos Minc**, disse neste domingo (8) que as obras de despoluição do Canal de o Cunha, na Ilha do Fundão, no Rio de Janeiro, serão iniciadas em breve pela **Petrobras**. Segundo reportagem da **Agência Brasil**, **Minc** participou da cerimônia de posse da nova diretoria da **Superintendência Estadual de Rios e Lagoas (Serla)**, e informou já ter conversado sobre a questão com o presidente da estatal, **José Sergio Gabrielli**.

Figure A.1: Illustrative example of NE annotation

We have produced eleven annotation guidelines for named entity in the annotation process. Below we describe them and present application examples.

1. A *polícia militar* teve de abrir caminho entre os torcedores que ocupavam a entrada do centro de treinamento.
2. A **Polícia** informou que a maioria de os torcedores se comportou de forma pacífica, tanto antes quanto durante e depois dos jogos.

Figure A.2: Application example of the first annotation guideline for NE

Uppercase is a determinant factor to named entity annotation. We only annotate capitalized phrases. In Figure A.2, we present an application example of this guideline. Annotated named entities are in **bold** and not annotated candidates are in *italic*. Even though both expressions refer to the organization

Police, we do not annotate the expression in the first sentence, since it is in lowercase.

1. O **Conselho de Orientação Fiscal do Corinthians** se reuniu na noite desta terça-feira, no Parque São Jorge, e sugeriu a suspensão do ex-presidente Alberto Dualib e do ex-vice Nesi Curi.
2. O árbitro foi denunciado nos artigos 266 e 267 do Código Brasileiro de *Justiça Desportiva*, sendo punido com 60 dias de suspensão em cada um.
3. Alex Mineiro marca e avança na briga pelo Prêmio *Friedenreich*.
4. Alex está fora do *Gre-Nal*.
5. Escute “*Jesus em Copacabana*”, o novo som da Banda Ganeshas.

Figure A.3: Application example of the second annotation guideline for NE

The context of the named entity candidate determines if it is annotated. We present an application example of this guideline in Figure A.3. Annotated named entities are in **bold** and not annotated candidates are in *italic*. In the first sentence, the highlighted phrase represents a board and, thus, is annotated. In the second sentence, the organization *Justiça Desportiva* is not annotated, since in the context, it qualifies the code of laws. In the third sentence, the person *Friedenreich* is not annotated, since in the context, it qualifies the prize. In the fourth sentence, the soccer teams *Grêmio*¹ e *Internacional*² are not annotated, since in the context, they refer to an event, the soccer game. In the fifth sentence, *Jesus* is not annotated, since it is part of a song name.

Sites, TV channels and newspaper, when not referring to physical or virtual places, are annotated. We present an application example of this guideline in Figure A.4. Annotated named entities are in **bold** and not annotated candidates are in *italic*. From the first to the third sentence, the site *Fôlego*, the TV channel *Premiere* and the newspaper *Marca* are annotated. From the fourth to the sixth sentence, the site *Globoesporte.com*, the TV channel *SporTV2* and the newspaper *Diário Oficial* represent virtual places and, therefore, are not annotated.

When an organization is mentioned, but not referring to the organization itself, but to a product produced by the organization or to the physical location of the organization, it is not annotated. We present an application example of this guideline in Figure A.5. Annotated named entities are in **bold** and not annotated candidates are in *italic*. In the first sentence, *Ferrari* is not

¹Gre

²Nal

1. O blog **FÔLEGO** publica série de programas de treinos para Meia Maratona do Rio.
2. O canal **Premiere** transmite o jogo ao vivo para todo o Brasil em sistema de pay-per-view.
3. (...) afirma Cerezo em entrevista ao jornal '**Marca**', da Espanha.
4. A partida, válida pelas quartas-de-final, terá início às 23h (de Brasília) deste sábado, com acompanhamento em Tempo Real no *GLOBOESPORTE.COM*.
5. Você assiste a programação no *SporTV2*.
6. No início deste mês, a decisão judicial para as novas eleições foi publicada no *Diário Oficial*.

Figure A.4: Application example of the third annotation guideline for NE

1. O polonês dominou a corrida após a primeira entrada do safety car, quando Adrian Sutil parou na pista e Lewis Hamilton acertou a traseira da *Ferrari* de Kimi Raikkonen na saída dos boxes, forçando o abandono de ambos.
2. Vídeo que aborda o fenômeno do 'bullying' entre adolescentes será lançado nesta terça-feira, no *Museu da Vida*.
3. Ciência e diversão não têm idade: **Museu da Vida** celebra o Dia do Idoso.

Figure A.5: Application example of the fourth annotation guideline for NE

mentioned as an automobilistic organization, but as a formula one car, and, therefore, it is not annotated. In the second sentence, *Museu da Vida* is mentioned as a place and is not annotated, while in the third sentence, it is mentioned as an organization and, thus, it is annotated.

1. A opinião é compartilhada pelo lateral Leo, ex-**Santos**, que joga no Benfica desde 2005.
2. O atacante Rafael Sobis, ex-**Internacional**, é o novo reforço do Al-Jazira, dos Emirados Árabes.

Figure A.6: Application example of the fifth annotation guideline for NE

Even when an organization is preceeded by *ex*, it may be annotated. We present an application example of that guideline in Figure A.6. Annotated named entities are in **bold**. In both cases, the highlighted soccer teams are annotated.

If a person or organization has a capitalized designation, it is included in

the annotation. We present an application example of that guideline in Figure A.7. Annotated named entity is in **bold**.

1. É o clube que revelou o **Rei Pelé**, tem muita repercussão no exterior.

Figure A.7: Application example of the sixth annotation guideline for NE

1. O ‘**Dirigente do Século**’ – ao lado do **Barão de Coubertin** – não é metido a atleta.
2. Após tantas andanças, o ‘**Senhor das Copas**’ não esquece de sua primeira experiência como dirigente da seleção brasileira...
3. **Luxa** acredita que com mais três vitórias seguidas o **Verdão** assume a liderança.
4. **Estado** regularizará posse de famílias carentes a imóveis públicos.
5. Tivemos vários jogadores sondados, alguns com propostas firmes, boas para o atleta e o **Clube**.

Figure A.8: Application example of the seventh annotation guideline for NE

Nickname, title or generic name of a person or organization is annotated. We present an application example of that guideline in Figure A.8. Annotated named entities are in **bold**. In the first and second sentence, the highlighted expressions are titles given to people and are annotated. In the third sentence, both highlighted words are nicknames and are annotated. In the fourth sentence, the highlighted word is a generic name that refers to the political institution of a country and is annotated. In the fifth sentence, the highlighted word is a generic name that refers to a soccer club and is annotated.

1. Livro conta a história das relações médico-científicas entre **Brasil** e **Japão**.
2. A equipe do **Reino Unido** foi prata, com 5m25s11.
3. Mike Carvalho busca segunda vitória do **Brasil** no boxe.
4. **EUA** vencem no feminino; **Canadá** leva ouro no masculino.
5. Alla Vazhenina, do *Cazaquistão*, ficou com a prata ao levantar 266 kg (119kg e 147kg, respectivamente).

Figure A.9: Application example of the eighth annotation guideline for NE

Country, when mentioned as a geopolitical organization or team, is annotated. We present an application example of that guideline in Figure A.9.

Annotated named entities are in **bold** and not annotated candidates are in *italic*. In the first sentence, *Brasil*³ e *Japão*⁴ are mentioned as geopolitical organizations and, therefore, are annotated. From the second to the fourth sentence, the country names represent teams and are annotated. In the fifth sentence, *Cazaquistão*⁵ is mentioned as a place, and, thus, is not annotated.

1. Weldon agradece o apoio e já pensa no futuro: –O time jogou bem, graças a **Deus**.
2. –Há dois anos participo desta maratona e graças a **Deus** consegui uma boa colocação neste ano.

Figure A.10: Application example of the ninth annotation guideline for NE

Superior entity is annotated. We present an application example of that guideline in Figure A.10. Annotated named entities are in **bold**.

1. Mais de 10 mil bolas coloridas também serão distribuídas na entrada das arquibancadas e o mascote do Fluminense, o **Cartola**, será o elemento surpresa da festa.

Figure A.11: Application example of the tenth annotation guideline for NE

Team mascot is annotated as person. We present an application example of that guideline in Figure A.11. Annotated named entities are in **bold**.

1. A festa no Mineirão, que começou com **Jota Quest** e passou por **Skank**, terminou com vaias.
2. Para exemplificar o que é possível fazer, a banda **Capital Inicial** consegue colocar um show inteiro à disposição do público em seu canal, algo que um usuário comum não conseguiria fazer.

Figure A.12: Application example of the eleventh annotation guideline for NE

Band is annotated. We present an application example of that guideline in Figure A.12. Annotated named entities are in **bold**.

³Brazil

⁴Japan

⁵Kazakhstan

A.0.2 Coreference Annotation

The coreference annotation is based on the coreference set, i.e., each coreferent element is tagged with a label that identifies the group it belongs to. Each coreference set must have at least one named entity of type person or organization. If a person or organization is mentioned only once in a document, they are not annotated. In Figure A.13, we present an illustrative example of identified coreferences in a text. Annotated coreferences are in **bold** and they are tagged with their respective coreference set label. Not annotated candidates are in *italic*. They are not annotated, since they are mentioned only once in the document.

Alexandre Gallo₁ deve ser apresentado como o novo técnico do **Atlético**₂ na manhã desta terça-feira. Depois de se desligar do *Figueirense*, no último domingo, **o treinador**₁ vai se reunir com a diretoria alvinegra para acertar os últimos detalhes da negociação iniciada após a saída de *Geninho*, na última quarta-feira. Como jogador, **Gallo**₁ viveu um momento marcante no **Atlético**₂.

Figure A.13: Illustrative example of coreference annotation

We have produced two annotation guidelines for coreference during the annotation process. Below we describe them and present application examples.

1. **Lucio Flavio**₁ foi um de os destaques de o **Botafogo**₂ em o Campeonato Carioca, além de ser **o capitão**₁ de **a equipe**₂.

Figure A.14: Application example of the first annotation guideline for coreference

1. O incidente ocorreu por volta das 18h local (19h de Brasília) e ‘a tripulação não informou qualquer ferido a bordo’, disse à AFP *a porta-voz do aeroporto*, **Stéphanie Lepage**₁.
2. –Se tudo correr bem ele se apresenta na quarta-feira – diz *o vice-presidente de futebol*, **Kléber Leite**₁.

Figure A.15: Application example of the second annotation guideline for coreference

If the coreference is a named entity, we do not include the article in the annotation. Otherwise, the article is included in the annotation. We present an application example of that guideline in Figure A.14. The sentence is presented after the tokenization. As *Botafogo* is a named entity, we do not include the

article *o* in the coreference annotation. However, *capitão* and *equipe* are not named entities. Thus, we include the articles in the coreference annotation.

When there is a role performed by a person and it is followed by his or her name, the role is not annotated. We present an application example of that guideline in Figure A.15. Annotated coreferences are in **bold** and not annotated roles are in *italic*.

A.0.3

Association Between Quotation and Author Annotation

We annotate quotations between quotation marks ' " - and dialogs. In Figure A.16, we present illustrative examples of different types of annotated quotations, in *italic*.

1. *–As páginas da internet são muito flexíveis – reconhece Wu, para quem uma das maiores dificuldades é a imprecisão nos dados destes endereços.*
2. Para ele, Mateus '*retrojeta*' uma situação da Igreja primitiva para a época em que Cristo ainda estava vivo.
3. "*Para qualquer dano sofrido, o consumidor tem todo o direito a pedir uma indenização*", afirma Fátima Lemos, técnica de defesa do consumidor, do Procon-SP.
4. EGO: *Já participou de outros leilões beneficentes desse tipo?*
LUCÍLIA DINIZ: *Sempre participo de eventos desse tipo. Procuro ajudar instituições e eventos beneficentes.*

Figure A.16: Illustrative examples of different types of annotated quotations

We have produced one annotation guideline for association between quotation and author during the annotation process. Below we describe it and present an application example.

We annotate as quotation author the coreference indicated by the verb of speech. When it is omitted, we annotate the nearest coreference author before the quotation. In Figure A.17, we present an application example of that annotation guideline. Quotations are in *italic*, coreferences are in **bold** and verbs of speech are underlined. We use a relative distance tag subscript of type $rd\pm$. For instance, the tag $r2+$ indicates that the quotation author is the second coreference after the quotation. In the first and second examples, the verbs of speech indicate the quotation author. In the third example, the verb of speech omits the quotation author. Thus, we annotate the nearest coreference author before the quotation. The fourth example is a dialog. Therefore, the quotation author is immediately before the quotation.

1. “*O cliente não fica isento de pagar na data. É preciso verificar se a empresa disponibiliza outra forma de pagamento*”_{r1+}, diz **Fátima**.
2. O incidente ocorreu por volta das 18h local (19h de Brasília) e ‘*a tripulação não informou qualquer ferido a bordo*’_{r2+}, disse à **AFP** a porta-voz do aeroporto, **Stéphanie Lepage**.
3. **A ex-diretora** saiu do Senado sem falar com os jornalistas. ‘*Preciso terminar de tirar cópia dos documentos para entregar aos senadores*’_{r1-}, limitou-se a dizer.
4. **EGO**: *Já participou de outros leilões beneficentes desse tipo?*_{r1-}
LUCÍLIA DINIZ: *Sempre participo de eventos desse tipo. Procuro ajudar instituições e eventos beneficentes.*_{r1-}

Figure A.17: Application example of the annotation guideline for association between quotation and coreference

The annotation of quote start, quote end and quote are automatically derived from the annotation of association between quotation and author.