

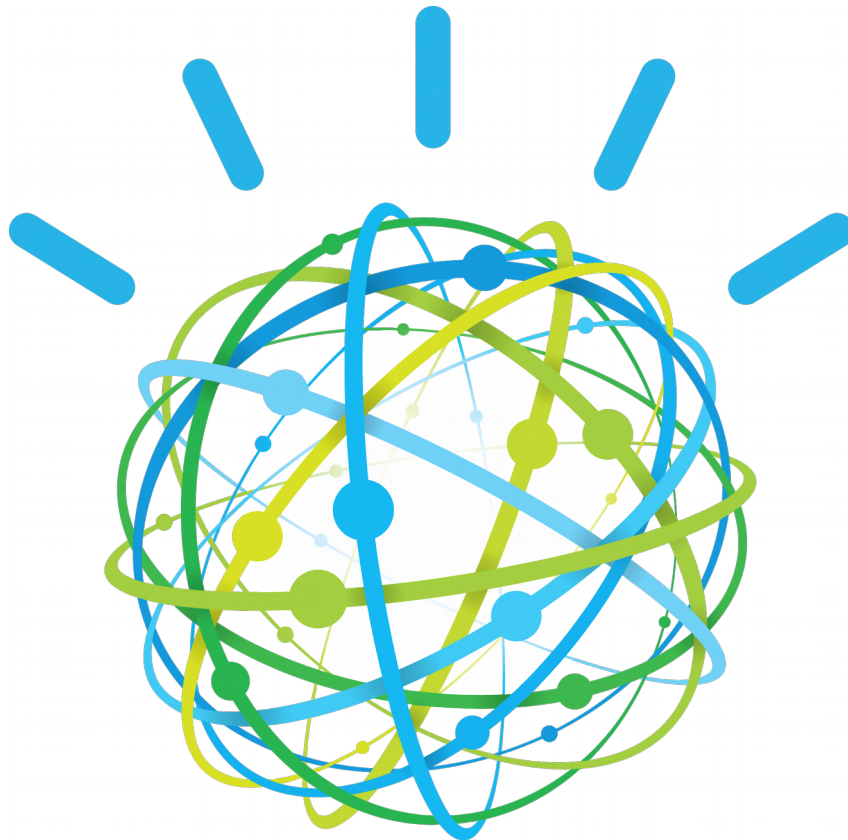
# IBM Watson Knowledge Studio:

## Demo Guide:

### Building a Machine-learning Annotator

#### Cognitive Solutions Application Development

IBM Global Business Partners  
Duration: 30 minutes  
Updated: September 24, 2018  
Klaus-Peter Schlotter  
[kps@de.ibm.com](mailto:kps@de.ibm.com)



# IBM

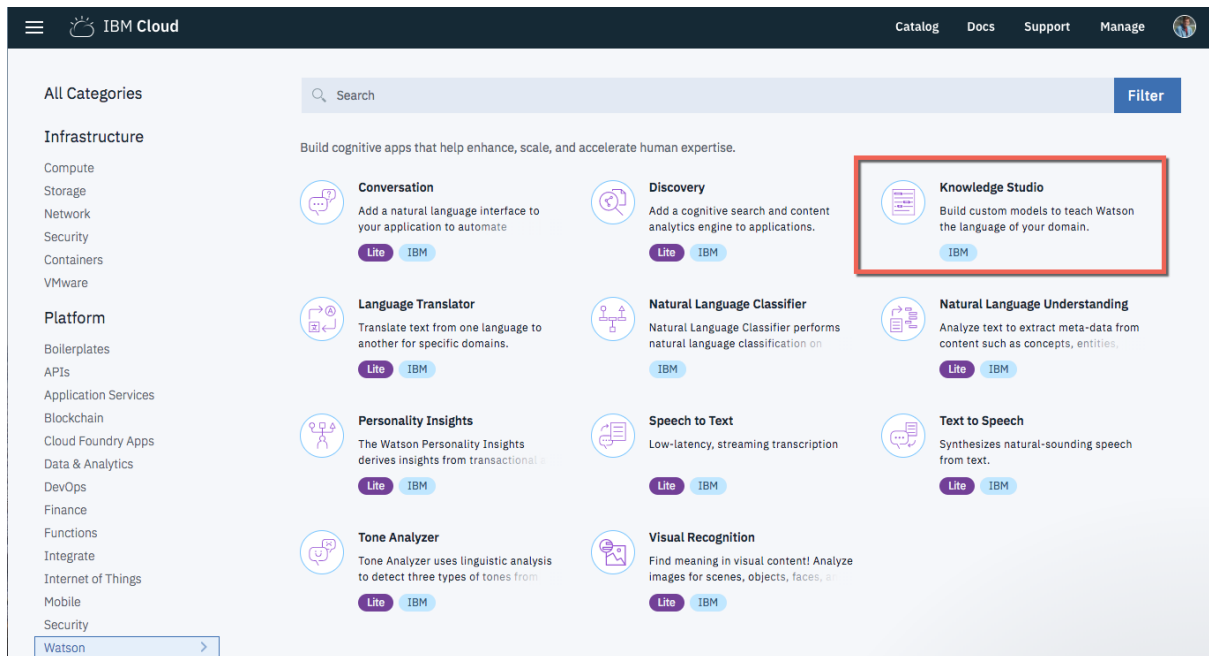
Version 2.0

The document describes the creation of a Machine Learning-annotator with IBM Watson Knowledge Studio (WKS) available as a service on the IBM Cloud Platform.

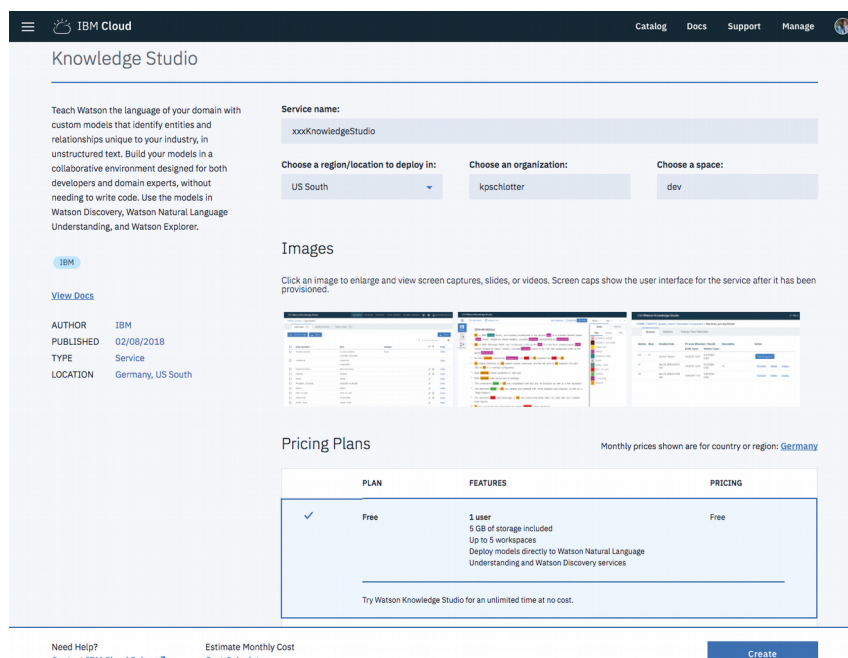
Version 1.x of this document describes this process with WKS available in the IBM Marketplace.

In the demo a new project will be created and configured until **Step 26**, then the model from an already existing project will be used due to the time needed for training and evaluating the new model.

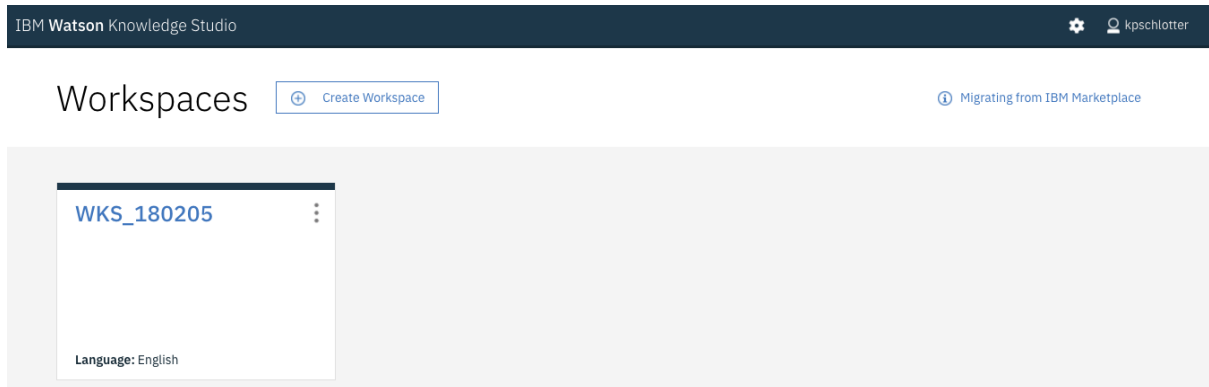
**Step 1** In your IBM Cloud Platform console in the Watson services category, click on the Knowledge Studio service.



**Step 2** Give it a unique *name* (you can also accept the default) and **select the Free plan**. **Click Create**.

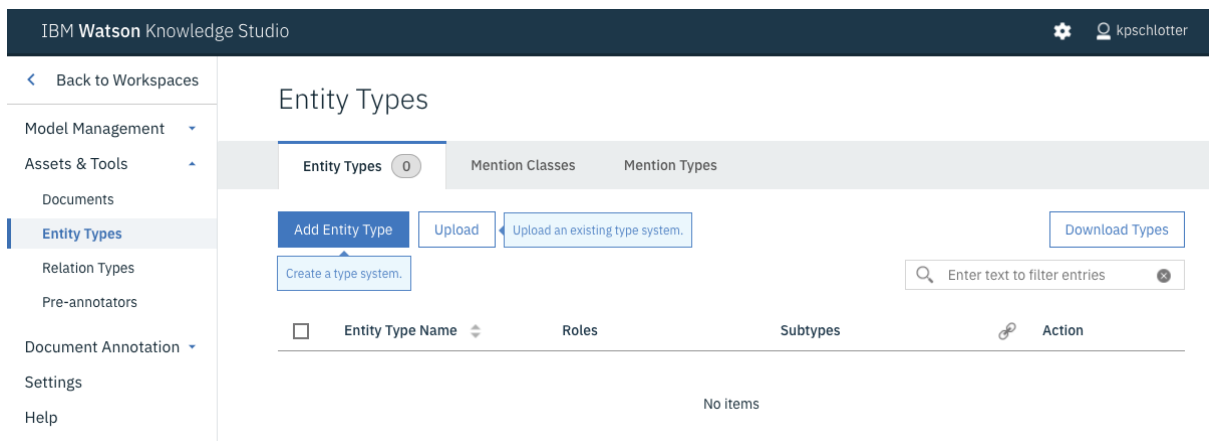


**Step 3** In the service instance click the [Launch](#) button to open Knowledge Studio.

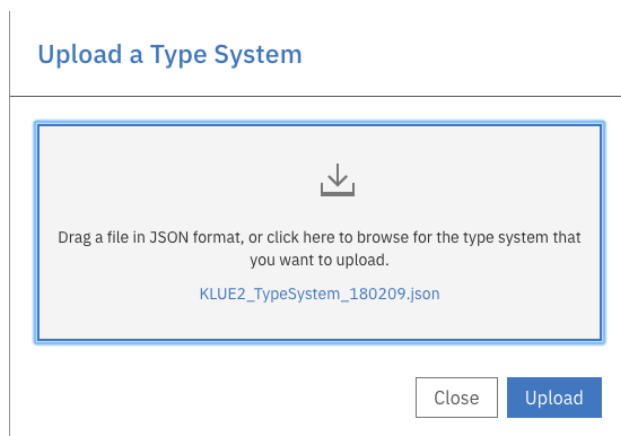


**Step 4** Click the [+ Create Workspace](#) button. Name the workspace and select *English* as the language. Click *Create*.

**Step 5** The workspace opens with the *Entity Types* page displayed.



**Step 6** Upload an existing Type system



The Type system is based on KLUE (Knowledge from Language Understanding and Extraction). It has 52 Entity Types and 2177 Relation Types.

**Step 7** On the *Documents* → *Document Sets* **upload** an existing documents set.

**Step 8** On *Documents* → *Annotation Sets* **Create** an *Annotation Set* with a *name* from the uploaded documents

## Create Annotation Sets

**Base set** All

**Overlap** 100 % of 14 documents

**New Sets** [+ Add another set and human annotator](#)

To view inter-annotator agreement scores, assign at least two human annotators and ensure that a percentage of documents overlap between the sets.

**Annotator** Klaus-Peter Schlotter

**Set name**

Cancel Generate

In the Free plan of WKS you can only have one Annotator. Therefore the Overlap must be 100%

**Step 9** On Assets → Dictionaries ...

**Step 10** Click *Create Dictionary*, enter a Name and click Save.

IBM Watson Knowledge Studio

← Back to Workspaces

Model Management ▾

Assets & Tools ▴

Documents

Entity Types

Relation Types

Pre-annotators

← Manage Dictionaries

Create Dictionary

Name: org\_dict

Cancel Save

**Step 11** In this newly created dictionary, **upload** a dictionary that lists organizations and **map** it to the *Entity Type* “ORGANIZATION”.

**Note:** Creating a dictionary and then importing allows you to modify the imported data.

IBM Watson Knowledge Studio

← Back to Workspaces

Model Management ▾

Assets & Tools ▴

Documents

Entity Types

Relation Types

Pre-annotators

Document Annotation ▾

Settings

Help

← Manage Dictionaries

Create Dictionary

org\_dict (25)

org\_dict

Language: English | 25 entries

Entity type: ORGANIZATION

Rule Class: None

Add Entry Upload Download

<input type="checkbox"/>	Lemma	Surface Forms	Part of Spee...	Action
<input type="checkbox"/>	Beverly Hills Hotel	Beverly Hills Hotel	Noun	<a href="#">Edit</a> <a href="#">Delete</a>
<input type="checkbox"/>	British Academy	British Academy, BAFTA/LA	Noun	<a href="#">Edit</a> <a href="#">Delete</a>
<input type="checkbox"/>	Golden Globes	Golden Globes	Noun	<a href="#">Edit</a> <a href="#">Delete</a>
<input type="checkbox"/>	Writers Guild	Writers Guild, WGA	Noun	<a href="#">Edit</a> <a href="#">Delete</a>

**Step 12** Add an *Entry* for IBM ( it's a **Noun**) and **click Save**.

<input type="checkbox"/>	Lemma	Surface Forms	Part of Speech	Action
<input checked="" type="checkbox"/>	IBM	<input checked="" type="radio"/> IBM <input type="radio"/> International Business M: <input type="radio"/> Big Blue <input type="radio"/> Enter text	Noun	<input type="button" value="Cancel"/> <input type="button" value="Save"/>
<input type="checkbox"/>	Beverly Hills	Beverly Hills Hotel	Noun	<a href="#">Edit</a> <a href="#">Delete</a>

**Step 13** On Machine Learning Model → *Pre-annotator* – The *org\_dict* is mapped to the Entity Type *Organization*. **Click** [Apply This Pre-annotator](#)

**Step 14** **Select** the Annotation document set created in Step 8 and **click Run**.

## Run Annotator

### Documents

Select the document sets or annotation sets that you want to pre-annotate.

☐ documents-new.csv

☒ kpsAnnotation

**Step 15** On *Machine Learning Models* → *Annotation Tasks*, **Click**  . Give it a name, set a Deadline and **click**  .

IBM Watson Knowledge Studio ⚙️ 👤 kpschlottter

[Back to Workspaces](#)

Model Management ▾  
Assets & Tools ▴  
**Documents**  
Entity Types  
Relation Types  
Pre-annotators  
Document Annotation ▾  
Settings  
Help

### Documents

Document Sets (2)   Annotation Sets (1)   Documents (kpsAnnotation, 14)   **Tasks**

▾

**Task name**

kpsAnnoTask

**Deadline**

02/28/2018

**Step 16** Select the annotation set and click *Create Task*.

Add Annotation Sets to the Task

Cancel

Create Task

Annotation Set Name	Annotator Name	Task Name
<input checked="" type="checkbox"/> kpsAnnotation	Klaus-Peter Schlotter	

## Documents

Document Sets (2)

Annotation Sets (1)

Documents (kpsAnnotation, 14)

Tasks

Add Task

Refresh

kpsAnnoTask

Deadline: 02/28/2018

Progress

0%

**Step 17** Click the task and click the Action [Annotate](#).

**Step 18** In the **gmanews.tv** document the organization **IBM** is already annotated by the annotation run in Step 13.

IBM Watson Knowledge Studio

Back to Workspaces

Ready

Alphab... 14pt 1

Entity Mention

Type Subtype Role

j AGE

h ANIMAL

a AWARD

q CARDINAL

d DATE

G DEGREE

J DISEASE

n DURATION

- EMAIL

Technology - gmanews.tv

1 NCR, which counts **IBM** founder Thomas Watson as one of its early employees, said its products and services account for more than \$400 billion in annual commerce and 23 billion consumer serf-service transactions.

2 The company's Philippine office established in the 1960s currently operates a research and development facility in Cebu with about 500 employees.

**Step 19** In this document we **annotate** *Thomas Watson* as a **PERSON** and in the *Relations* section we **drag** the *PERSON* Type **onto** the *ORGANIZATION* type to build the **founderOf** Relation.

IBM Watson Knowledge Studio

Back to Workspaces

Ready

Alphab... 14pt 1

Relation Type

- affectedBy

- affiliatedWith

- agentOf

- authorOf

- awardedBy

- awardedTo

- basedIn

- before

- bornAt

- bornOn

Technology - gmanews.tv

1


founderOf

OR... PERSON

NCR, which counts **IBM** founder Thomas Watson as one of its early employees, said its products and services account for more than \$400 billion in annual commerce and 23 billion consumer serf-service transactions.

2

The company's Philippine office established in the 1960s currently operates a research and development facility in Cebu with about 500 employees.

**Step 20** **Save** the annotated document and return to the document list by **clicking**  at the top.

**Step 21** Back on the documents list, **click** . Now all documents have the status **Completed**.



**Step 22** On *Machine Learning Model* → *Annotation Tasks*, **select** the SUBMITTED *Annotation Set* and **click Accept** the submitted annotations done in previous step.

IBM Watson Knowledge Studio

Back to Workspaces | kpsAnnoTask

Deadline: 09/30/2018

Annotations added to annotation sets are not considered ground truth until the annotation sets are submitted and accepted. When an annotation set is accepted, documents that are annotated in only one annotation set immediately become ground truth. Documents that are annotated in two or more annotation sets become overlapping documents that will become ground truth after conflicts are resolved.

Buttons: Edit Annotation Sets, Refresh, Calculate Inter-Annotator Agreement, Check Overlapping Documents for Conflicts, Apply Type System Updates

Buttons: Accept (highlighted), Reject

Annotation Set Name	Annotator Name	Status	Action
<input checked="" type="checkbox"/> kpsAnnotationSet	Klaus-Peter Schlotter	SUBMITTED	<a href="#">View</a>

Left sidebar: Assets (Documents, Entity Types, Relation Types, Dictionaries), Rule-based Model (Rules, Versions), Machine Learning Model (Pre-annotation, **Annotation Tasks**, Performance, Versions), Settings, Help

The status of the annotation set changes to **COMPLETED**

**Step 23** On *Machine Learning Models* → *Performance* click [Train and evaluate](#) to work with the annotations done in previous steps. Click [Train & Evaluate](#).

**Note:** This step takes approximately 10 - 15 minutes.

IBM Watson Knowledge Studio | kpschlotter

Back to Workspaces | Training / Test / Blind Sets

Select Training/ Test/ Blind Sets

Buttons: Train, Train & Evaluate

Document Set	Task Status
<input type="checkbox"/> All	
<input type="checkbox"/> documents-new.csv	
<input checked="" type="checkbox"/> kpsAnnotation	COMPLETED

Create new sets by splitting the selected document s...

Ratio: Enter the percentage of documents to include in each set.

70	Training Set (70% Recommended)
23	Test Set (23% Recommended)
7	Blind Set (7% Recommended)

Left sidebar: Model Management (Performance, Versions), Assets & Tools, Document Annotation, Settings, Help

With the following Result:

IBM Watson Knowledge Studio | kpschlotter

Back to Workspaces | Performance

Success: Feb 9 10:52 Machine learning model evaluation complete.

WKS\_180209  
Language of documents: English

Donut chart showing: 9 Training Set, 3 Test Set, 2 Blind Set

Buttons: [Train and evaluate](#), [View Ground Truth](#) (Training Set), [View Ground Truth](#) (Test Set), [View Decoding Results](#)

Last trained on: Feb 9, 2018 10:49 AM  
Last evaluated on: Feb 9, 2018 10:52 AM

Low performance? Click here to train.

Left sidebar: Model Management (Performance, Versions), Assets & Tools, Document Annotation, Settings, Help

**Step 24** On *Machine Learning Models* → *Versions* click **Take Snapshot** from the trained model. **Click Deploy**.

#### Version History and Deployment

Version	Base	Creation Date	F1 score (Precision / Recall)		Description	Action	Status
			Entity Types	Relation Types			
1.1		Current Version	0.9 (1 / 0.82)	N/A		<b>Take Snapshot</b>	N/A
1.0		02/09/2018	0.9 (1 / 0.82)	N/A	1st Snapshot	<a href="#">Promote</a> <a href="#">Delete</a> <a href="#">Deploy</a>	N/A

**Step 25** Select *Natural Language Understanding* and click **Next**.

#### Deploy Model v1.0

Select a service to deploy to.

All services require a subscription. [Learn more](#)

☒ Natural Language Understanding

☐ Discovery

☐ AlchemyLanguage

Cancel

Next

**Step 26** Select your *Cloud Region*, application *Space* and *Service instance* and click **Deploy**. As a result, the Model ID is displayed.

#### Deploy Model v1.0

Deploying to Natural Language Understanding

You must have a subscription to the IBM Watson™ Natural Language Understanding service, and know the names of your IBM Cloud space and service instance. [Learn more](#)

##### IBM Cloud Information

IBM Cloud is the IBM cloud platform. Click [here](#) to open IBM Cloud and create an account or look up details for an existing service.

Region

US South

Space

dev

Service instance

my-nlu

Cancel

Deploy

#### Deployment Started.

##### Deploying to Natural Language Understanding

It might take a few minutes for publishing and deployment to complete, and for this model to be available to your applications. You can view your deployed models, withdraw a model from deployment, or deploy a newer version.

Model ID `10:f3222fca-76f7-4d25-ae3e-016cec77d9a0`

You can [view documentation](#) to learn how to implement the deployed model into your application.

OK

## Step 27 Testing the Model with Postman.

- a) **List your deployed models.** You need Basic Auth (username and password) from your Natural Language Understanding Service from your Bluemix account and the following URL:

```
https://gateway.watsonplatform.net/natural-language-understanding/
api/v1/models?version=2017-02-27
```

▶ NLU\_ListModel

Examples (0) ▼

GET ▼

https://gateway.watsonplatform.net/natural-language-understanding/api/v1/models?version=2017-02-27

Params

Send ▼

Save ▼

Authorization ●

Headers

Body

Pre-request Script

Tests

Code

Type

Basic Auth ▼

Clear

Update Request

Username

c461252-995f-445d-b51c-f89cbfb83...

The authorization header will be generated and added as a custom header

Password

.....

☐ Save helper data to request

☐ Show Password

Response:

```
Body    Cookies    Headers (12)    Tests
Pretty  Raw    Preview    JSON
1 {
2   "models": [
3     {
4       "status": "available",
5       "model_id": "10:8e[REDACTED]9",
6       "description": ""
7     }
8   ]
9 }
```

- b) Test the model with the Basic Auth information from previous step and in the header section add**

```
key: Content-type      value: application/json
```

## In the URL exchange models with analyze

In the Body section enter text you want to analyze and service features you want to use:

```
{ "text": "NCR, which counts later IBM Thomas Watson as one of its early employees, said its products and services account for more than $400 billion in annual commerce and 23 billion consumer selfservice transactions2", "features": { "relations": { "model": "10:8ec45dbd-5daf-450f-8a99-c924a7bce649" } } }
```

```

Authorization  Headers (2)  Body  Pre-request Script  Tests
form-data  x-www-form-urlencoded  raw  binary  JSON (application/json)
1 {
2   "text": "NCR, which counts later IBM Thomas Watson as one of its early employees, said its products and services
3   account for more than $400 billion in annual commerce and 23 billion consumer selfservice transactions2",
4   "features": {
5     "relations": {
6       "model": "10:8ec45dbd-5daf-450f-8a99-c924a7bce649"
7     }
8   }
9 }

```

Response:

```

Body  Cookies  Headers (12)  Tests  Status: 200 OK  Time: 1184 ms
Pretty  Raw  Preview  JSON
1 {
2   "relations": [
3     {
4       "type": "founderOf",
5       "sentence": "NCR, which counts later IBM Thomas Watson as one of its early employees, said its products and
6       services account for more than $400 billion in annual commerce and 23 billion consumer selfservice
7       transactions2",
8       "score": 0.790493,
9       "arguments": [
10        {
11          "text": "Watson",
12          "entities": [
13            {
14              "type": "PERSON",
15              "text": "Watson"
16            }
17          ],
18          "text": "IBM",
19          "entities": [
20            {
21              "type": "ORGANIZATION",
22              "text": "IBM"
23            }
24          ]
25        }
26      ]
27    },
28    "language": "en"
29  ]
30 }

```