

Annotation Guidelines for Product NER

***Product* Mentions
and
CompanyProvidesProduct Relation Mentions**

Introduction	3
The CompanyProvidesProduct Relation	3
Company Names as Part of the Product Mention	4
Explicitness of Relation Mentions	5
Several Arguments Within a Sentence	5
Product Mentions as Part of a Different Concept	6
Companies, Products and Triggers	6
Companies (Organizations)	6
Products	8
Definition of Product Entity	8
Elements of a Product Mention	8
Elements Excluded from Product Mentions	10
Triggers	14
The Identity Relation (Coreference)	15

1. Introduction

In this document, we will present the guidelines for annotation of Company, Product and Trigger concepts, as well as the *CompanyProvidesProduct* relation, in the DFKI Product Corpus. We first describe the relation *CompanyProvidesProduct* because it includes products as one of its essential arguments and can be found in a large number of sentences. As we will show, company arguments can be found within a product mention, which is another argument for selecting this particular relation. In the second part of the annotation guidelines we will present the possible extent of a product mention. First, the elements that a product mention can consist of are presented. Subsequently, we will give an overview of the elements that commonly appear close to a product mention and might be mistaken as part of it, also illustrating exceptions by means of examples of what are often problematic cases.

Please note that in the following examples, the font is changed to highlight certain aspects:

- **company mentions**: bold
- triggers: italic and underlined
- *[product mentions]*: italic and in square brackets

2. The CompanyProvidesProduct Relation

The *CompanyProvidesProduct* relation describes a producer or “provider” relation between a single company or organization, and one or more products. It consists of two mandatory arguments, a company (organization) and a product, and an optional argument type, namely a (set of) trigger phrases. A company can serve as the first argument if it is stated as the creator, manufacturer, provider, distributor or vendor of the mentioned product(s). The slot for the product argument can be filled by one or more product mentions (e.g., in the case of conjunctive enumerations). The following example illustrates a *CompanyProductRelation* mention:

- **Sony** is one of the biggest producers of [image sensors].

company	Sony
trigger	producers
product	image sensors

We annotate relation mentions (RMs) only within a sentence, not cross-sentential. A colon ":" is not a sentence separator, therefore we annotate relations in sentences such as:

- **Siemens'** products include: [gas turbines], [wind mills], and [power plants].

- World's 10 biggest *[smartphone]* companies: **Samsung, Apple, Huawei, Oppo, Vivo, OnePlus, Xiaomi, Lenovo, LG, Sony**

2.1. Company Names as Part of the Product Mention

In a number of cases, the product mention includes the company name, creating an RM within the mention. The company name is, however, *not* part of the product extent if the company name is followed by a possessive marker ('s', ') or a different punctuation mark (other than a hyphen) that is neither part of the company's or the product's proper name. Note that the company name, when part of the product mention, may be reduced to a shorter form.

- **[Apple Watch Series 2]**
- **[BMW Z3]**

but:

- **BMW's [Z3]**
- Find nutrition facts for **Maggi's [So Juicy Bbq Ribs]**.
- Change is coming to **Mercedes' [small-car]** range.

An exception to this rule are company names that already include one of the aforementioned possessive markers:

- There are 260 calories in a 1/2 cup serving of **[Ben & Jerry's Coffee Coffee Buzz Buzz Buzz Ice Cream]**.
- The new **[McDonald's Nutella burger]** has been launched in Italy and it's creating a lot of buzz.

In a nested RM we tag the company name as part of the product, and individually as both a company mention and the company argument in the RM. If there is another mention of the company outside the product extent and the relation is explicitly triggered, we annotate it as a second RM:

- **Lumo Bodytech**, the creators of **[Lumo Run]**, today announced the addition of Mark Mastalir to the team as Vice President of Marketing and Partnerships.

Here, we annotate two RMs, connecting both "Lumo Bodytech" and "Lumo" with the product "Lumo Run".

In the following example, we annotate only two RMs although there are two mentions of each potential argument:

- When paired with **Sentient Energy's** family of *[intelligent line sensors]*, the *[Ample Analytics platform]* provides everything needed to manage and analyze the immense amount of data captured by **[Sentient MM3™ sensors]** when deployed on the distribution grid.

There is a relation between “Sentient Energy” and “intelligent line sensors” as well as between “Sentient” and “Sentient MM3 TM sensors”. The reason why there are not four RMs here is that the full mention of the company name (“Sentient Energy”) is not connected explicitly to the product “Sentient MM3 TM sensors” (unless we can resolve the coreference between “Sentient Energy” and “Sentient”, which we do not). “Ample Analytics platform” should be annotated as a product but cannot be linked with a company.

2.2. Explicitness of Relation Mentions

In sentences that include a company mention and a product mention but feature no trigger word or structure that makes the relation explicit, we do not annotate a RM. Context, in which the relation might be expressed or implied, must be ignored. If the annotator has to rely on his or her world knowledge to deduce the relation between the product and the company, no RM is annotated.

In the following example, the second product mention “Ample Analytics platform” can only be identified as a product by company “Sentient Energy” as a result of reading the context. Therefore, we only annotate it as a product, not, however, as part of an RM. The only RM in this example sentence is between “Sentient Energy” and “Sentient Energy intelligent sensors”.

- Together, [***Sentient Energy** intelligent sensors*] and [*Ample Analytics platform*] hold the key to accurately predicting future grid equipment failures and preventing outages before they occur...

The same applies in the next example. Only with world knowledge does the reader know that the “B-Class” is part of “Mercedes’ small-car range” because the sentence is not linguistically explicit. Therefore “B-Class” will be annotated as a product but not as part of the RM.

- Change is coming to **Mercedes’** [*small-car*] range, and this new [*B-Class*] is the first taste of it.

We also consider RMs in sentences that are not grammatically “complete”. Examples are often found in titles.

- **Sentient Energy** Announces *Release* of [*Ample Analytics 3.0*]
- *companies* that sell [*tablets*], [*sweets*] and [*snacks*] (eg **Nestlé**, **Ferrero** and **Mars**)
- World’s 10 biggest [*smartphone*] *companies*: **Samsung**, **Apple**, **Huawei**, **Oppo**, **Vivo**, **OnePlus**, **Xiaomi**, **Lenovo**, **LG**, **Sony**

2.3. Several Arguments Within a Sentence

In the *CompanyProvidesProduct* relation, there can only be one company mention per RM serving as the company argument. This is not usually true for the product argument. In the following examples, a company is mentioned alongside several of their products.

- FPL will use **Sentient Energy's** *[Grid Analytics System]* – *[MM3 sensors]*, *[high-resolution oscillography]*, and the *[Ample Analytics platform]* – to detect minute disturbances.
- List of **Distech Controls'** *products* installed: - *[1500 ECL-PTU-208]* - *[1500 ECx-Light-4]* - *[1500 ECx-Blind-4]* - *[3000 EC-Multi-Sensor]* - *[3000 EC-Smart-Vue communicating sensors]* - *[60 EC-BOS]* - *[600 I/O modules]* ...

In the first example, there is a hierarchy between the products, whereas in the second example, the products are simply listed. Accordingly, we annotate RMs between “Sentient Energy” and “Grid Analytics System”, “MM3 sensors”, “high-resolution oscillography” and “Ample Analytics platform” in the first sentence. In the second example, there is one RM between the company mention and each element tagged as a product.

2.4. Product Mentions as Part of a Different Concept

In some cases, the product mention is part of a enclosing noun phrase describing a separate concept. In the following example, the product is “Anti-Collision Sensors” and is part of the concept “Anti-Collision Sensors Market”. In this case, we annotate a (nested) product, and treat “market” as a trigger phrase:

- Key *players* in the Global *[Anti-Collision Sensors] Market*: **Continental AG, Delphi Automotive, Denso, Bosch and TRW Automotive**.

3. Companies, Products and Triggers

3.1. Companies (Organizations)

We annotate **organizations** referenced by names / proper nouns anywhere in a document. In addition, we annotate (pro-) nominal references to companies and connect them via the *Identity* relation. Organizations must always refer to a specific organization (proper name) at least once in the document. See the ACE Guidelines for examples: <https://www ldc.upenn.edu/collaborations/past-projects/ace/annotation-tasks-and-specifications>)

- **Parkifi** *sells* *[smart sensors]*.
- With **our** *[sensors]*, we can reduce city traffic and parking hassles by directing drivers to a parking space.
- **We** *deliver* *[Innovative Solutions for Greener Buildings™]* through our passion for innovation, quality, customer satisfaction, and sustainability.

- In August 2014, the **company** launched its [*linear Hall-effect magnetic position sensor IC*]...
- The **company** offers a broad range of advanced [*magnetic sensor*] products.
- The top five global [*MEMS magnetic sensors*] vendors are: ...
- As a [*landscape*] **company** in Atlanta, we know which plants thrive in this planting zone and know the optimum landscaping designs for local yards and business."

If there is a pronoun or a common noun that stands for the company in addition to the explicit company name, we tag the latter as the argument, not the pronoun or the common noun. In these cases the pronoun or common noun is part of the pattern and can be treated equivalent to a trigger.

- **Parkifi** has **their** [*smart sensors*] installed across the nation and continues to see consistent growth.
- **Sentient Energy** will showcase [*Ample Analytics 3.0*] along with **its** [*MM3 intelligent sensors*], and [*high-resolution waveform analysis*] at DistribuTECH 2016 – booth #849.

Sometimes, after a first full extent mention, the company name appears in the text in an abbreviated form. Even if there is no full extent mention at all, we annotate the mention as the company argument for the relation CompanyProvidesProduct.

- **Lumo Bodytech**, the creators of [*Lumo Run*], today announced the addition of Mark Mastalir to the team as Vice President of Marketing and Partnerships.

When the company name is stated in full-length and then followed by an abbreviation in brackets, we annotate both mentions as individual company mentions.

If there is a relation mentioned in the sentence, we neglect the mention in the brackets and only create an RM with the full-length mention as the company argument. However, the relation between the company abbreviation and the product is implicitly stated, since both company mentions are connected via Identity/Coref relation. Note that the brackets determine which mention is used as an argument, not the brevity of the mention.

We annotate organizations even if mentioned in introductory/bulk text, or if mentioned as a news source:

- To find out more about **Aflac** and One Day PaySM, visit aflac.com or espanol.aflac.com.
- Featured Company **Aflac** Recognized for Technology Excellence and Leadership **PR Newswire** Monday, January 25th 2016 Company Receives Two Top Honors from **IDG Enterprise/Computerworld** COLUMBUS, Ga. , Jan. 25, 2016 /**PRNewswire**/

If applicable, a organization name might contain Domain-Endings or consist of a Domain-Name:

- About **ReportsnReports.com** **ReportsnReports.com** is an online market research reports library of 500,000+ in-depth studies of over 5000 micro markets.

3.2. Products

We annotate products referenced by names / proper nouns anywhere in a document. We annotate generic product classes ONLY as product arguments (!) of a *CompanyProvidesProduct* relation. When annotating a generic product reference as a relation argument, also annotate any other text occurrences of the same (or slightly changed) character sequence in the document as a product mention, and connect them via the *Identity* relation.

In this chapter, we will present the different elements that a product mention can consist of. We will discuss the categories, their respective functions, their forms of linguistic manifestation as well as their importance. In the second half of this chapter we will portray elements that are (usually) not considered part of a product mention, along with exceptions to the rules.

3.2.1. Definition of Product Entity

We define as a product any commercially available good, be it a finished product, a pre-product, or a part or component of a larger product. While the focus of this work is on non-consumer products, this definition also includes consumer products.

A product does not have to be a tangible object, but can be a service or virtual object. An entity of a very abstract nature can serve as product, if it mentioned in this capacity. The following examples illustrate the different abstraction levels of product mentions:

- **Sensata Technologies** *develops* [sensors] and [controls]
- Part 1 of the giveaway is a paperback novel by Elizabeth Hoyt, [*Wicked Intentions*].
- [*Global Anti-Collision Sensors Market 2016-2020 report*] has been prepared based on an in-depth market analysis with inputs from industry experts.
- **ReportsnReports.com** *offers* [research studies] on agriculture, energy and power, automotive, Semiconductor and Electronics, Manufacturing & Construction, chemicals, environment, medical devices, healthcare, food and beverages, water and much more.

Although they are semantically closely related, product-like entity mentions that refer to an industry sector or branch of business are not treated as products. The industry term can, however, be part of the product mention. Categories such as brand names and trademarks also often appear as part of the product mention.

3.2.2. Elements of a Product Mention

We found a limited set of elements that products usually consist of. This set can be subdivided into seven categories: *company name*, *brand name*, *series*, *model*, *trademark symbol*, *type* and *feature*. Not all of these elements appear in every product mention.

Product mentions can vary strongly in length and complexity, from a single element (a-c), to a combination of any of the categories (d-f) to a coverage of all of the categories (g).

- a) sensors (*type*)
- b) Kleenex (*brand name*)
- c) Q7 (*model*)
- d) Audi Q7 (*company name* and *model*)
- e) Innocent Drinks smoothies (*company name* and *type*)
- f) white iPhone 6 (*feature*, *brand name* and *model*)
- g) Toyota Land Cruiser 100 Series VX SUV diesel turbo (all of the above)

Some categories have shown to be more essential to a product mention than others. A product mention contains at least a common noun, representing the product *type*, or a proper noun that can either stand for a *brand name* or a specific *model*.

Like *brand names* and *models*, the company name that often appears as part of the product mention (but is not essential to it) can also consist of a proper name and therefore include any kind of word class and even punctuation. Most of the seven categories can include or be realized as nouns. However, *series*, *generations* and *models* tend to consist of letters or numbers or combinations of the two. See Table (1) for further specification of the individual categories as well as examples. In the table, the different categories of elements that we consider to be parts of product mentions are listed in the order in which they usually appear when a product mention contains more than one category, with the exception of the category *feature*, which can be found in any position.

Category	Example	Description
company name	[Dunlop] Sport M3 winters	This is the name of the company that provides the product when it is mentioned as part of the product name. It is usually found in the beginning of the product mention. It tends to be a proper noun that does not always follow orthographic rules and sometimes appears as an abbreviation (e.g. "Tumblr", "BMW"). The complexity of a company name tends to be reduced to the main word(s) when it is part of the product mention, i.e., "Toyota" instead of "Toyota Motor Corporation".
brand name	Apple [iPhone] 6S	The brand name is the name under which a certain product is marketed. Usually this is a proper name, however it is not always capitalized or otherwise orthographically correct. Like all categories that can be realized as a proper name, a brand name can include any word class, special characters and punctuation (e.g. "FILL OR BUST!").

series	VW Golf [VII]	This is the part of the product mention that denotes the series, generation, edition or model range. It is often realized as a number, sometimes as a name. Sometimes it includes the word “series” or “generation” or an equivalent abbreviation. Whether or not this can be part of a product mention is dependent on the domain.
model	BMW [i8]	This part of the product mention denotes the specific model of a product in a product series. It often consists of letters and numbers or a combination of the two.
trademark symbol	McRib[®]	The trademark symbol usually appears right after the brand name.
type	Nike Air Max 2016 [running [shoes]]	The type of the product is the broader category or subcategory a product falls into. The type is usually a common noun and can often be found at the end of the product mention. The category can include an attribute that serves as a specifier to the noun and it is the category in which nonspecific terms such as “product” or “solutions” can be included, serving as the head of the product mention, but only if specified further (e.g. “cosmetic product”).
feature	[2006] Ford Mustang [GT] Convertible [2-Door] Samsung Galaxy S7 [32 GB] [black]	Most other relevant aspects of products we will categorize as a feature. Since this is the vaguest category and the one that is most highly domain-dependent, it can be represented by a broad variety of linguistic manifestations. It includes elements such as the year of fabrication, colors, sizes, variants, and special features. Features can appear in almost any position in the product mention.

Table (1): Categories of product mention elements. Square brackets denote the extent of elements.

3.2.3. Elements Excluded from Product Mentions

We do not annotate the following elements as part of the product mention:

Articles

Articles are not annotated as part of the product extent unless they are part of the proper name.

- *[The Tonight Dough TM]*
- **Ben & Jerry's** launches the *[Tonight Dough ice cream]* inspired by Jimmy Fallon.

Annotating the example sentence is a bit intricate, because the article specifies grammatical definiteness of “ice cream” whereas the actual product name “The Tonight Dough” is shortened to only “Tonight Dough” and serves as a specifier of “the ... ice cream”. Capitalization of an article within a sentence can be considered an indication that it is part of the name and therefore the extent of the product.

Prepositional Phrases

Prepositional phrases are generally not part of the extent of the product mention unless they are part of the proper name.

Common cases that may seem like integral parts of the product extent are PPs that begin with the words “for”, “with” or “including”, (introducing purpose, certain features or sub-categories of a product), to name only a few examples. PPs expressing affiliation to a company are also common, and even though they are probably components of some productive patterns, they are not part of the product extent.

- *[High-Frequency 600mA DC-DC Buck Converter] from* **Diodes Incorporated**
- **Asahi Kasei Microsystems** launched a *[6-axis electronic compass]* for mobile devices and GPS-enabled digital cameras.
- **Allegro Microsystems** announced a unique *[dual-channel latch]* with 2D sensing for determining position and direction.
- **Infineon Technologies** launched a highly accurate *[3D magnetic sensor]* with low power consumption for consumer and industrial applications in May 2015.

but:

- **Distech Contols** delivers *[Innovative Solutions for Greener BuildingsTM].*

Relative Clauses

We do not annotate relative clauses following a product as part of the product mention.

- We provide the industry's only *[Grid Analytics System]* that covers the entire distribution network with quickly deployed intelligent sensors and analytics that identify and analyze potential faults and other grid events.
- New South Wales Police has taken delivery of a **[BMW i8]**, the *[hybrid sports car]* that consumes just 2.1-litres per 100km - making it the most fuel efficient car on the force.

Appositions

Appositions are not tagged as part of the product mention they follow but may in some cases be or entail a separate product mention.

- [Sensors], the most important part of your industrial process...
- **Distech Controls'** *[Smart Room Control solution]*, a unified system for the control of HVAC room terminal equipment, lighting, and shades/sunblind, will be installed throughout the building.
- New South Wales Police has taken delivery of a *[BMW i8]*, the [hybrid sports car] that consumes just 2.1-litres per 100km - making it the most fuel efficient car on the force.

Pronouns

Pronouns are not annotated as part of the product extent. As mentioned before, they may serve as a trigger in certain patterns but they are never part of either of the arguments in CompanyProvidesProduct RMs.

- We are pleased **Samsung** has selected us to work in their advanced *[14nm LPP process]*, enabling our solutions to achieve the tremendous speeds and power efficiencies required by today's high-performance systems.

Non-Specifying Attributes

We do not annotate attributes that accompany the product even if they are part of the same noun phrase unless they are used to specify the class or category of the product. While "advanced", for example, is merely a descriptive element, "magnetic" serves as a distinguishing element and is part of the NP compound (see third example).

- We provide the industry's only *[Grid Analytics System]* that covers the entire distribution network with quickly deployed intelligent sensors and analytics that identify and analyze potential faults and other grid events.
- **Parkifi** uses their unique *[sensors]* to detect occupancy in individual parking spaces while delivering such data, along with weather, event and other valuable data, to parking operators in real time.

but:

- The company offers a broad range of advanced *[magnetic sensor products]* and makes vital contributions to...
- The company launched a *[6-axis electronic compass]* for mobile devices and GPS-enabled digital cameras which reduces the number of required components by half and enhances the process efficiency of incorporating electronic compasses in products.
- smart sensors (adjective), communicating sensors (verb, gerund), vision sensors (noun), Hall sensors (proper noun) etc.

When "unnecessary" adjectives or other attributes appear within the noun phrase in between relevant elements of the product, we divide the product mention in two and annotate one RM, only including the product mention closer to the company mention.

- **Rambus'** *[R+]* industry-standard *[interface solutions]* are high-quality, complete PHY solutions designed with...

In this example, descriptive elements that do not serve to specify the product (class) further separate the product's proper name and its type, both of which are relevant elements. Instead of annotating "industry-standard", we annotate two product mentions and create a relation mention between "Rambus" and "R+".

Conjunctions and Punctuation Elements

In most cases, conjunctions and punctuation elements that appear between different parts of product mentions serve to separate two or more product mentions. The distinction between the conjunction "and" or a comma, for example, that links *different elements* describing the same product and the "and" or the comma that links *different products* sharing the same head, however, is not always easy.

- **Rambus** *creates* cutting-edge *[semiconductor]* and *[IP products]*, spanning memory and interfaces to security, smart sensors and lighting.
- **Pericom's** *[analog]*, *[digital]* and *[mixed-signal integrated circuits]* ...

In the first sentence, "products" is both part of "IP products" and "semiconductor products", the latter of which is not spelled in full in this sentence. The second example is similar – here the circuits are distinguished by means of adjectives.

The next example illustrates the segmentation of several product mentions with unclear borders. "Lumo Lift" and "Lumo Back" each might be either a kind of "posture coach" or a kind of "activity tracker" or both, but for lack of specialist knowledge we will analyze such instances as separate mentions.

- Current **Lumo Bodytech** *products include* the *[Lumo Lift]* and *[Lumo Back]* *[posture coaches]* and *[activity trackers]*, as well as the *[Lumo Run smart running shorts]*.

•

Elements other than the head can also be shared by different product mentions:

- *[car tires]* and *[rims]* (= car tires and car rims)

Exception(s):

However, sometimes conjunctions or punctuation elements can be included in the product's extent. This happens in two cases. In the first case the conjunction or punctuation element is part of the product's (proper) name, as the following examples illustrate:

- *[Coffee, Coffee BuzzBuzzBuzz!®]*
- *[FILL OR BUST!]* is a great family game that any number can play.
- Everything But The...®

In the second case, several essential attributes serve to describe the same product. The conjunction linking those attributes is included in the product extent, as shown in the following example:

- **EnOcean** now offers [*wireless and self-powered LED controls*] under the brand name [*Easyfit*].

Here, the conjunction connects two different aspects of the same product, the “LED controls”, that are both “wireless” *and* “self-empowered”.

Sometimes, one or more of the product mention elements appears in brackets. It does not matter whether the material contained in the brackets is an abbreviation of a product feature mentioned right before or if it is an additional element - as long as the relevant information within brackets is placed *in-between other essential elements* of the product, we include it in the mention:

- [Micro-fused Strain Gage (MSG) pressure sensors]
- [insulating (non-metallic) floor standing, wall-mount enclosures]
- [electronic (non-magnetic) enclosures]
- [eXtra-small Form Factor (XFF) sensor]

If the respective element is placed in the bracket after the product mention, then do not include it in the product extent:

- [electronic enclosures] (APS and polyethurane)

Product-Related, Non-Specific Nouns

Nouns like “product”, “product line”, “article”, “manufacture”, “good”, “merchandise” etc. should not be tagged as a product mention when standing alone. These words can serve as triggers in certain sentences. Only when found in the end of a more complex product mention, we annotate them as part of it. Sometimes, the word itself is part of a fixed compound that represents a product type, e.g. “cosmetic product”. As always, if the term is part of the proper name, it is included in the extent also. Whether or not words like “system” or “solution” trigger or are part of a product mention depends on whether or not they refer to a product or a technology. Even more specific terms like “device” or “tool” should only be tagged if they appear in combination with a specifier, such as “medical device”.

3.3. Triggers

Trigger concepts are a generic class of annotations that cover lexical expressions (terms or phrases) or syntactical elements (e.g. possessive marker -s or prepositional constructs) that indicate a specific event type.

Examples of triggers indicating CompanyProvidesProduct relation:

- Lexical trigger: **Sensata Technologies Holding** produces [sensors]
- Syntactic trigger: **BMW's** [Z3]

If a sentence contains more than one trigger for the same relation instance, then as many relation mentions are annotated as there are triggers. The following example contains one company, one product, and the three triggers developer, manufacturer and vendor, each of them referring to a different way of how the product is related to the company -- therefore three relation mentions are created. The motivation for generating a relation mention for each trigger lays in keeping the mentions as simple as possible.

- **FUJIFILM** invested in **Japan Biomedical Co.**, a developer, manufacturer, and vendor of [additives for cell culture media].

4. *The Identity Relation (Coreference)*

In addition to the *CompanyProvidesProduct* relation, we also annotate coreference via a “special” relation type, namely “Identity”. Each *Identity* relation mention corresponds to a single lexical chain connecting all mentions of a specific company or product within the document. Annotating coreferential company and product mentions allows us to label significantly more *CompanyProvidesProduct* relation mentions than by relying on named entity mentions alone. We follow the guidelines for annotating the *Identity* relation with its arguments “source” and “target” described in the supporting coreference annotation guidelines document. An example *Identity* relation is marked up as follows:

- The history of **De Jong Combustion: A. de Jong bv (De Jong)** was founded in 1877, when Albert de Jong bought a shipyard in Scheveningen and started a shipping company. In the 1930's **De Jong** became involved in combustion technology. In 1955 **De Jong** settled in Schiedam with an engineering office and a construction workshop. From that time, the **company** is developing rapidly. Nowadays **it** still designs and develops gas and oil burners for various industrial applications.

In this example, the first mention of “De Jong Combustion” is labeled as the “source” argument, and all other mentions are labeled as “target” mentions of the source concept. We select the most explicit (typically longest) mention found in the document as the source argument.