31st Annual INCOSE international symposium

virtual event

July 17 - 22, 2021

Michael Riesener, Christian Dölle, **Annika Becker**, **Sofia Gorbatcheva**, Eric Rebentisch, Günther Schuh

Application of natural language processing for systematic requirement management in model-based systems engineering

www.incose.org/symp2021

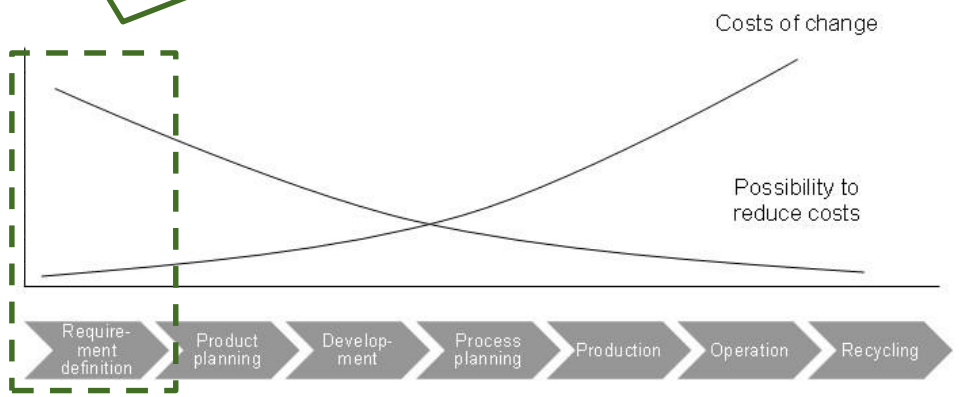
Motivation

Challenges during the product development…

- Growing global competition
- Demand for individualized products
- Shorter Time-to-Market
- Interaction of the different engineering disciplines mechanics, electric/electronic (E/E) and software

… end up in a high number of requirements

Unstructured documents and libraries
Unclear processes and dependencies
Time and cost intensive review and structuring



Model-based Systems Engineering (MBSE) as a possible solution.

Sources: Pinque et. al. 2016, Abele 2016, Bernard 2012, Arora et al. 2015, Arellano et al. 2015, Micouin et al. 2018, Walter et al. 2019

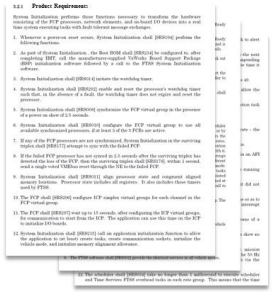
What is meant by MBSE in our context?

MBSE

- Enables structured processing of requirements within the product development process
- Uses model-based simulation
- Allows early validation of requirements and their fulfilment
- Various modeling languages have been introduced (e.g. SysML)
- Creates system models and shows dependencies
- **Needs formalized and structured input**

NLP links MBSE and text-based requirements

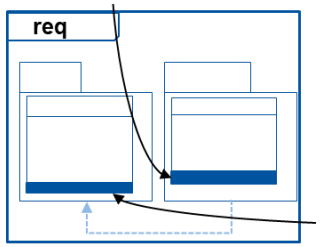
TEXT-BASED REQUIREMENTS



- In practice, the **natural language** is the most commonly used language for documenting requirements
- No specific syntax or structure
- Expressions without restriction are allowed and are **easier to understand** for stakeholders without technical background

NLP

MBSE



ID	Name	Text
2	Performance	The Hybrid SUV shall have a barking, acceleration, and road capability of a typical SUV, but have dramatically better fuel economy.
2.1	Braking	The Hybrid SUV shall have a barking capability of a typical SUV.
2.2	FuelEconomy	The Hybrid SUV shall have dramatically better fuel economy than a typical SUV.
2.3	OffRoadCapability	The Hybrid SUV shall have off-road capability of a typical SUV.
2.4	Acceleration	The Hybrid SUV shall have acceleration of a typical SUV.

ID	Name	Relation	ID	Name	Relation	ID	Name
2.1	Braking	deriveReqt	d.1	RegenerativeBraking			
2.2	FuelEconomy	deriveReqt	d.1	RegenerativeBraking			
2.2	FuelEconomy	deriveReqt	d.2	Range			
4.2	FuelEconomy	deriveReqt	d.2	Range			
2.3	OffRoadCapability	deriveReqt	d.4	Power	deriveReqt	d.2	PowerSourceManagement
2.4	Acceleration	deriveReqt	d.4	Power	deriveReqt	d.2	PowerSourceManagement
4.1	CargoCapacity	deriveReqt	d.4	Power	deriveReqt	d.2	PowerSourceManagement

- SysML models requirements according to a **defined syntax**
- One diagram type for requirements is the **tabular format**, among others, it contains a requirement ID, a title and the requirement text.
- In addition, **user-defined property categories** can be added.

A method is needed to convert text-based requirements into MBSE fitting structures by using **Natural Language Processing**.

What is meant by NLP in our context?

Natural Language Processing

- NLP is defined as a field in computer science, engineering and artificial intelligence (AI) with roots in computational linguistics
- NLP enables computers to understand unstructured content, make derivations and add context to lanague
- The main task of NLP is to develop applications that provide an interaction between machines and Natural Language by processing and understanding human Natural Language and to perform specific tasks to generate useful output
- Thus, it enables to handle large amounts of natural language in order to transform it into data which can be processed and analysed
- → **Offers the potential to process unstructured text-based requirements**

Sources: Banerjee 2020

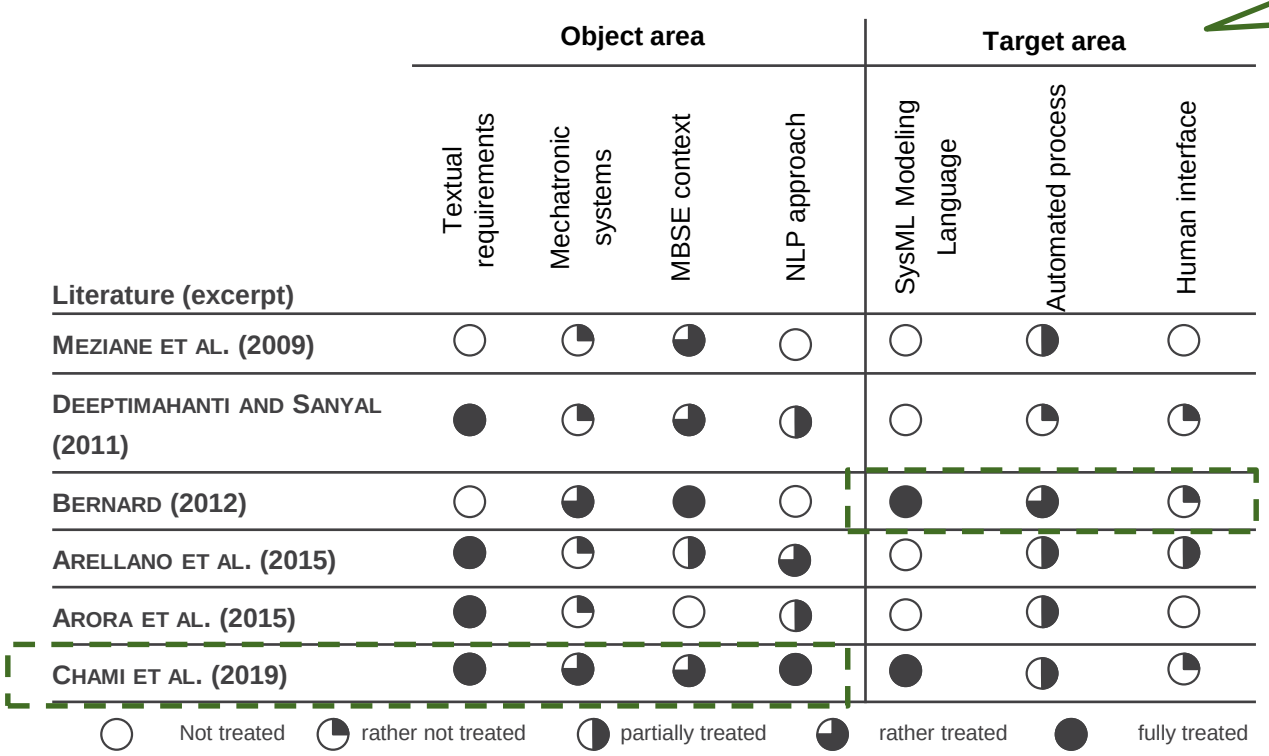
Existing Approaches and Research Gap

Identified Research Gaps

Object area: „context“ of the methodology

Target area: sub-targets to be examined

Object area mostly covered. NLP is applied to create models from semi-structured input text. Performs poorly when used on new input text



Target area mostly covered. Method transforms property-based requirements directly into MBSE common modeling languages. Cannot be formally applied for textual written requirements

Sources: Meziane et al. 2009, Deeptimahanti and Sanyal 2011, Bernard 2012, Arellano et al. 2015, Arora et al. 2015, Chami et al. 2019

The developed model consists of three steps

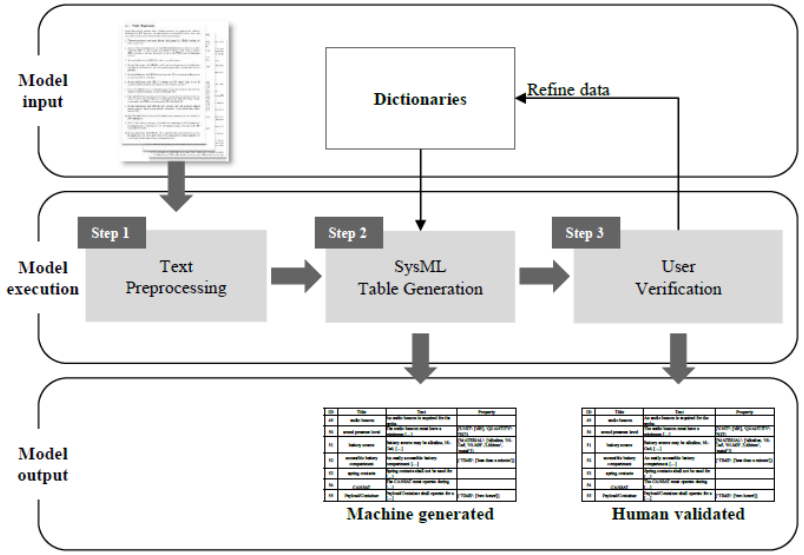
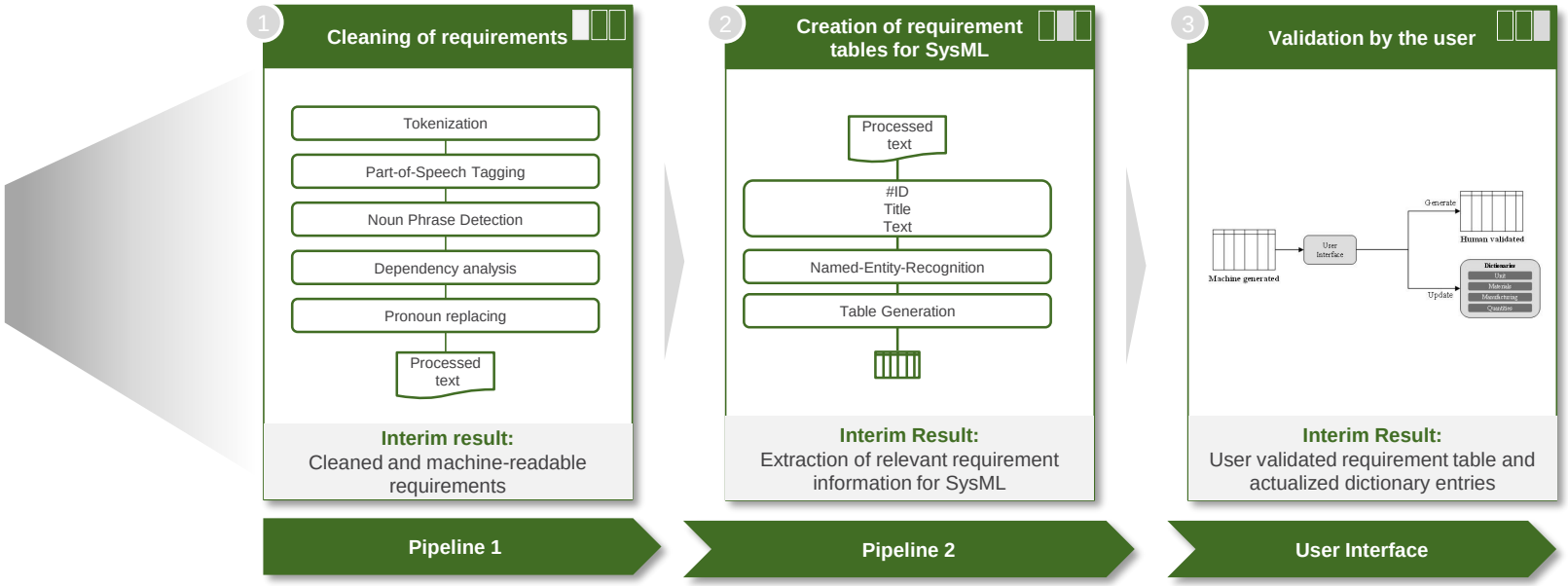
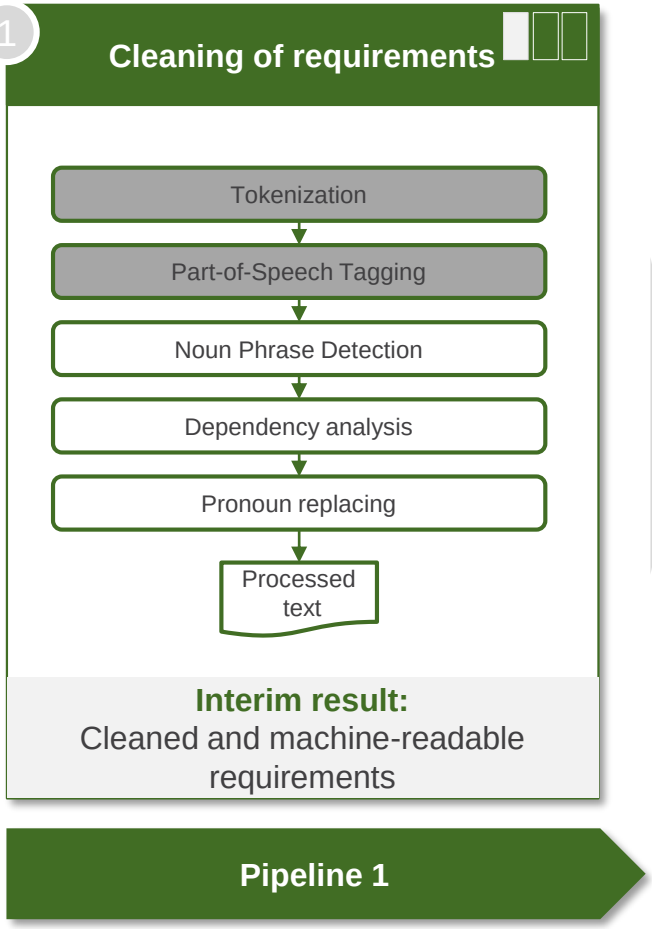


Figure 1. Model for requirement transformation compatible for MBSE



Model execution

Step 1 – Preprocessing of the Text



TONKENIZATION – Separation of a sentence in single words

| An | audio | beacon | is | required | for | the | probe | . | It | may | be | powered | after | landing | or | operate | continuously | . |

PART-OF-SPEECH TAGGING – Assignment of a function to each token

An	DT
audio	JJ
beacon	NN
is	VBZ
required	VBN
for	IN
the	DT
probe	NN
.	.
It	PRP
may	MD
be	VB
powered	VBN
after	IN
landing	VBG
or	CC
operate	VBP
continuously	RB
.	.

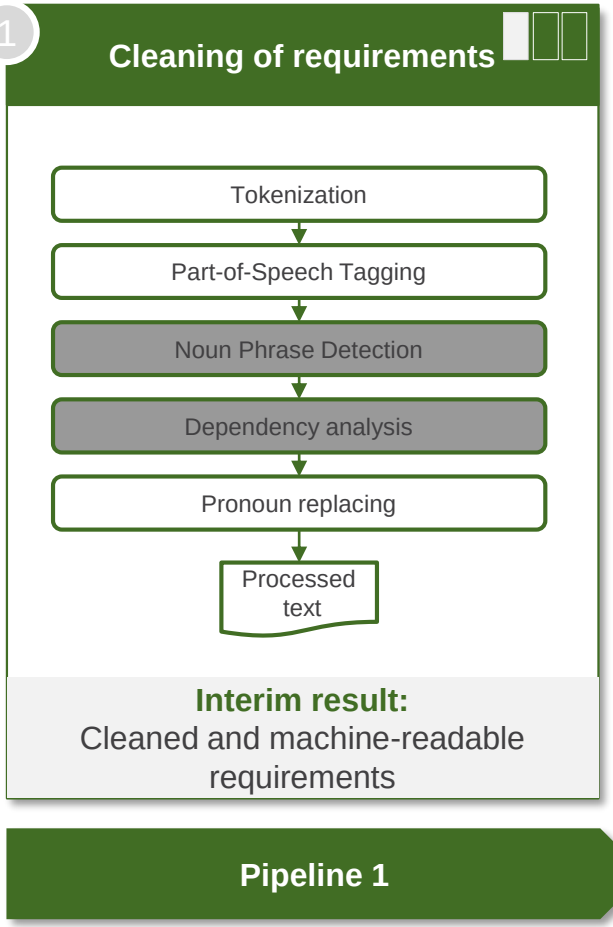
JJ: adjective

NN: noun

Functions according to ISO/IEC 19514:2017 2020

Model execution

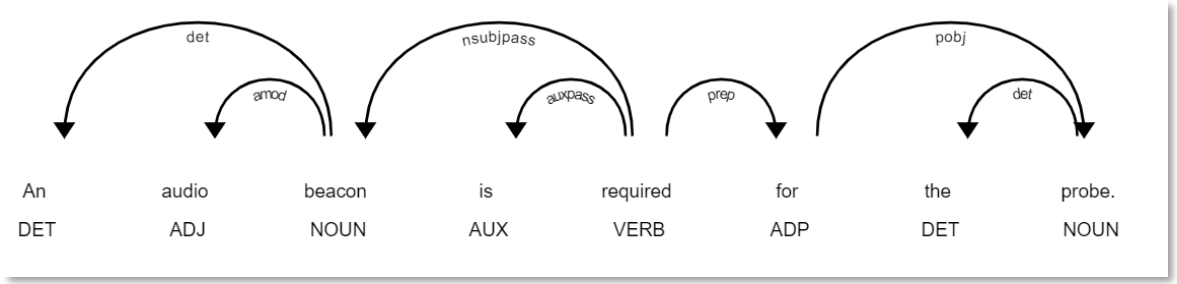
Step 1 – Preprocessing of the Text



NOUN PHRASE DETECTION – Grouping of token, that are related by content

```
The audio beacon
a minimum sound pressure level
92 dB
Battery source
Lithium polymer batteries
Lithium cells
a metal package
18650 cells
An easily accessible battery compartment
```

DEPENDENCY ANALYSIS – Determination of semantic and syntactic relations



Model execution

Step 1 – Preprocessing of the Text

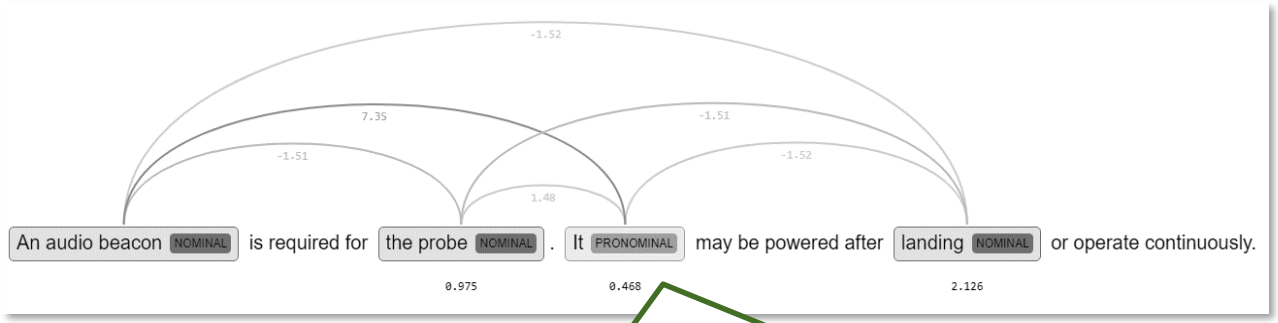
1 **Cleaning of requirements**

- Tokenization
- Part-of-Speech Tagging
- Noun Phrase Detection
- Dependency analysis
- Pronoun replacing
- Processed text

Interim result:
Cleaned and machine-readable requirements

Pipeline 1

PRONOUN REPLACING – Reduce ambiguity by using pronouns

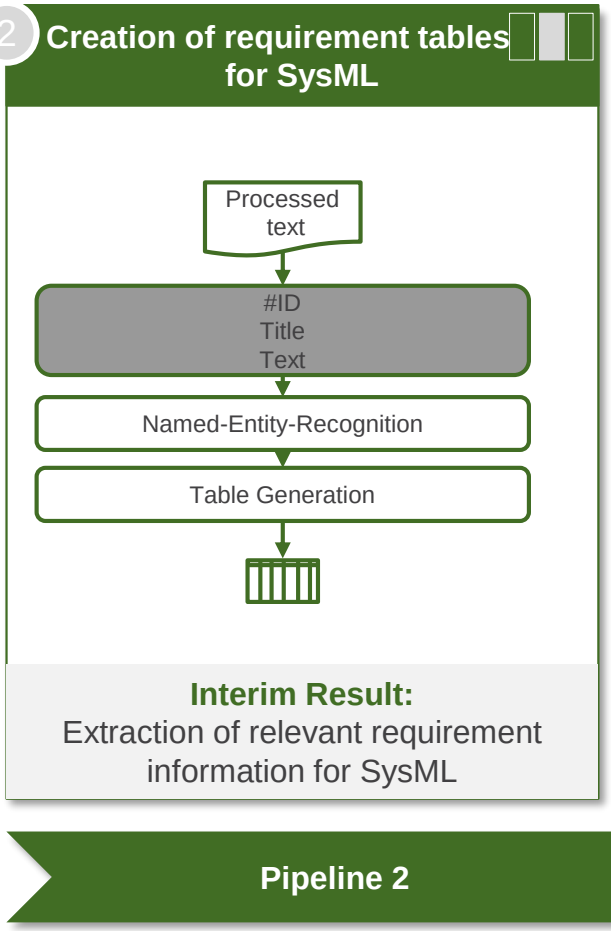


Co-referencing: use of two of more terms in a text which rely on the same entity, typically a pronoun → the pronoun will be replaced by the relative noun to reduce ambiguity

Outcome: Cleaned and machine-readable requirements

Model execution

Step 2 – Creation of tables



ID-GENERATION – Creation of an ID and a title

ID Generation

ID	Requirement
54.1	An audio beacon is required for the probe.
54.2	An audio beacon may be powered after landing or operate continuously.

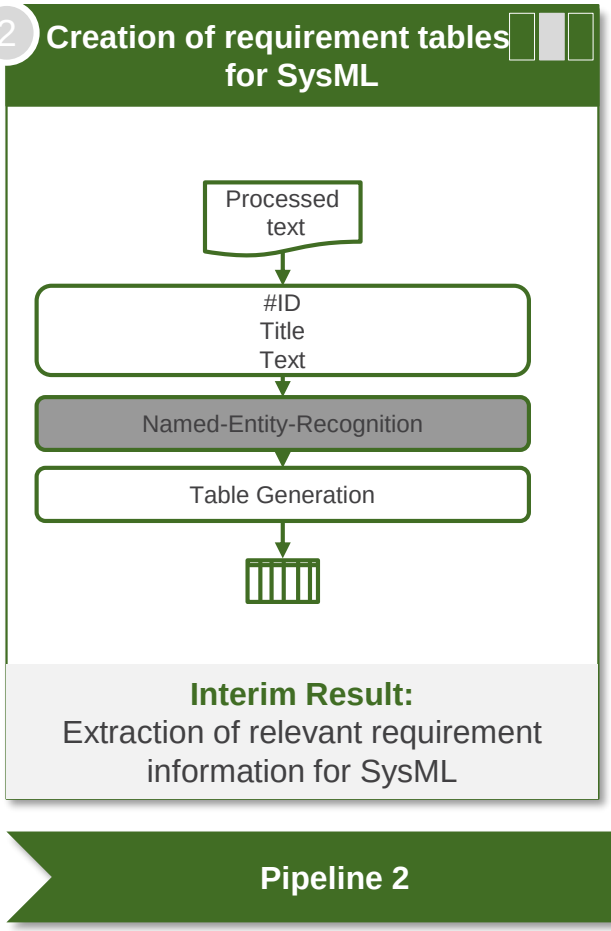
Title Generation

Option	Title	Example	Requirement Title
1	Direct Object	The system shall purify water.	Water
2	Direct Object with verb	The system shall provide water.	Provide water
3	Indirect object	The system shall provide water to the boiler.	Boiler
4	Subject	The system shall provide water to the boiler within 2s.	System
5	Requirement ID	RQ4: The system shall provide water.	4

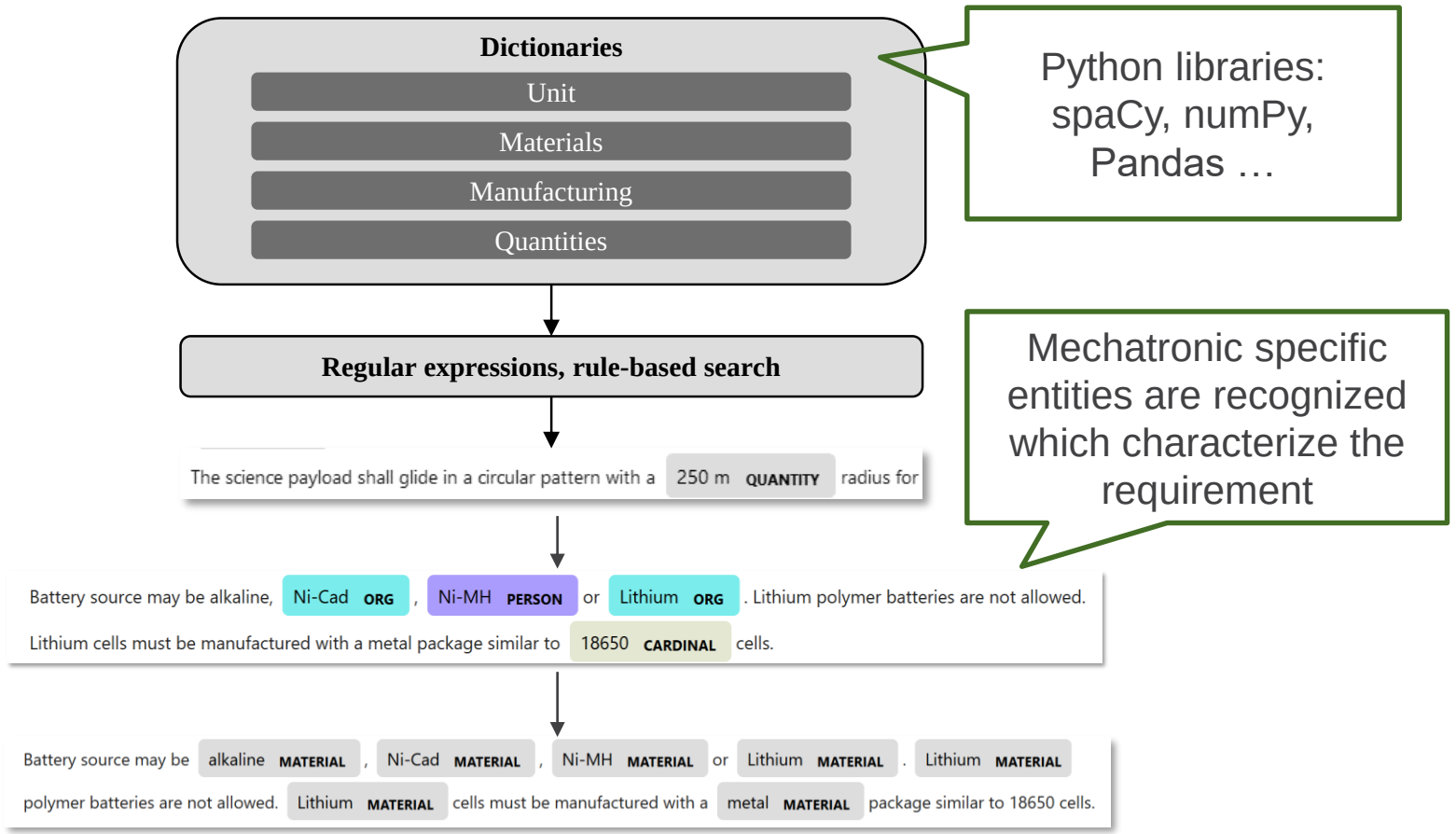
Requirement Titles based on recognized parts of speech

Model execution

Step 2



Named-Entity-Recognition – identification of mechatronic specific entities



Model execution

Step 2

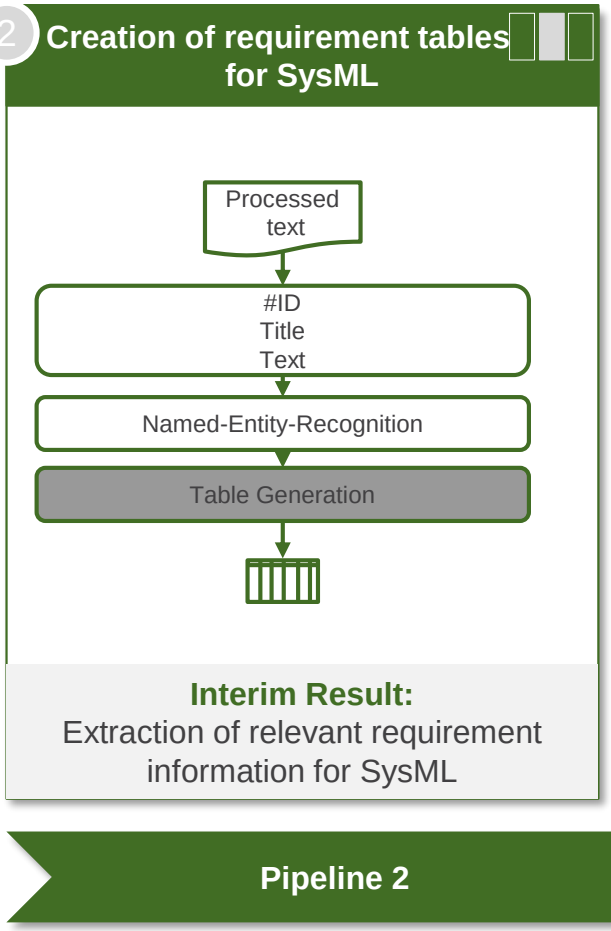


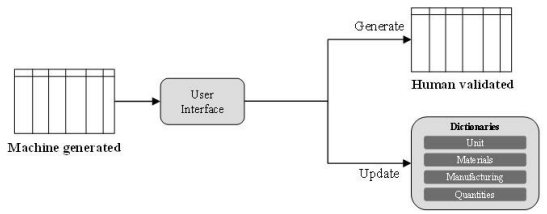
TABLE GENERATION – Summary of information in a table of requirements

id	Title	Text	Entity
18	Proper mounts	All electronics shall be hard mounted using proper mounts such as standoffs, screws, or high performance adhesives.	{'MANUFACTURING': ['screws, ', ' adhesives', ' mounted ']}
		…	
50.1	An audio beacon	An audio beacon is required for the probe.	{}
50.2	Operate an audio beacon	An audio beacon may be powered after landing or operate continuously.	{}
51	A minimum sound pressure level	The audio beacon must have a minimum sound pressure level of 92 dB, unobstructed.	{'UNIT': ['92 db, ']}
52.1	Battery source	Battery source may be alkaline, Ni-Cad, Ni-MH or Lithium.	{'MATERIAL': [' alkaline, ', ' lithium']}
52.2	Lithium polymer batteries	Lithium polymer batteries are not allowed.	{'MATERIAL': ['lithium ']}
52.3	Lithium cells	Lithium cells must be manufactured with a metal package similar to 18650 cells.	{'MATERIAL': ['lithium , 'metal']}

Model execution

Step 3

3 **Validation by the user**



Interim Result:
User validated requirement table and actualized dictionary entries

User Interface

USER INTERFACE – Enables the user to improve the performance of the algorithm without knowledge of NLP

```
Choose validation mode: [1: line-by-line, 2: id range, 3: id list, 4: help]
>>3

Id list selected
Pleas pass list of ids as comma separated values i.e.: 4,7,9,12
>>52.1

1 elements found in the list. Correct?[y/n]
>>y

id      name                  text                                                  entity                                          correct
52.1    Battery source        Battery source may be alkaline, Ni-Cad, Ni-MH or Lithium.      {'MATERIAL': [' alkaline, ', ' lithium']}
Are the entities correct? [y/n] (stop: s)
>>n

Available entities: ['materials','units','manufacturing']
Which entity is missing? (0 for materials, 1 for units, 2 for manufacturing)
>>0

What is the label of the missing entity?
>>Ni-Cad

Updated materials.xlsx successfully with 1 new label(s).
Updated units.xlsx successfully with 0 new label(s).
Updated manufacturing.xlsx successfully with 0 new label(s).
Success! Thanks for validating.
```

E.g.: User can have a check if all entities are correctly recognized by the algorithm

E.g.: Missing entities are inserted by the user and added to the specific dictionary.
A new corrected requirement table is generated and saved

Evaluation and critical reflection of the method

ID Generation

id
50.1
50.2
51
52.1
52.2
52.3
53

- All IDs are correctly allocated
- Requirements consisting of two sentences are devided into two sentences
- Sub-IDs are correctly allocated

Co-Referencing

The science payload shall measure **its** battery voltage.
The science payload shall measure The science payload battery voltage.

- Succesful implementation of Co-Referencing
- The more convoluted the requirements are formulated, the more difficult it is to replace the pronoun

Title Generation

id	Name
18	Proper mounts
19	All mechanisms configuration
20	Pyrotechnics
21	Vegetation
22	An air pressure sensor
23	Gps
24	Science payload battery voltage
25	Measure science payload
26	Particulates
27	Air speed

- Each title is uniquely chosen and describes the requirement as specifically as possible
- For some requirements the results of title generation are rather inaccurate

Dictionary Analysis

- The specific keyword of the text request is recognized, and the corresponding entity is assigned
- Chemical formulas are recognized less frequently

Entity
{'UNIT': ['900 mhz ']}
{'UNIT': ['92 db, ']}
{'MATERIAL': [' lithium', ' alkaline, ']}
{'MATERIAL': ['lithium ']}
{'MATERIAL': ['lithium', 'metal']}

Conclusion and Outlook

Conclusion

- Method to automatically generate SysML requirement tables out of text-based requirement specifications
- Improved exchange of information and communication between different disciplines
- Reduction of the manual workload
- Resulting tables can be used as input for further model analysis

Outlook

- In order to establish intensive domain understanding, ML algorithms need to be trained to detect requirement properties based on the given context
- Natural Language Processing needs to be established in the industrial context

Thank you for your kind attention!

31st Annual INCOSE international symposium

virtual event

July 17 - 22, 2021

www.incose.org/symp2021

Annika Becker, M.Sc. RWTH

Group leader Complexity Management
Department Innovation Management
Laboratory of Machine Tools and Production Engineering of RWTH Aachen
Campus-Boulevard 30, 52074 Aachen
Mobile: +49 (0) 151 431 90382
Mail: a.becker@wzl.rwth-aachen.de

Enter your title here

Enter your title here

Enter your title here

Enter your title here

Enter your title here

What is meant by MBSE and NLP?

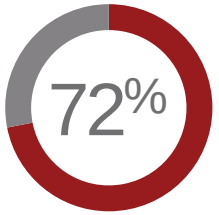
MBSE

- Enables structured processing of requirements within the product development process
- Uses model-based simulation
- Allows early validation
- Various modeling languages have been introduced (e.g. SysML)
- creates system models and shows dependencies
- Needs formalized and structured input

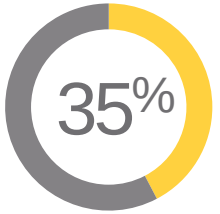
Natural Language Processing

- Offers the potential to process unstructured text-based requirements
- Necessary to structure text-based requirements from heterogeneous sources used in the industries

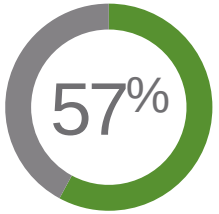
Both necessary elements to enable an automized process from unstructured information to MBSE readable elements.



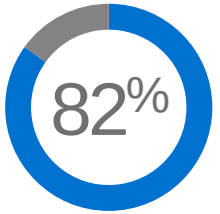
This is a sample text. You can replace Enter your text here.



This is a sample text. You can replace Enter your text here.



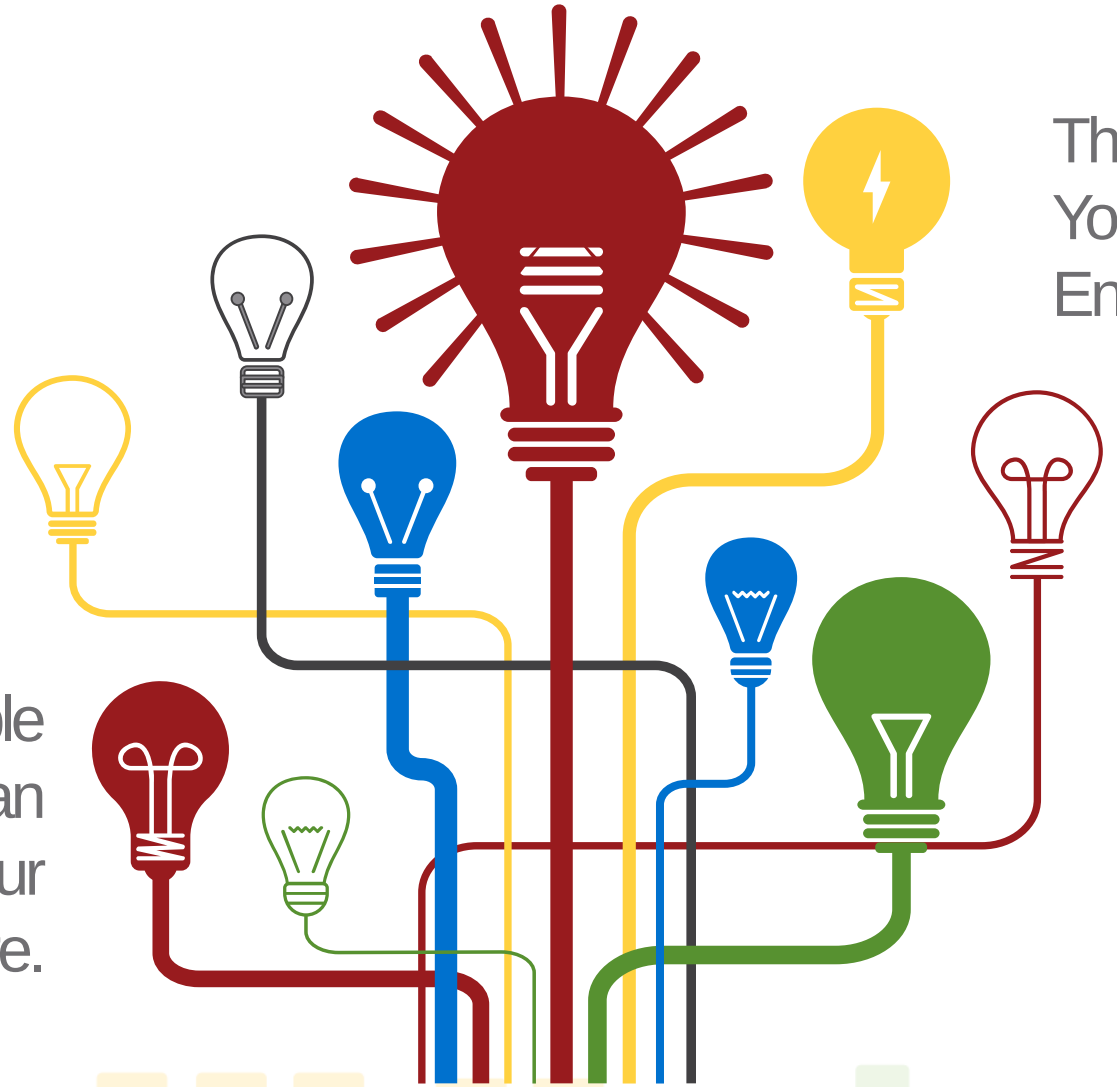
This is a sample text. You can replace Enter your text here.



This is a sample text. You can replace Enter your text here.

This is a sample text. You can replace Enter your text here.

This is a sample text. You can replace Enter your text here.



This is a sample text. You can replace Enter your text here.

This is a sample text. You can replace Enter your text here.

Enter your text here

This is a sample text. Enter your text here

Enter your text here

This is a sample text. You can replace this text. Enter your text here. This is a sample text. You can replace this text. Enter your text here

Enter your text here

This is a sample text. Enter your text here

Enter your text here

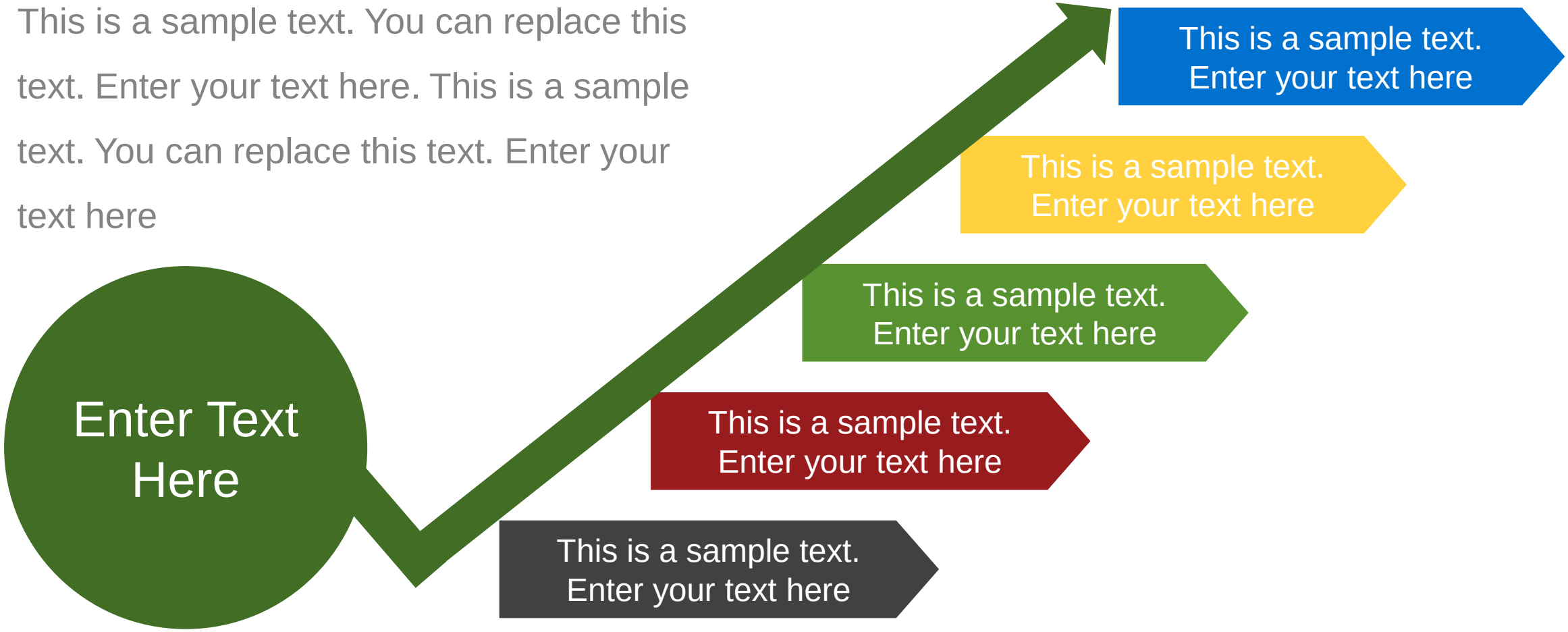
This is a sample text. Enter your text here

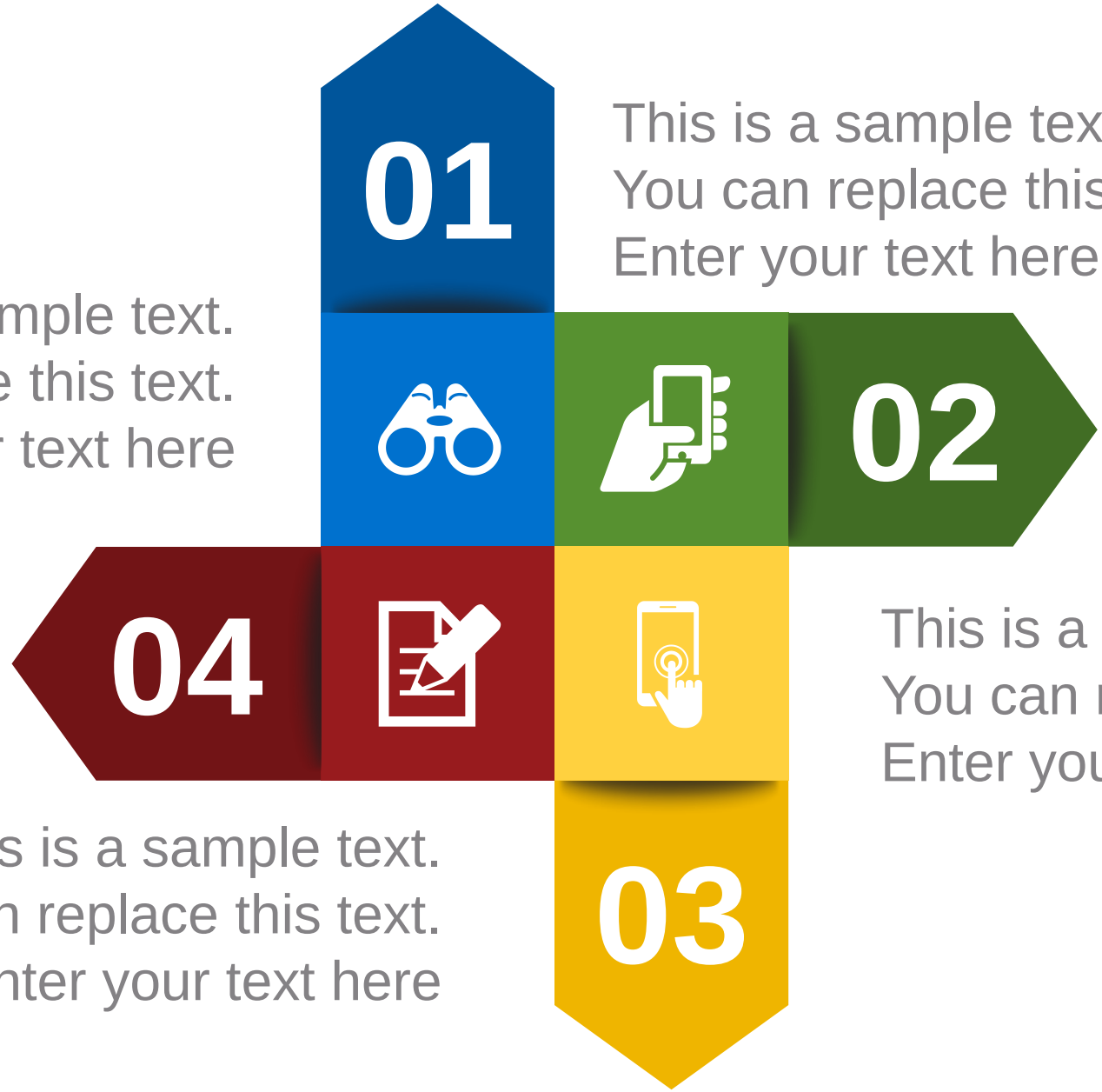
56%

Enter your text here

This is a sample text. Enter your text here

This is a sample text. You can replace this text. Enter your text here. This is a sample text. You can replace this text. Enter your text here





This is a sample text.
You can replace this text.
Enter your text here

This is a sample text.
You can replace this text.
Enter your text here

This is a sample text.
You can replace this text.
Enter your text here

This is a sample text.
You can replace this text.
Enter your text here

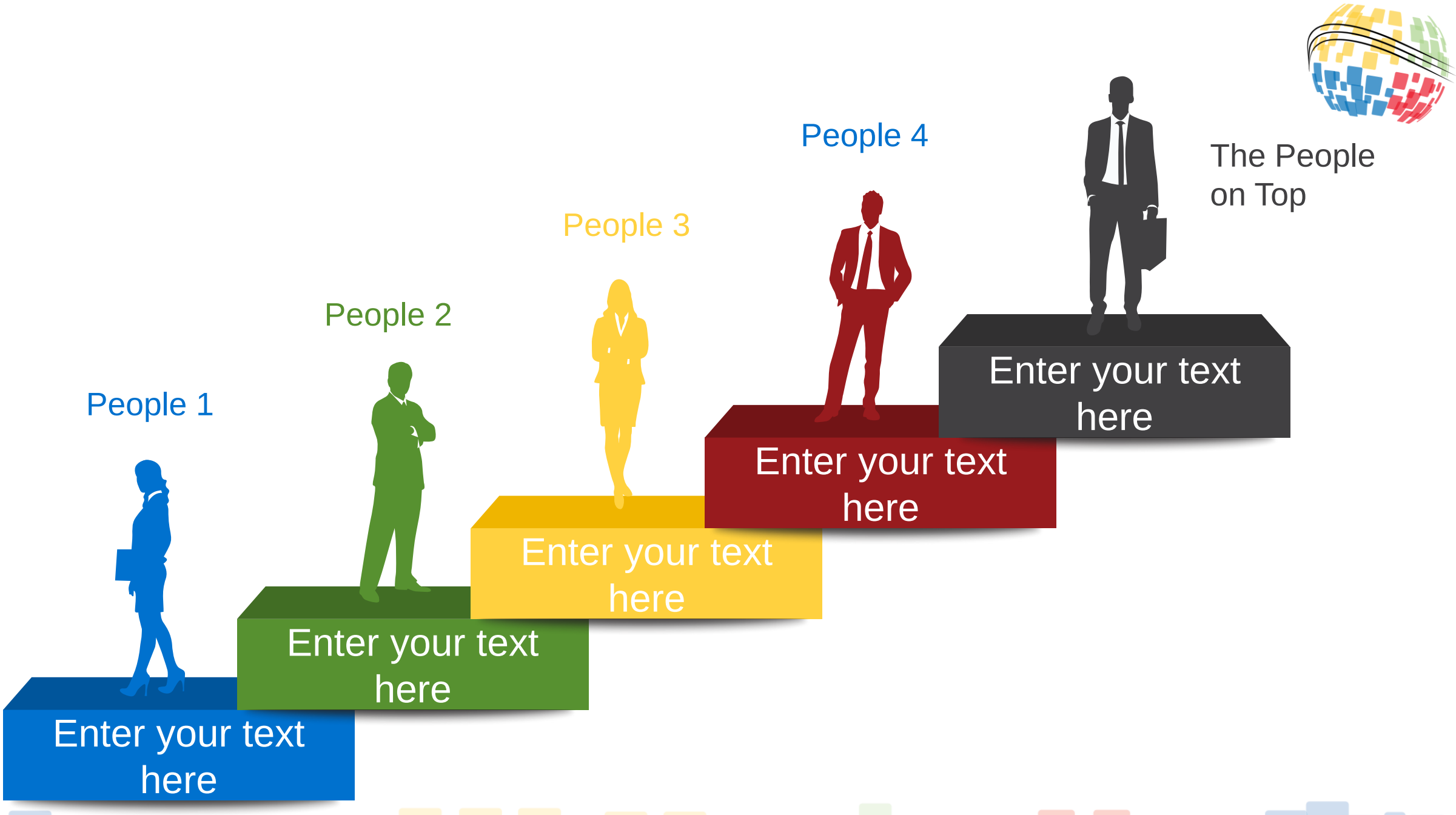
Section 01

Section 02

Section 03

Section 04

This is a sample text. You can replace this text. Enter your text here


People 1
People 2
People 3
People 4
The People on Top
Enter your text here
Enter your text here
Enter your text here
Enter your text here
Enter your text here

This is a sample text. You can replace this text. Enter your text here.

This is a sample text. Enter Your text Here

This is a sample text. Enter Your text Here

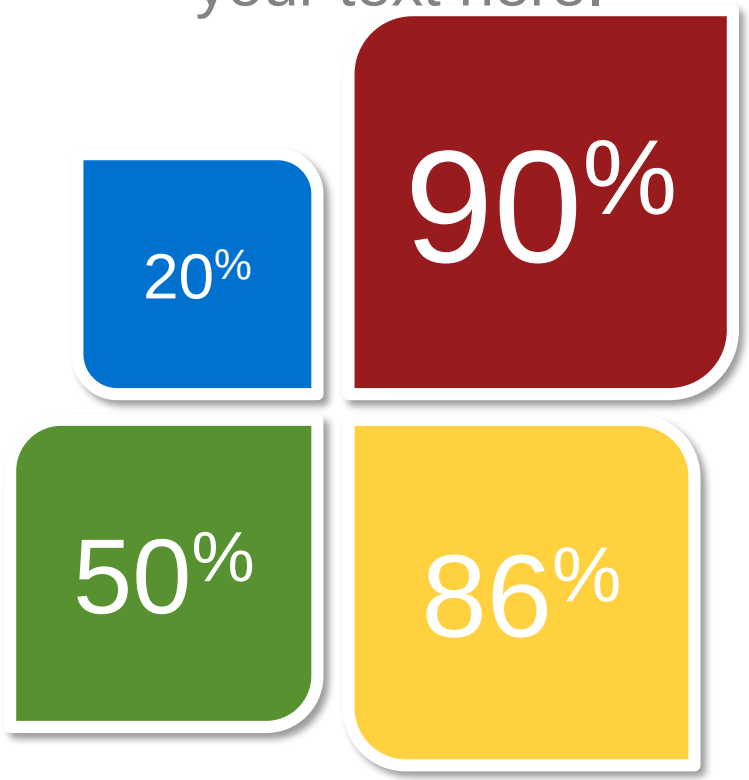
This is a sample text. Enter Your text Here

This is a sample text. Enter Your text Here

This is a sample text. Enter Your text Here

This is a sample text. You can replace this text. Enter your text here.

This is a sample text.
Enter Your text Here



This is a sample text.
Enter Your text Here

This is a sample text.
Enter Your text Here

This is a sample text.
Enter Your text Here

This is a sample text.
Enter Your text Here

This is a sample text. You can replace this text. Enter Your text Here

This is a sample text.
Enter Your text Here

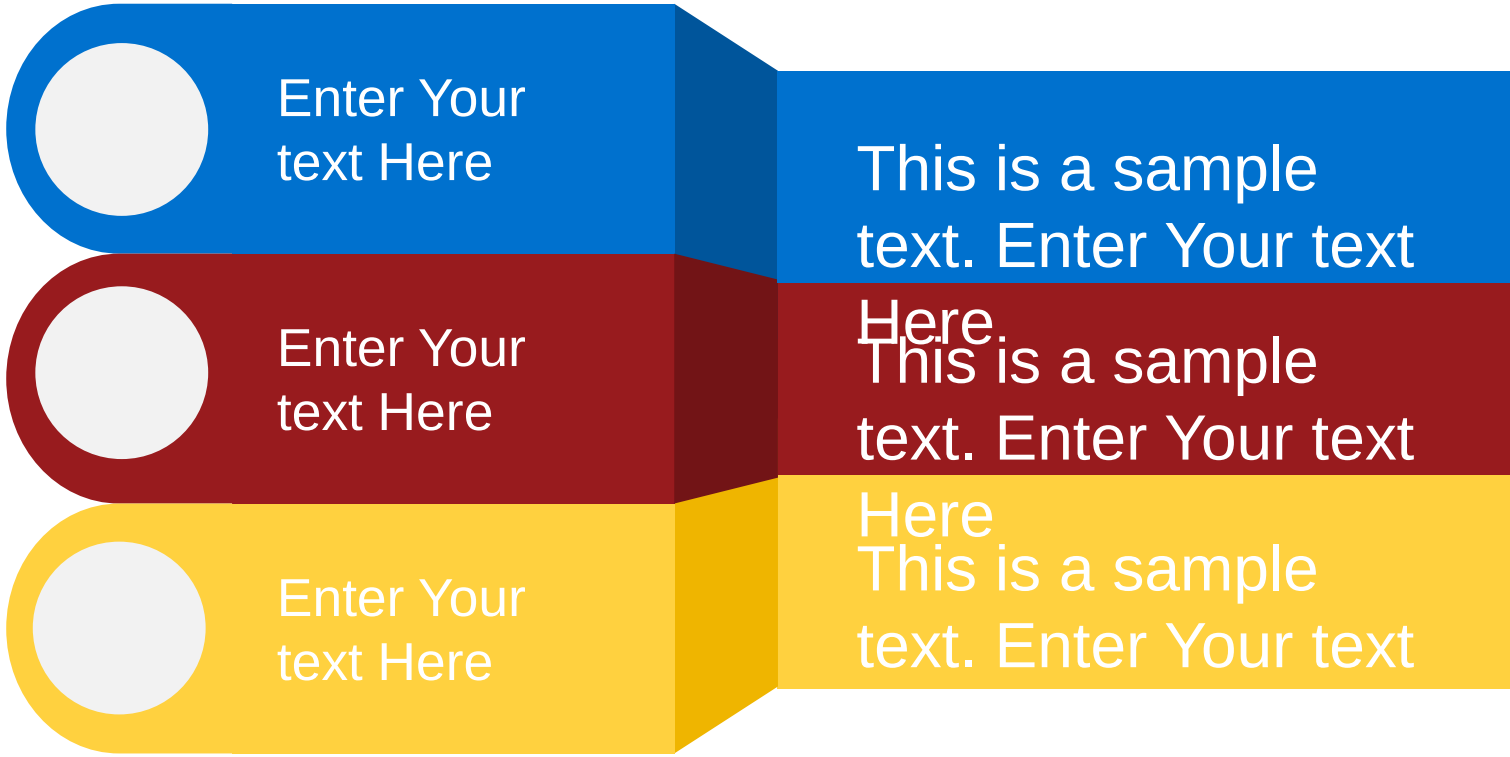
This is a sample text. You can replace this text. Enter Your text Here

This is a sample text.
Enter Your text Here

This is a sample text. You can replace this text. Enter Your text Here

This is a sample text. You can replace this text. Enter your text here.

1. This is a sample text. Enter Your text Here
2. This is a sample text. Enter Your text Here
3. This is a sample text. Enter Your text Here



Sample text

This is a sample text.
Insert your desired text here.

Sample text

This is a sample text.
Insert your desired text here.

Sample text

This is a sample text.
Insert your desired text here.

Sample text

This is a sample text.
Insert your desired text here.

Sample text

This is a sample text.
Insert your desired text here.

Sample text

This is a sample text.
Insert your desired text here.

Sample text

This is a sample text.
Insert your desired text here.

Sample text

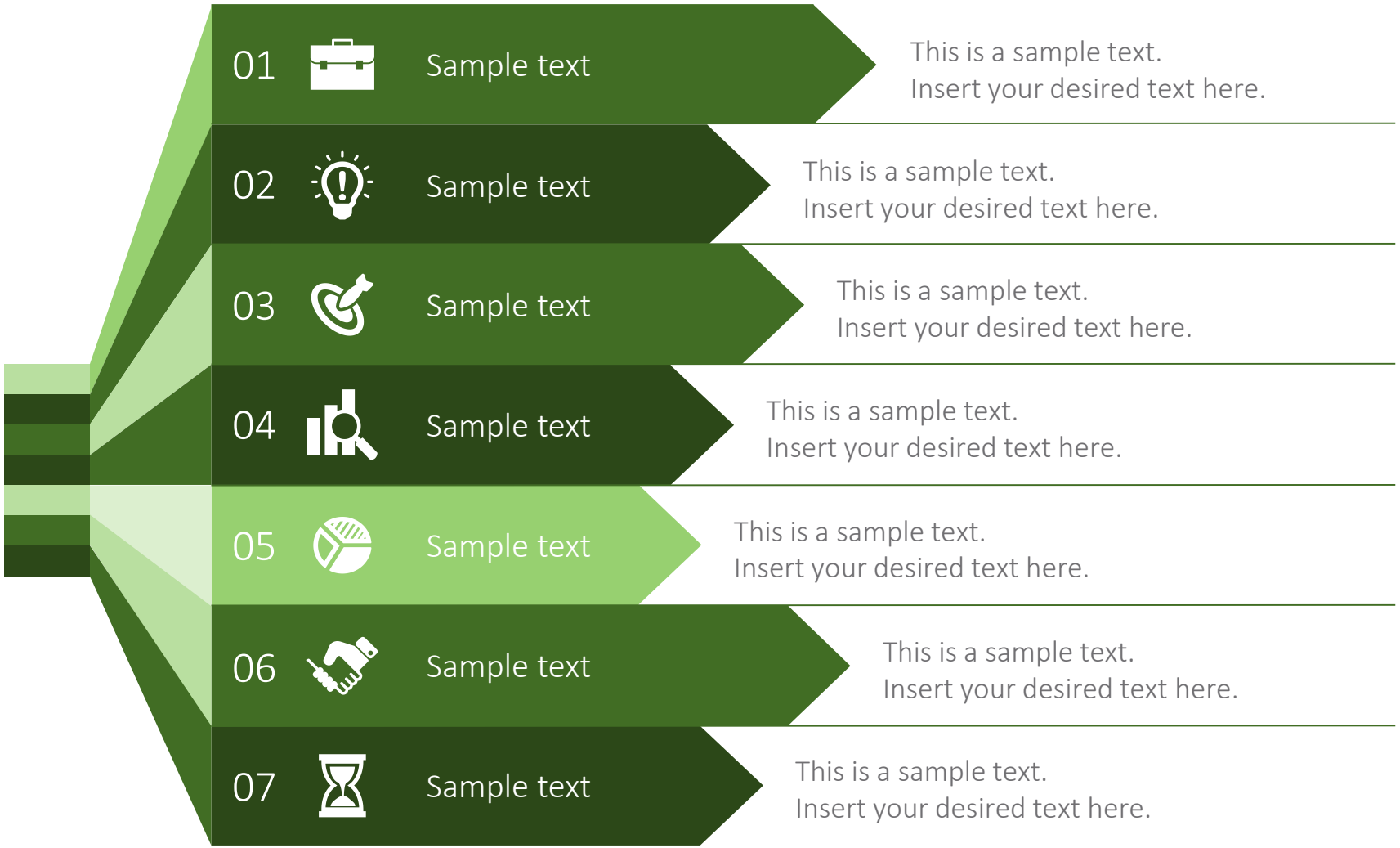
This is a sample text.
Insert your desired text here.

Sample text

This is a sample text.
Insert your desired text here.

Sample text

This is a sample text.
Insert your desired
text here.



Sample text

This is a sample text.
Insert your desired text here.

Sample text

This is a sample text.
Insert your desired text here.

Sample text

This is a sample text.
Insert your desired text here.

Sample text

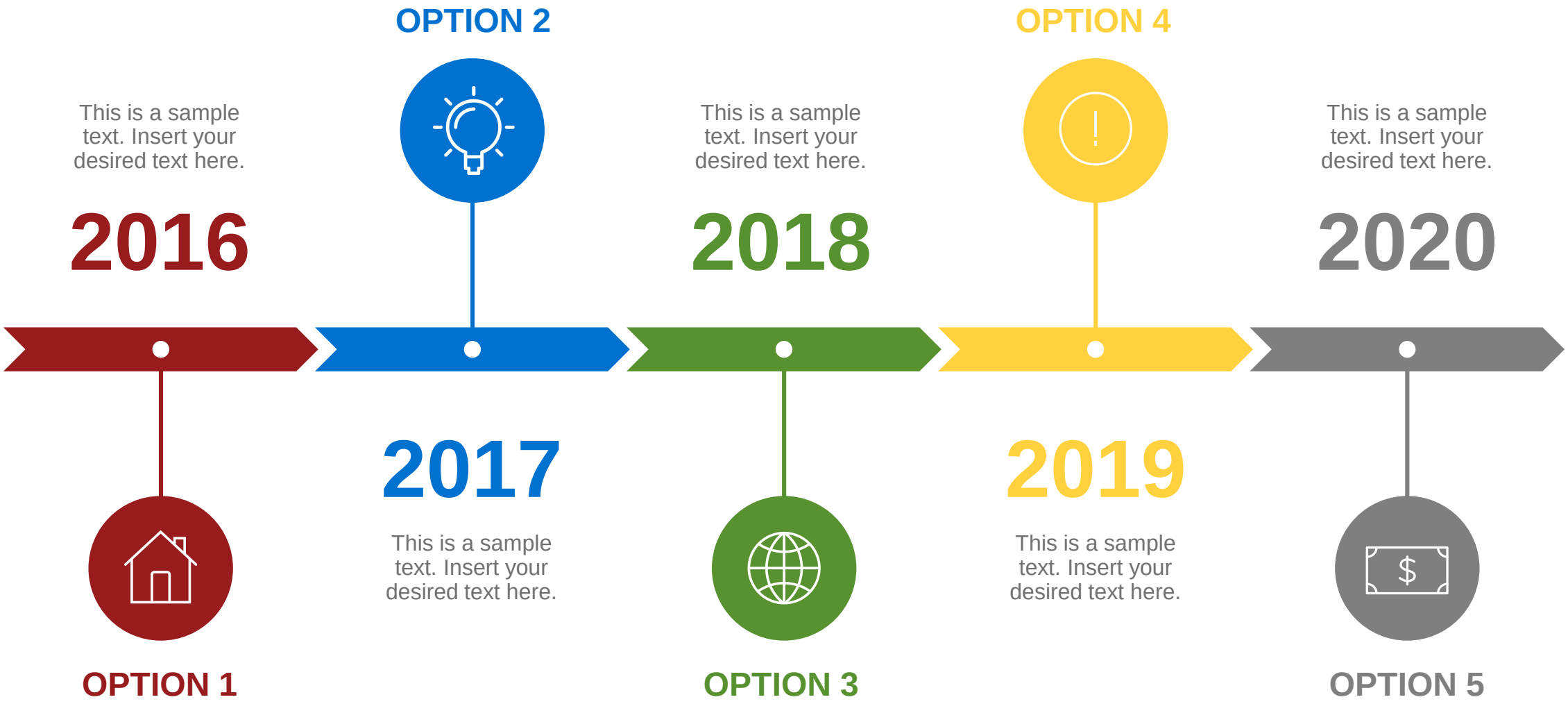
This is a sample text.
Insert your desired text here.

Sample text

This is a sample text.
Insert your desired text here.

Sample text

This is a sample text.
Insert your desired text here.


OPTION 2
OPTION 4
This is a sample text. Insert your desired text here.
2016
This is a sample text. Insert your desired text here.
2018
This is a sample text. Insert your desired text here.
2020
2017
This is a sample text. Insert your desired text here.
2019
This is a sample text. Insert your desired text here.
OPTION 1
OPTION 3
OPTION 5