https://vsr.informatik.tu-chemnitz.de/projects/2019/growth

TECHNISCHE UNIVERSITÄT CHEMNITZ

Helsinki, FINLAND
ICWE 2020

20TH INTERNATIONAL CONFERENCE ON WEB ENGINEERING

VISH: Does Your Smart Home Dialogue System Also Need Training Data?

Mahda Noura, Sebastian Heil, and Martin Gaedke

*VSR.Informatik.*TU-Chemnitz.*de*

Professorship for Distributed and Self-Organizing Systems
Chemnitz University of Technology, Germany

Mahda Noura
Research Associate

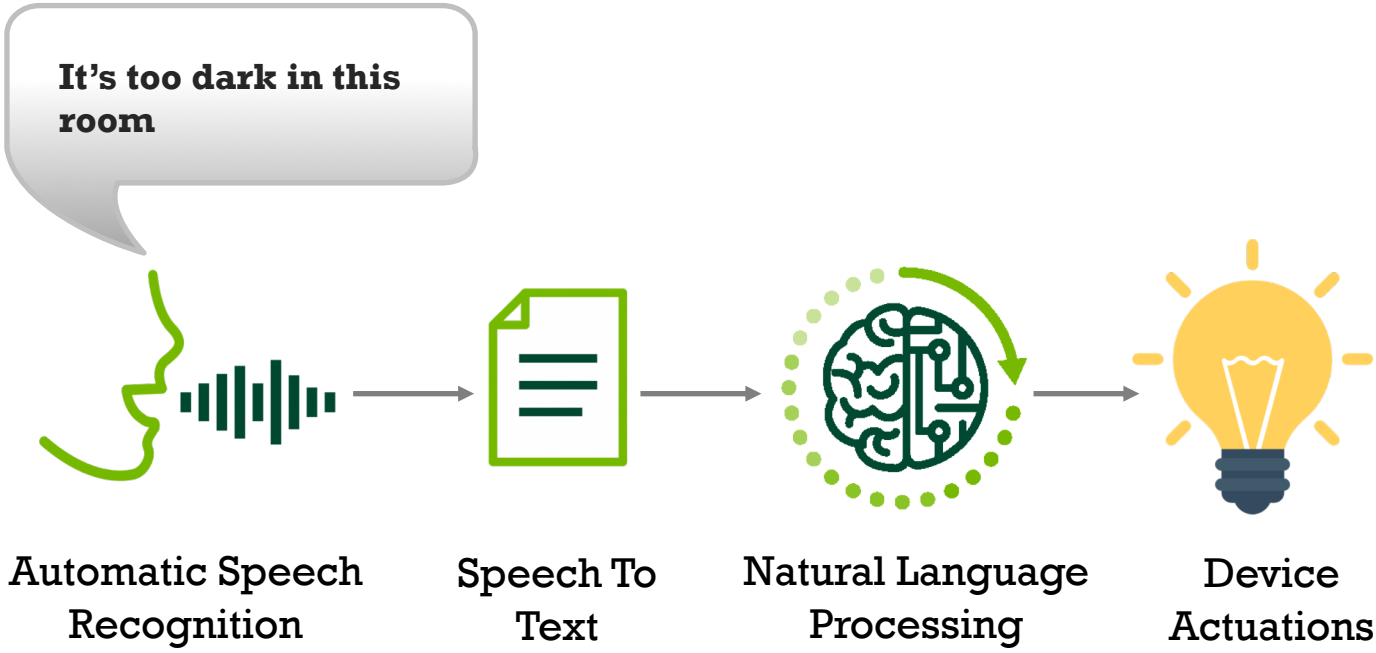
Sebastian Heil
Research Associate

Martin Gaedke
Full Professor

1 Introduction

https://www.o-ton.at/

Natural Language Understanding in Smart Home



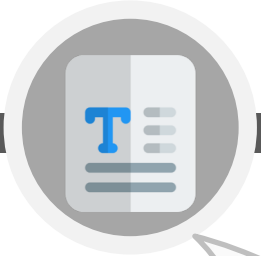
Research Challenge

To ensure understanding of the user goal and provide good user experience, a **high-quality** and **sufficiently** large training corpus is required

Common Process

Record sample utterances from users

Transcribe utterances to text



Create labelled training dataset

Train NLU model

Approach

Goal-oriented WoT Interfaces

GWoT is defined as $\mathbf{GWoT} = \langle S, s^0, A, s^*, \gamma \rangle$

S

S = $\{s_1, s_2, \ldots, s_n\}$ includes s^0 and s^*

s_i represented by $Par_s = \{p_1, p_2, \ldots, p_k\}$

A

$a \in A$ is $a = \langle effects(a), actuation(a) \rangle$

$effects(a) = \langle p_i, t, v \rangle$

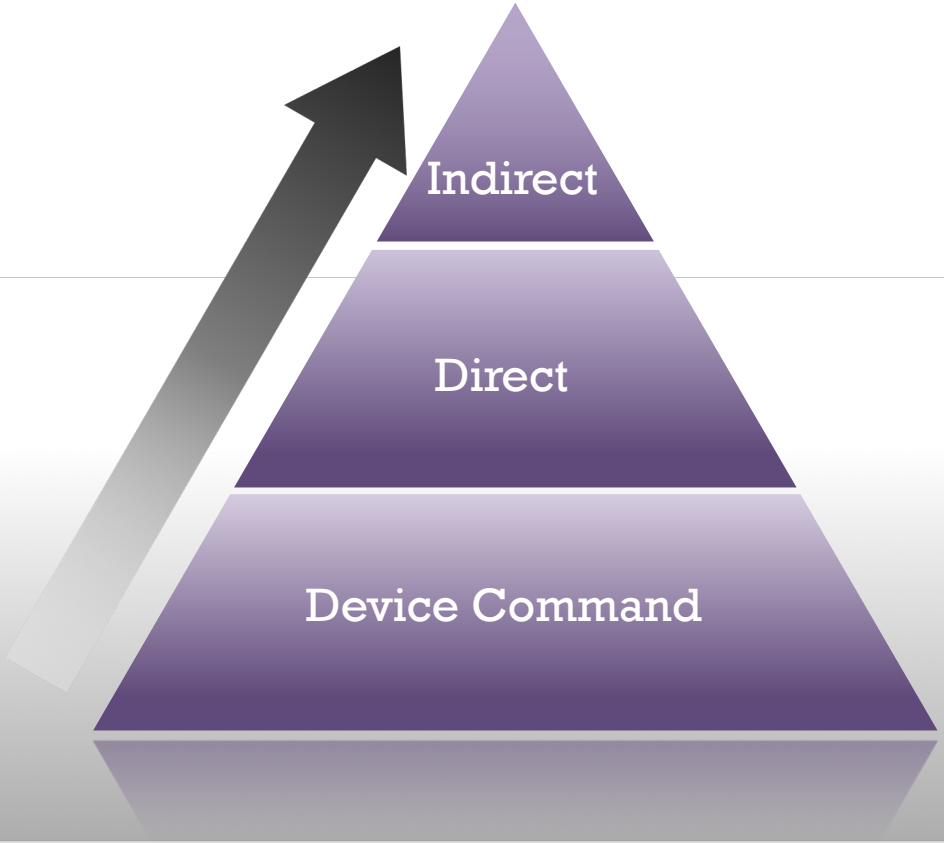
$effects(a) \in \{increase, decrease, assign, toggle\}$

$actuation(a)$: set of HTTP / MQTT requests

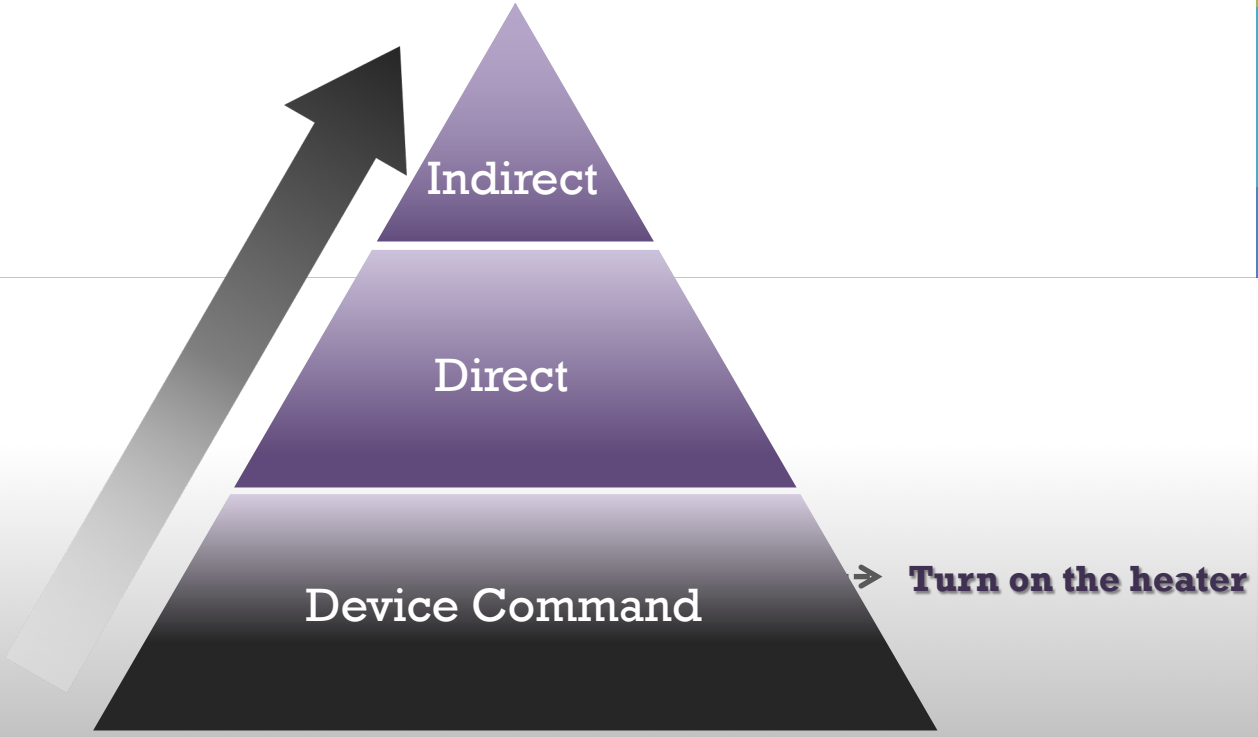
γ

$\gamma: S \times A \longrightarrow S$

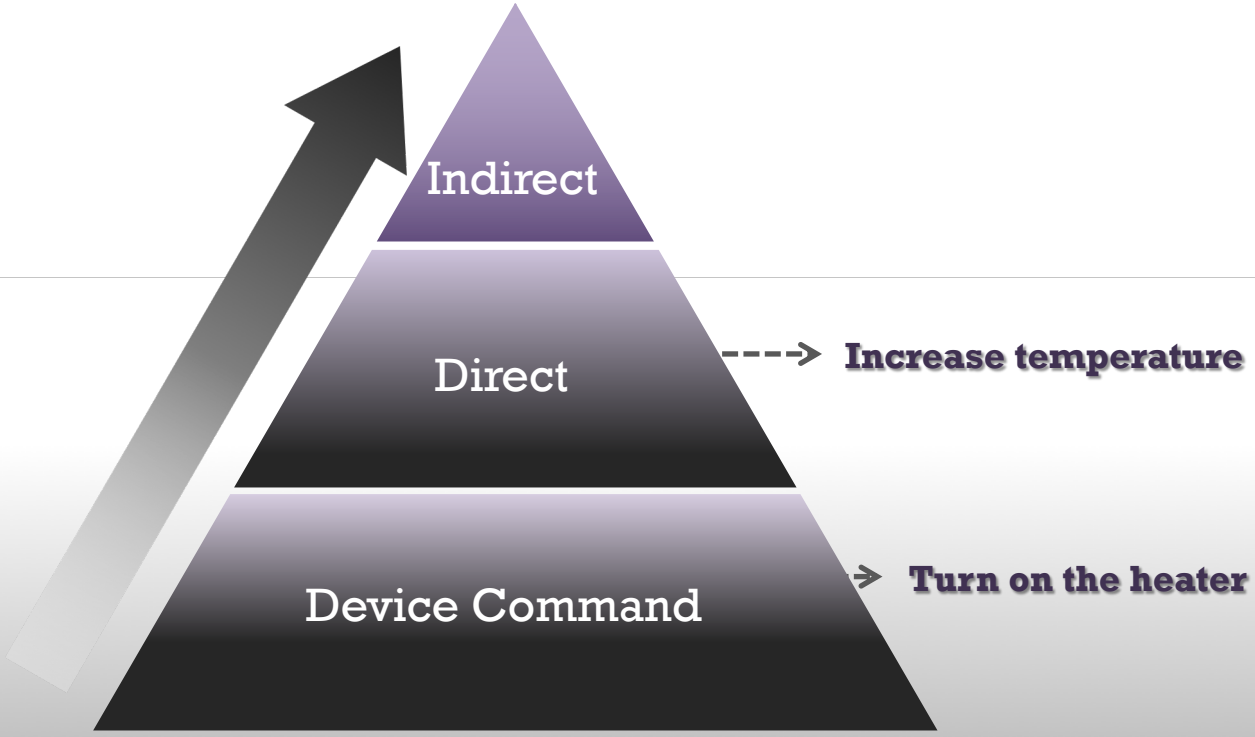
User Goals in Smart Home



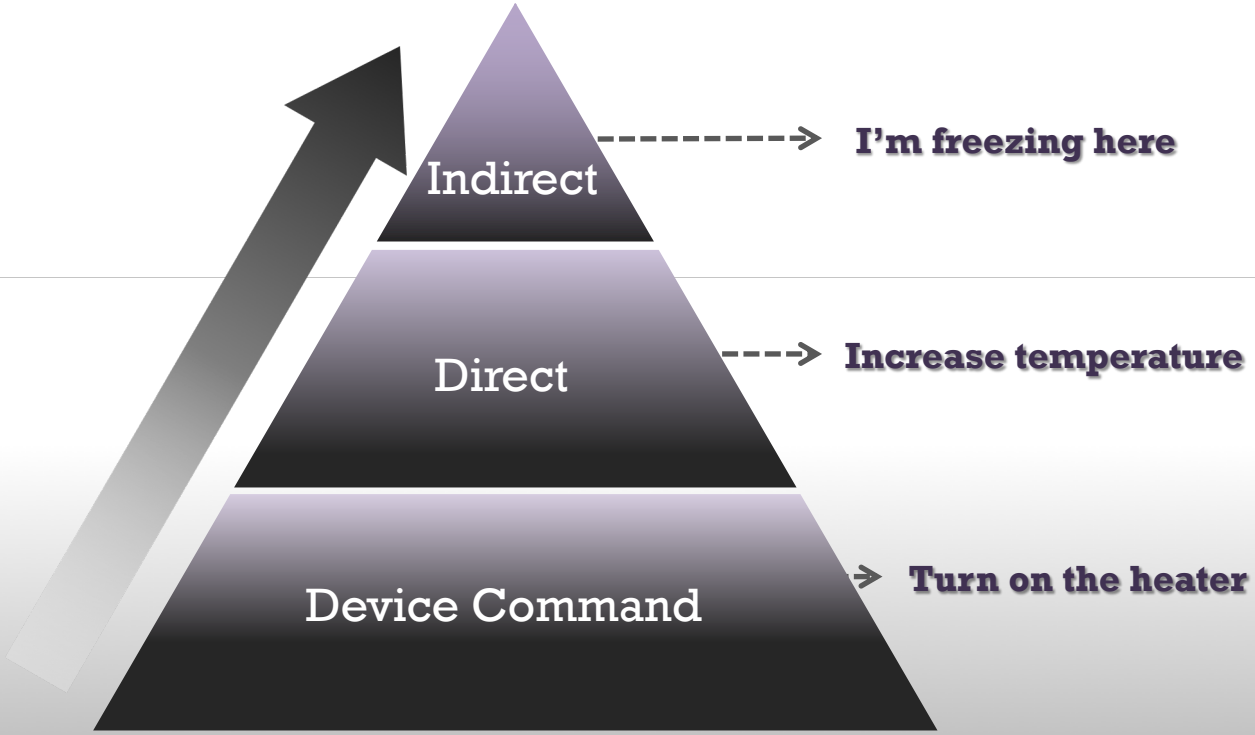
User Goals in Smart Home



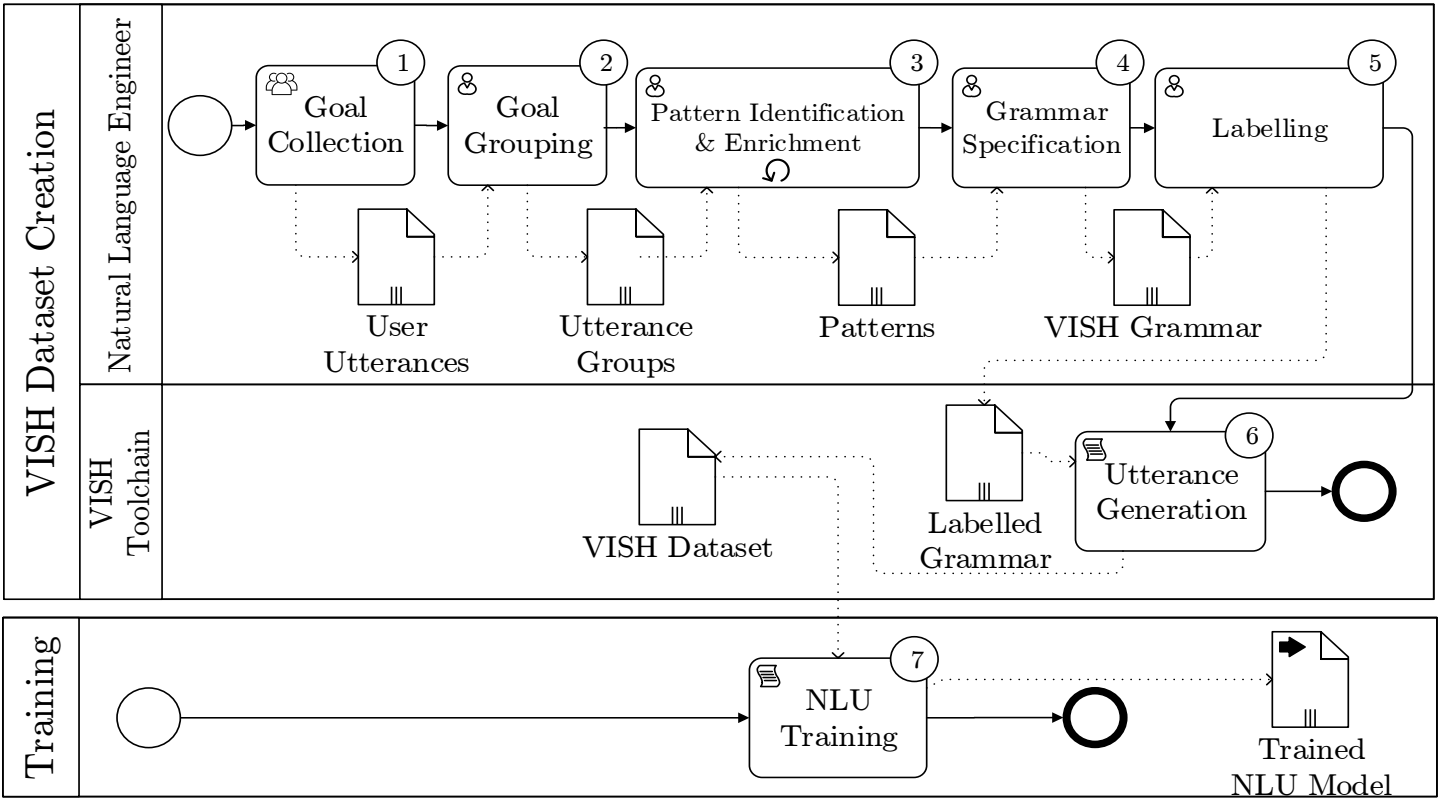
User Goals in Smart Home



User Goals in Smart Home

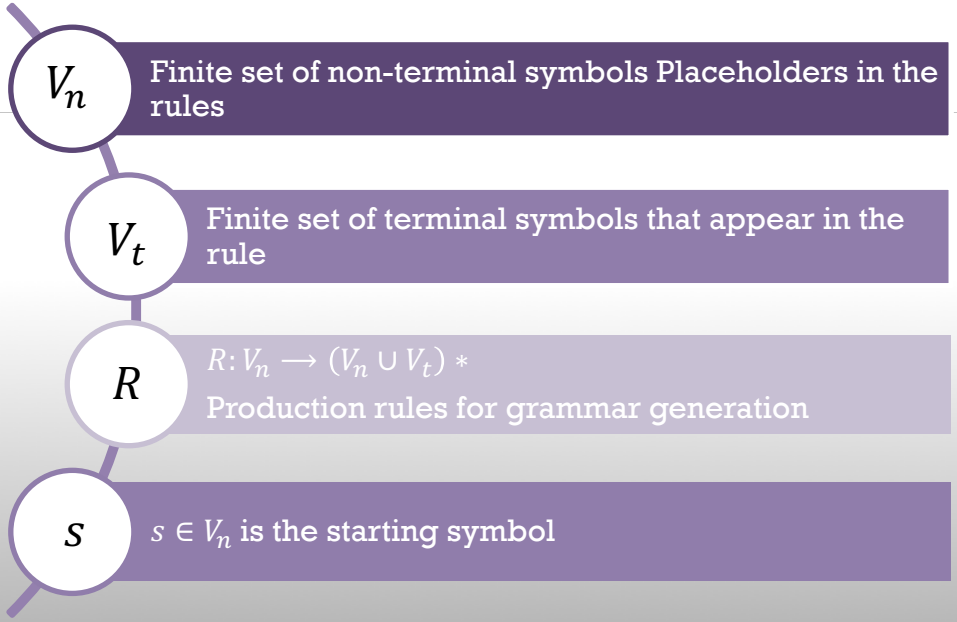


VISH: Voice Interface for Smart Home

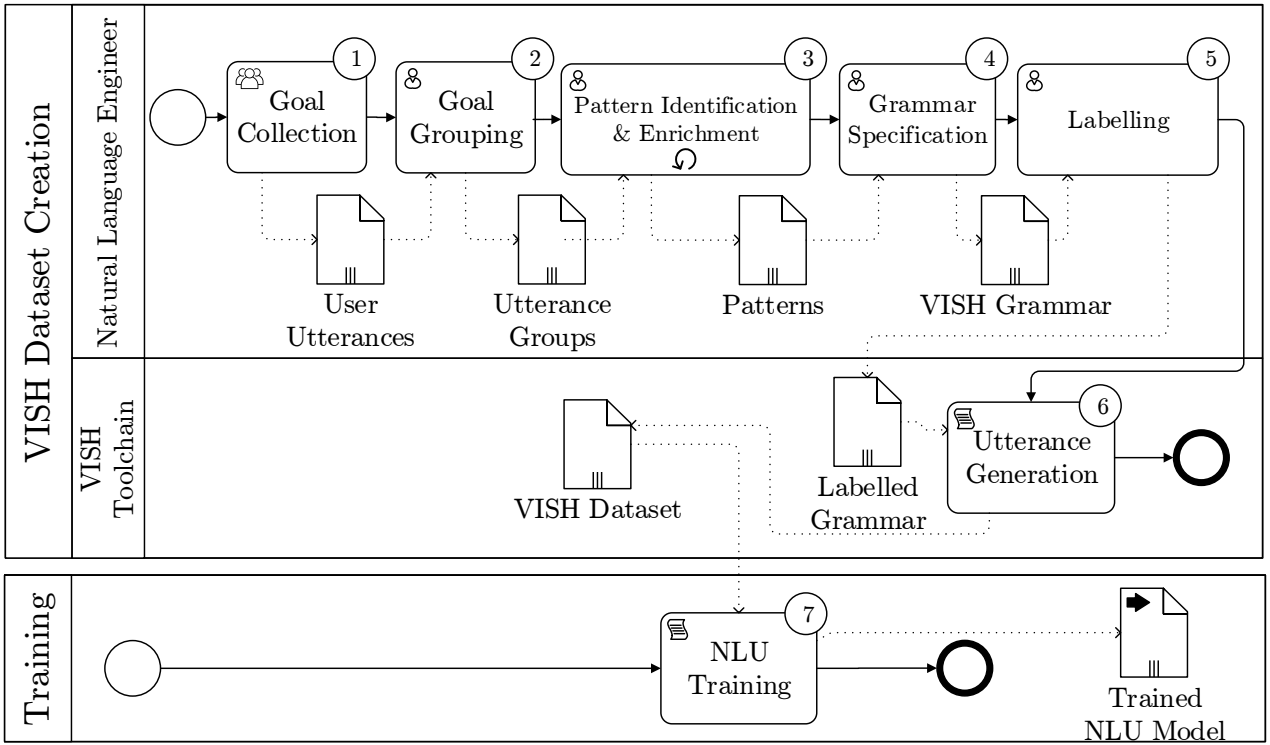


VISH Grammar Specification

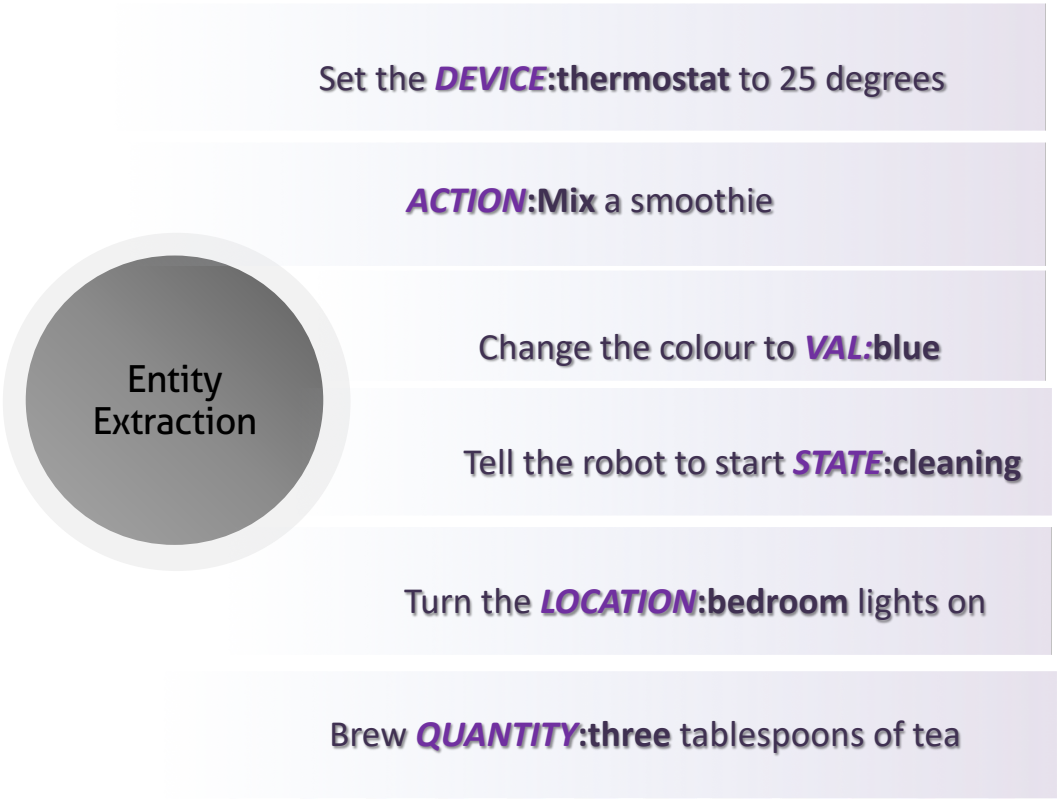
VISH Grammar G is defined as $G = \langle V_n, V_t, R, s \rangle$



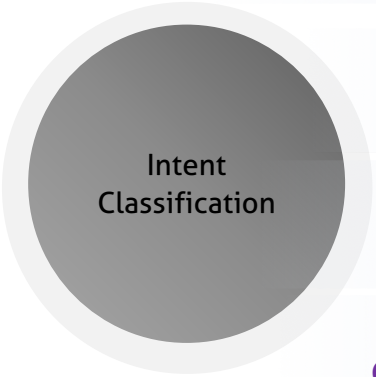
VISH: Voice Interface for Smart Home



Labelling for Extraction



Labelling for Classification



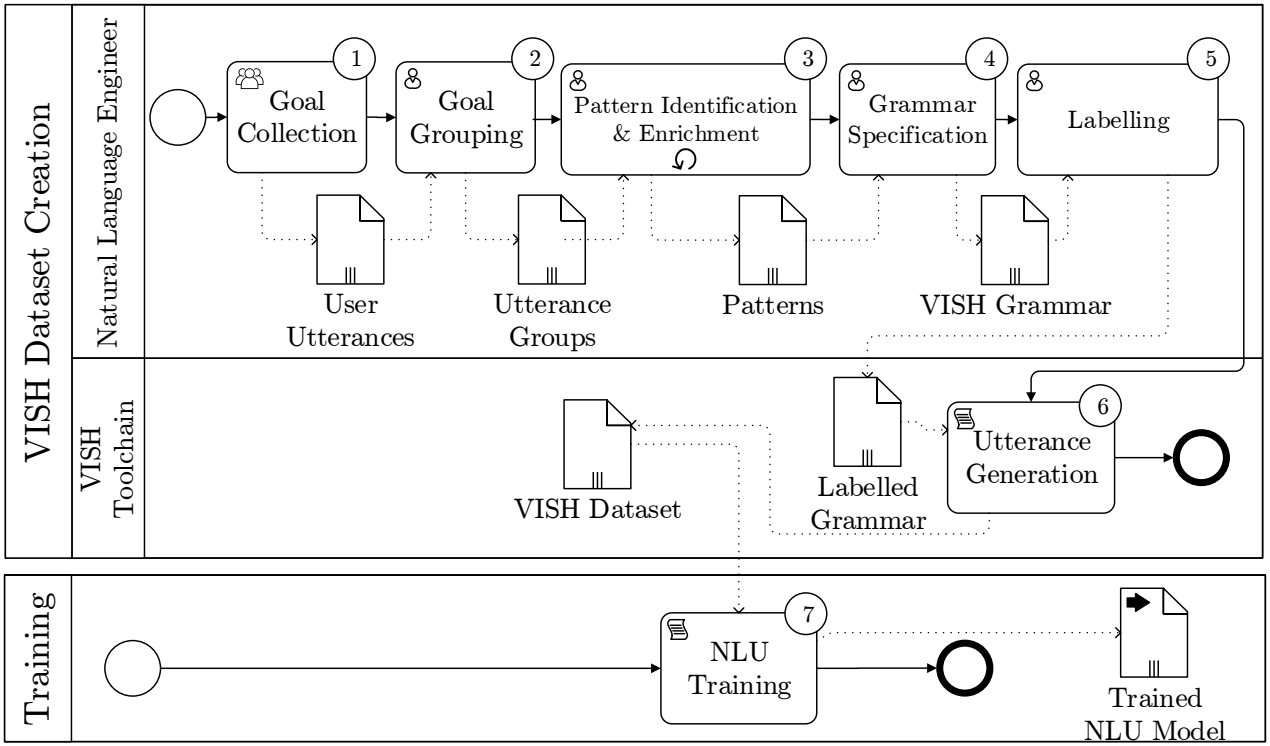
Physical Parameter : temperature, humidity, brightness

Effect Type : increase, decrease, toggle, assign

Scope : global, device

Goal Type : device command, direct, indirect

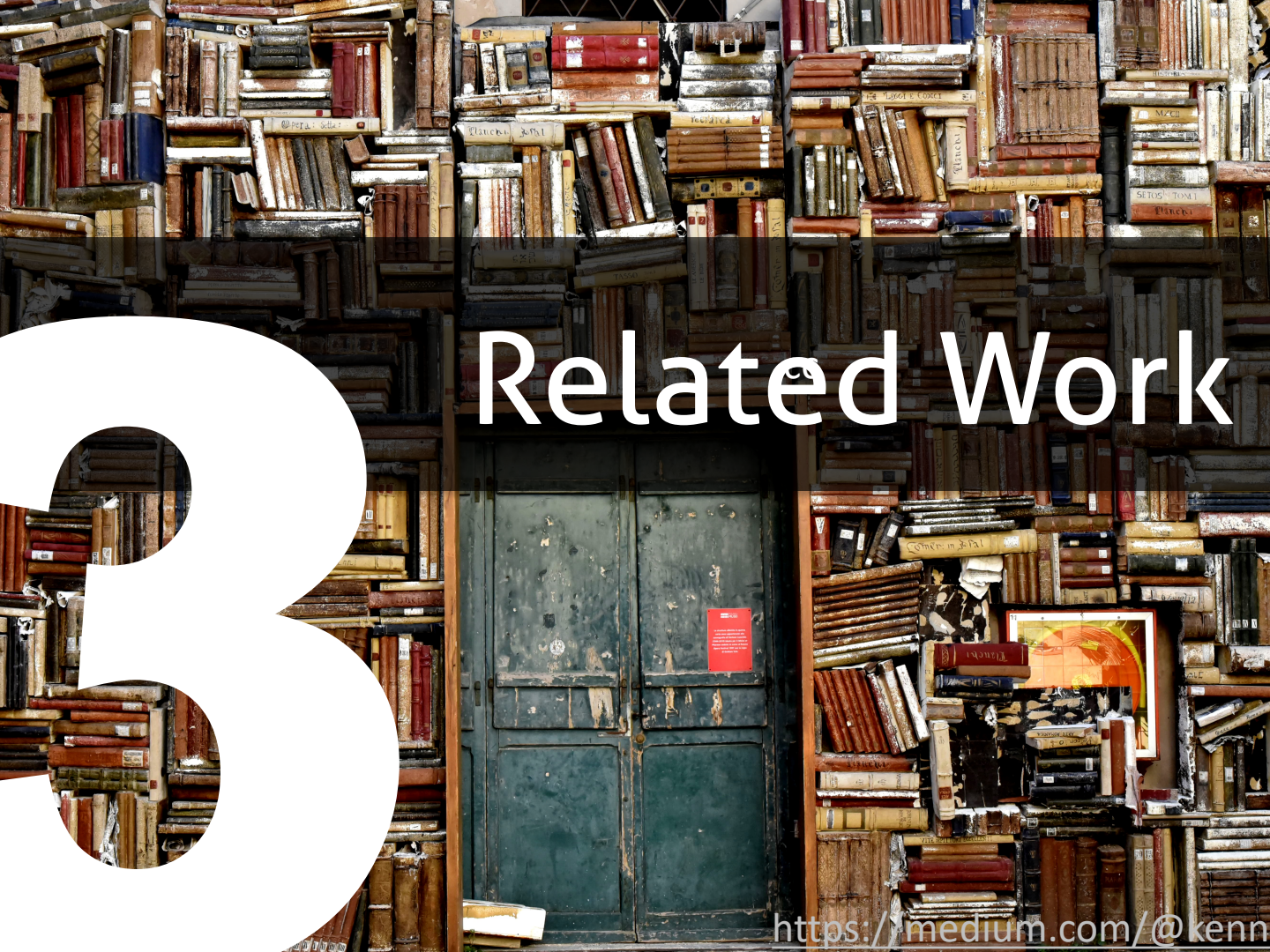
VISH: Voice Interface for Smart Home



VISH Dataset

https://vsr.informatik.tu-chemnitz.de/projects/2019/growth/

Total Words: 33,618,437	**Entities:** 6,819,062	**Effect Types:** 4
Utterances: 4,743,745	**Indirect Goals:** 4,00,854	**Device Types:** 305
Unique Words: 822	**Direct Goals:** 742,891	**Action Types:** 130
Entity Types: 7	**Total Actions:** 674,711	**Parameter Types:** 8


3
Related Work
https://medium.com/@kenn

Dataset	Size	Lan	Direct Goal	Indirect Goal	Availability	Pr	E	De	Ac	Approach
Google Home	NA	en	✓	✗	✗	✗	✗	✓	✓	User
Alexa	20K	en	✓	✗	✗	✗	✗	✓	✓	User
IFTTT	NA	en	◐	✗	✓	✗	✗	✓	✓	User, ECA
DIRHA	183	en	✗	✗	◐	✗	✗	✗	✗	Experiment
VocADom@4H	7K	fr	✗	✗	✗	✗	✗	✗	✗	Synthetic & Experiment
Voice-Home2	1,5K	fr	✓	✗	✓	✗	✗	✗	✗	Experiment
Sweet-Home	1,5K	fr	◐	✗	◐	✗	✗	✗	✗	Experiment
ATHENA	370	el	◐	✗	✗	✗	✗	✗	✗	Experiment
Genie	3M	en	✓	✗	✓	✗	✗	✓	✓	Synthetic & Crowdsourcing
VISH	5M	en	✓	✓	✓	✓	✓	✓	✓	User-derived generative grammar

4 Evaluation

Evaluation Procedure

Dataset preparation : Rasa markdown syntax

Dataset division: 80% training, 20% holdout

Dataset training on NLU models: supervised embedding, pretrained embeddings convert, pretrained embedding spacy

Testing the NLU models: accuracy, precision, F1-score

Evaluation Results

	Supervised 100	Supervised 300	Supervised 500	ConveRT	Mean
Precision	.9834	.9849	.9886	.9756	.9831
Recall	.9834	.9849	.9886	.9756	.9831
F_1	.9834	.9849	.9886	.9756	.9831

Suitability of VISH dataset for training state-of-the-art NLU intent classification for smart home utterances

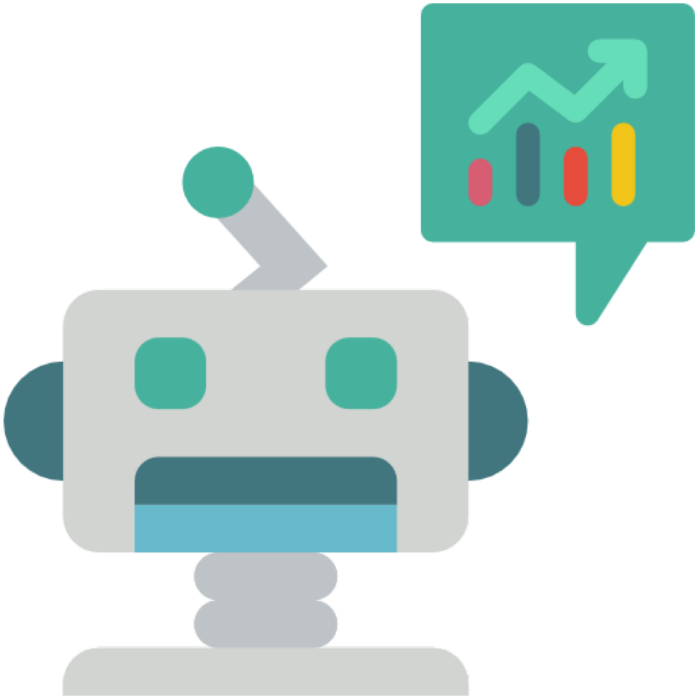
Evaluation Results

		State	Quantity	Value	Location	Device	Unit	Action	Mean
Custom	P	1.0	1.0	.9980	1.0	.9999	1.0	.9998	.9997
	R	1.0	1.0	.9999	1.0	.9998	1.0	.9995	.9999
	F_1	1.0	1.0	.9989	1.0	.9998	1.0	.9997	.9998
SpaCy	P	1.0	1.0	.9992	1.0	.9999	1.0	.9998	.9998
	R	1.0	1.0	.9983	1.0	.9998	1.0	.9995	.9997
	F_1	1.0	1.0	.9988	1.0	.9998	1.0	.9996	.9998

Suitability of VISH dataset for training state-of-the-art NLU entity extraction for smart home utterances


5
Conclusion

Conclusion & Future Work



VISH dataset can be used to train goal-oriented dialogue systems for smart homes with **high accuracy**

Integrate custom NLU pipeline trained on the VISH dataset

Evaluate the dataset with end-user goals in live IoT settings

VSR

Thank You!

mahda.noura@informatik.tu-chemnitz.de

VSR.*Informatik*.TU-Chemnitz.*de*

@mahdanoura /mahdanoura