Have We achieved Personalized Dialogue Generation yet?

S. **Mahed** Mousavi - Ph.D. Candidate
Signals and Interactive Systems Lab

UNIVERSITY OF TRENTO - Italy
Department of Information Engineering and Computer Science
SiS LAB
Signals & Interactive Systems

What is Natural Language Generation?

- What is Natural Language Generation?
  - Data-2-Text

What is Natural Language Generation?

- What is Natural Language Generation?
  - Data-2-Text

- Image Captioning

Input:

Output:

```
<start> a close up of a person
wearing a bow tie <end>
```

What is Natural Language Generation?

- What is Natural Language Generation?
  - Data-2-Text

- Table to Text

Input:

Denise Margaret Scott	
Born	24 April 1955 Melbourne, Victoria
Nationality	Australian
Other names	Scotty
Occupation	Comedian, actor, television and radio presenter

Output:

Denise Margaret Scott (born 24 April 1955) is an Australian comedian, actor and television presenter.

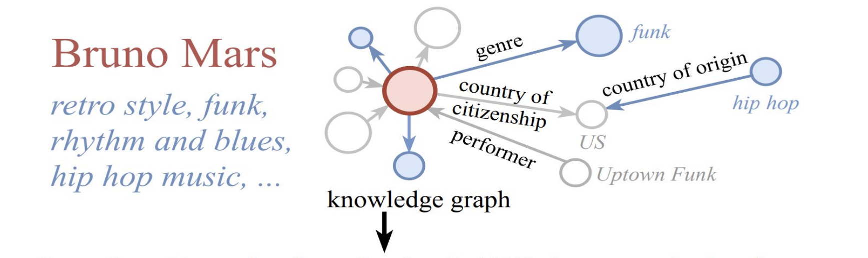
UNIVERSITY OF TRENTO - Italy
Department of Information Engineering and Computer Science
SiS LAB Signals & Interactive Systems

What is Natural Language Generation?

- What is Natural Language Generation?
    - Data-2-Text

- Graph to Text

Input:



Output:

Peter Gene Hernandez (born October 8, 1985), known professionally as Bruno Mars, is an American singer, songwriter, multi-instrumentalist, record producer, and dancer. He is known for his stage performances, *retro* showmanship and for performing in a wide range of musical styles, including *R&B*, *funk*, *pop*, *soul*, *reggae*, *hip hop*, and *rock*.

What is Natural Language Generation?

- What is Natural Language Generation?
  - Data-2-Text
  - Text-2-Text
    - Generating a lexicalized human-readable response based on a textual context
    - Traditionally: Converting Mean Representation (a non-linguistic intermediate representation) to a lexicalized output
    - Properties of the Generation output:
      - Correctness
      - Appropriateness
      - Coherent
      - Engaging

What is Natural Language Generation?

- What is Natural Language Generation?
  - Data-2-Text
  - Text-2-Text
    - By Domain:
      - Summarization
        - Abstractive vs. Extractive

Extractive Summarization

Source Text: Peter and Elizabeth took a taxi to attend the night party in the city. While in the party, Elizabeth collapsed and was rushed to the hospital.

Summary: Peter

Abstractive Summarization

Source Text: Peter and Elizabeth took a taxi to attend the night party in the city. While in the party, Elizabeth collapsed and was rushed to the hospital.

Summary: Elizabeth was hospitalized after attending a party with Peter.

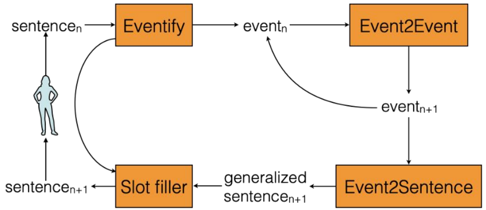
What is Natural Language Generation?

- What is Natural Language Generation?
  - Data-2-Text
  - Text-2-Text
    - By Domain:
      - Story Generation

Context: **Tom and Sheryl have been together for two years. One day, they went to a carnival together. He won her several stuffed bears, and bought her cakes. When they reached the Ferris wheel, he got down on one knee**

Prediction: *Tom asked Sheryl to marry him.*

- Generation

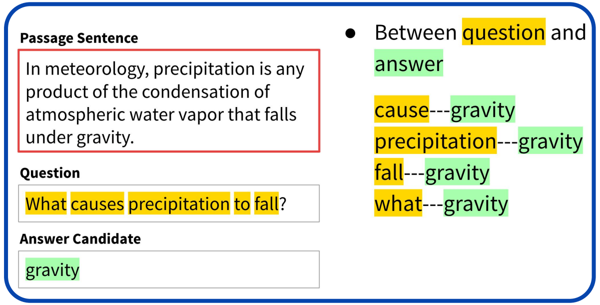


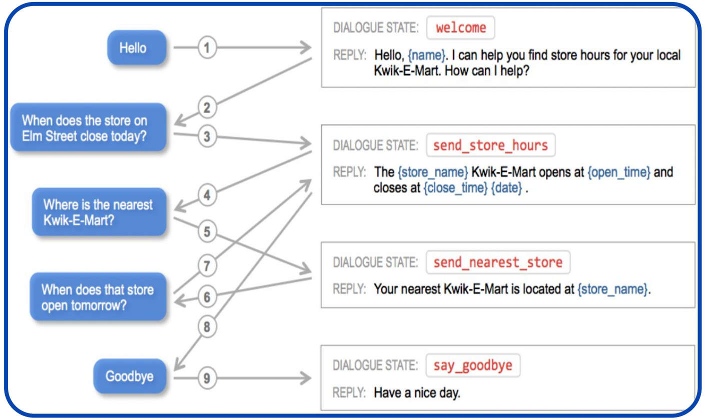
- Selection

Right Ending	Wrong Ending
Tom asked Sheryl to marry him.	He wiped mud off of his boot.

What is Natural Language Generation?

- What is Natural Language Generation?
  - Data-2-Text
  - Text-2-Text
    - By Domain:
      - Response Generation
        - Task-Based
        - Open-Domain
        - Q & A



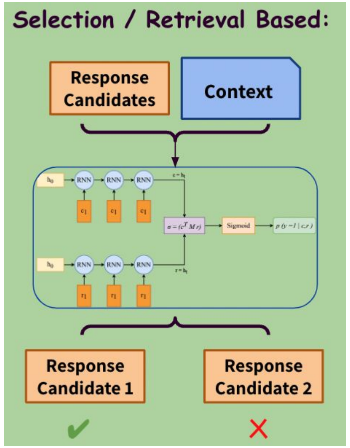


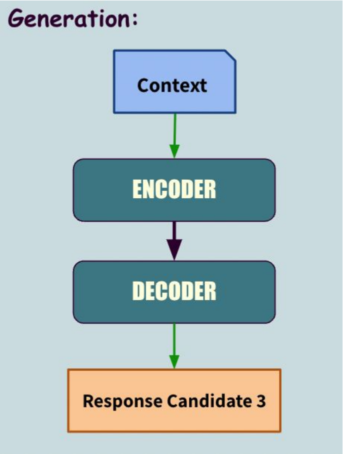
What is Natural Language Generation?

- What is Natural Language Generation?
  - Data-2-Text
  - Text-2-Text
    - By Domain:
      - Summarization
        - Abstractive vs. Extractive
      - Story Generation
      - Response Generation
        - Task-Based vs. Open-Domain vs. Q & A
      - Translation
    - By Policy:
      - Template-based vs. Selection vs. Generation

Generative Vs. Retrieval?

- What is Natural Language Generation?
  - Data-2-Text
  - Text-2-Text
    - By Policy:
      - Template-based vs. Selection vs. Generation





What is Personal Dialogue Generation?

- Personal dialogues are conversations which include user-specific recollections of events, objects, entities and their relations.
    - They may also encompass personal feelings, thoughts, and emotions.

- Personal dialogue generation requires to:
    1) Hold dialogues about personal events, relationships, and participants
    2) carrying out **multi-session conversations** over extended period of time.
    3) obtain the **required knowledge** during each interaction with the user and from her Personal Narratives.

What is Personal Dialogue Generation?

- Personal dialogues are conversations which include user-specific recollections of events, objects, entities and their relations.
  - They may also encompass personal feelings, thoughts, and emotions.

- Personal dialogue generation requires to:
  1) Hold dialogues about personal events, relationships, and participants
  2) carrying out **multi-session conversations** over extended period of time.
  3) obtain the **required knowledge** during each interaction with the user and from her Personal Narratives.

We need Dialogue Data!

- The acquisition of a dialogue corpus is a key step in the process of training a dialogue model.

- Corpora acquisitions have been designed either for open-domain information retrieval about a finite set of topics (e.g. news, music, weather, games etc.) or slot-filling tasks (e.g. restaurant booking).

- However, neither of the above approaches can address the need for personal conversations which include user-specific recollections of events, objects, entities and their relations.

We need Dialogue Data!

- The acquisition of a dialogue corpus is a key step in the process of training a dialogue model.
- Corpora acquisitions have been designed either for open-domain information retrieval about a finite set of topics (e.g. news, music, weather, games etc.) or slot-filling tasks (e.g. restaurant booking).
- However, neither of the above approaches can address the need for personal conversations which include user-specific recollections of events, objects, entities and their relations.

We need Dialogue Data!

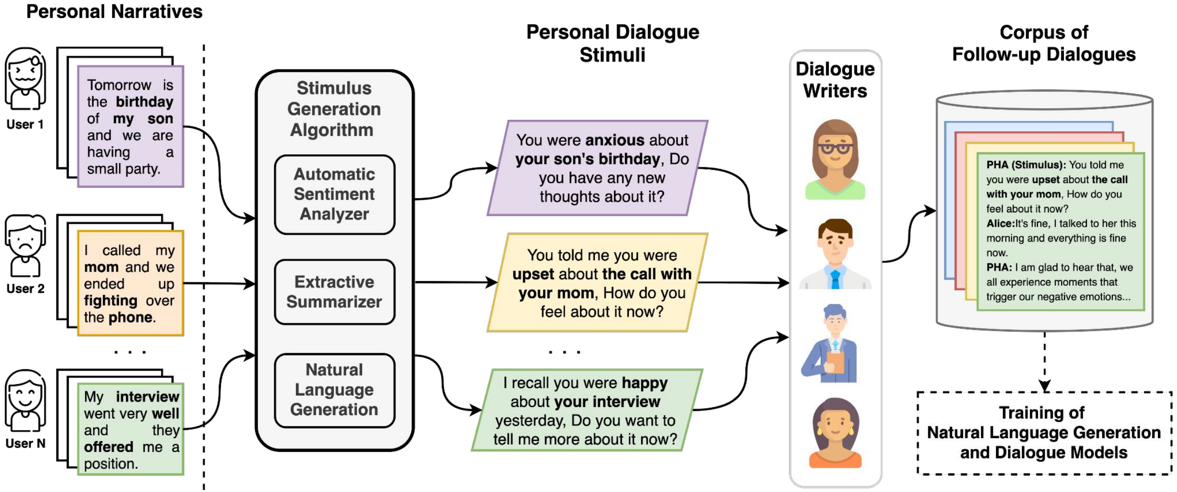
- The acquisition of a dialogue corpus is a key step in the process of training a dialogue model.

- Corpora acquisitions have been designed either for open-domain information retrieval about a finite set of topics (e.g. news, music, weather, games etc.) or slot-filling tasks (e.g. restaurant booking).

- However, neither of the above approaches can address the need for personal conversations which include user-specific recollections of events, objects, entities and their relations.

We Collect Dialogue Data!

- We proposed a novel methodology to collect corpora of personalized follow-up dialogues.[1]
    - Dialogue follow-ups are a critical resource for learning about the life events of the narrator as well as his/her corresponding thoughts and emotions in a timely manner.



UNIVERSITY OF TRENTO - Italy
Department of Information Engineering and Computer Science
SiS LAB
Signals & Interactive Systems

[1] Mousavi, S. M., Cervone, A., Danieli, M., & Riccardi, G. (2021, June). ***Would you like to tell me more? Generating a corpus of psychotherapy dialogues.*** In Proceedings of the Second Workshop on Natural Language Processing for Medical Conversations (pp. 1-9).

We need Personal Narrative Understanding!

- We need to obtain the **knowledge** required for responding during each interaction with the user

- A Personal Narrative has a complex structure.
  - It is different than Intent(or Dialogue Act) classification and slot filling.
  - We need to extract personal events, emotions, and participants.

- We need to construct the **Personal Space** of the user from her narratives

We need Personal Narrative Understanding!

- We need to obtain the **knowledge** required for responding during each interaction with the user

- A Personal Narrative has a complex structure.
    - It is different than Intent(or Dialogue Act) classification and slot filling.
    - We need to extract personal events, emotions, and participants.

- We need to construct the **Personal Space** of the user from her narratives

We need Personal Narrative Understanding!

- We need to obtain the **knowledge** required for responding during each interaction with the user

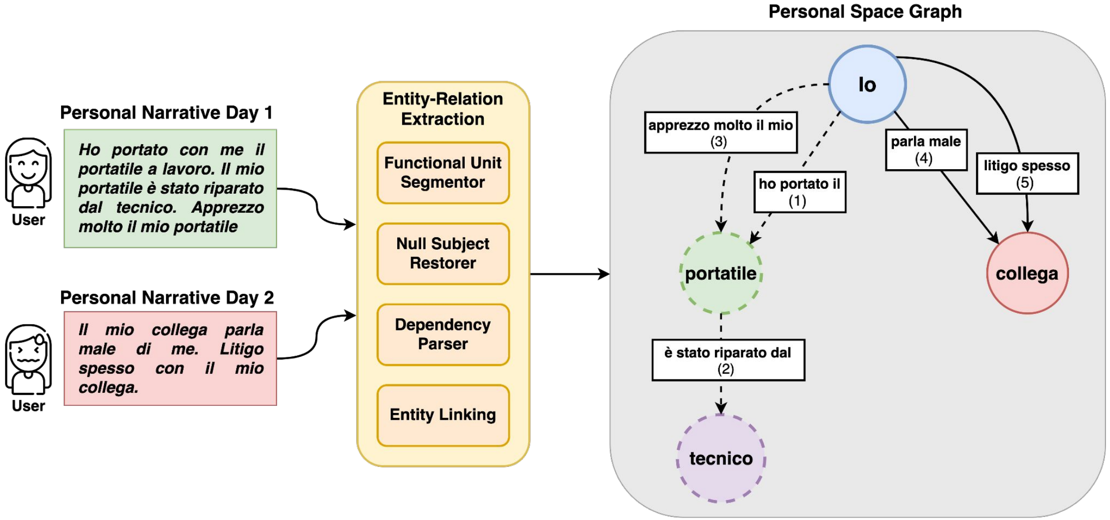
- A Personal Narrative has a complex structure.
  - It is different than Intent(or Dialogue Act) classification and slot filling.
  - We need to extract personal events, emotions, and participants.

- We need to construct the **Personal Space** of the user from her narratives

We develop an Event-Relation Extractor!

- We develop an unsupervised approach to automatically extract personal life-events and participants from the users' narratives and represent them as a personal graph.[2]

- This personal graph is then updated at each interaction with the patient.



[2] Mousavi, S. M., Negro, R., & Riccardi, G. (2021). ***An Unsupervised Approach to Extract Life-Events from Personal Narratives in the Mental Health Domain***. Eighth Italian Conference on Computational Linguistics.

Another Perspective: Emotion Carriers!

- Emotion Carriers (EC) are defined as the persons, objects or actions that explain the emotion felt by the narrator, after recollecting the event[3].

- EC tokens clearly convey the activation of the emotional state in the narrator, even though they may not explicitly manifest a sentiment.

- In contrast, emotion-laden words (such as happy, sad, enjoyed, and overwhelmed), explicitly express certain sentiment polarity.

Emotion-laden word *Emotion Carrier*
FU1: I experienced a bit of **distress** **in the office**

Emotion Carrier *Emotion-laden word*
FU2: because **talking with colleagues** *makes me* **anxious**!

Example: *Emotion-laden words manifest a sentiment explicitly while Emotion Carriers describe the events, persons or objects conveying emotions.*

UNIVERSITY OF TRENTO - Italy
Department of Information Engineering and Computer Science
SiS LAB Signals & Interactive Systems

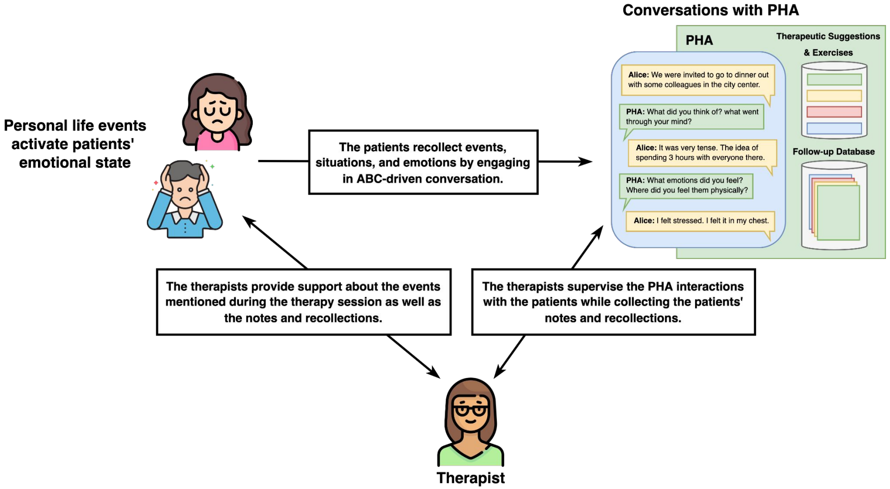
[3] Mousavi, S. M., et al. **"Can Emotion Carriers Explain Automatic Sentiment Prediction? A Study on Personal Narratives."** Proceedings of the 12th Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis. 2022.

Lets Try Retrieval Response Models!

- We developed a Retrieval Based PHA for the mental health domain.[4,5]

- The results show a significant positive trend in the reduction of symptoms related to distress, obsessivity, and compulsivity by the patient group that received support from PHA.

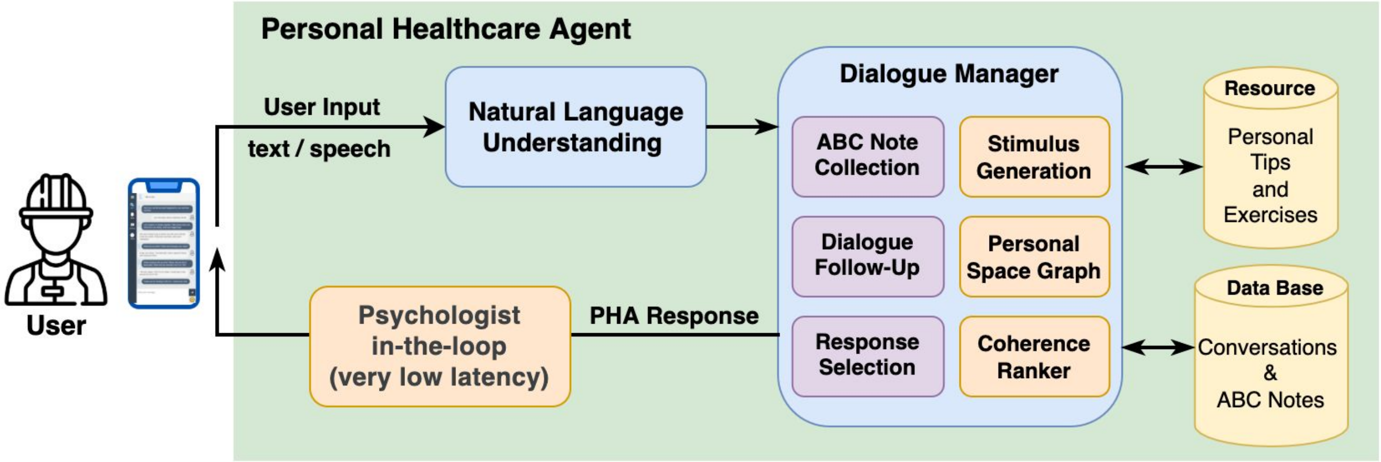
- The therapists engaged in the studies believe the blended intervention may improve the patients' mental health since it results in continuous support provision by the conversational agent.



[4] Danieli M, Ciulli T, Mousavi S, Riccardi G. ***A Conversational Artificial Intelligence Agent for a Mental Health Care App: Evaluation Study of Its Participatory Design.*** JMIR Form Res 2021;5(12)
[5] Danieli M, Ciulli T, Mousavi MS, Silvestri G, Barbato S, Di Natale L, Riccardi G. ***TEO: Assessing the Impact of Conversational AI in the Treatment of Stress and Anxiety in Aging Persons: A Randomized Controlled Trial Study.*** JMIR Preprints.

Lets Try Retrieval Response Models!

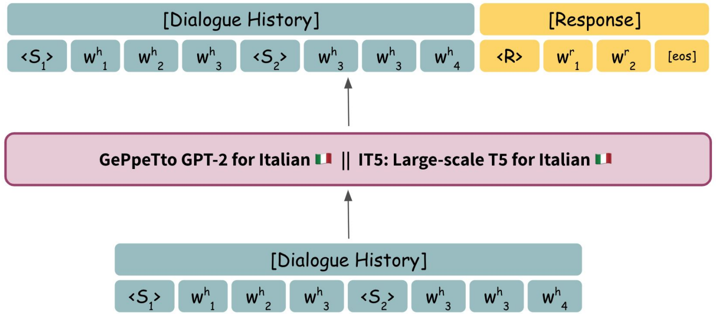
- We developed a Retrieval Based PHA for the mental health domain.[4,5]



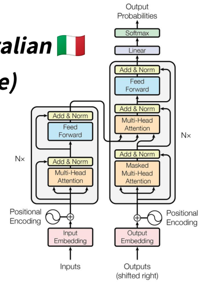
[4] Danieli M, Ciulli T, Mousavi S, Riccardi G. ***A Conversational Artificial Intelligence Agent for a Mental Health Care App: Evaluation Study of Its Participatory Design.*** JMIR Form Res 2021;5(12)
[5] Danieli M, Ciulli T, Mousavi MS, Silvestri G, Barbato S, Di Natale L, Riccardi G. ***TEO: Assessing the Impact of Conversational AI in the Treatment of Stress and Anxiety in Aging Persons: A Randomized Controlled Trial Study.*** JMIR Preprints.

Can we Train now?

- We are experimenting with generative models
- We are currently working on Conversational Fine-Tuning of Pre-Trained Language Models



IT5: Large-scale T5 for Italian 🇮🇹
(small, base, large)



GePpeTto GPT-2 for Italian 🇮🇹
(small)

GPT-2 SMALL

Issues & Challenges

- The design, training and development of a PHAs is strongly subject to factors of accuracy, appropriateness and acceptance by the user and ethical compliance.

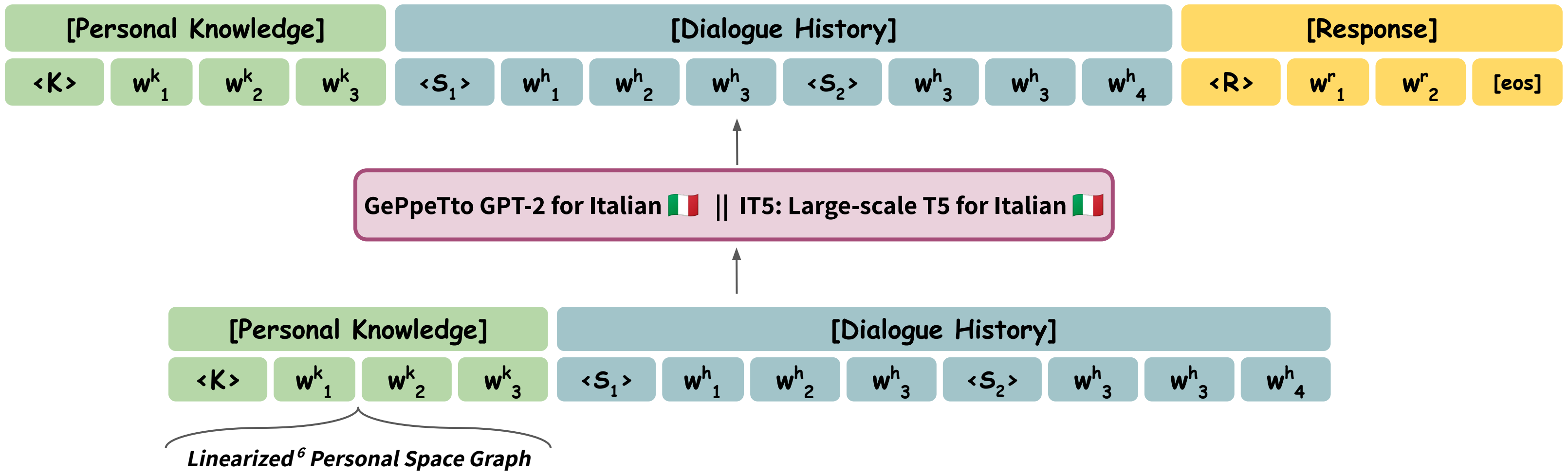
- Data-driven conversational models (including GPT-x, T5, …) suffer from the major issues of generating toxic responses, which are unethical, offensive, biased, and dangerous leading to ethical problems, and are unexplainable due to their nature.

- Meanwhile, detecting and controlling such toxic output in these models are yet not possible due to the nature of such models.

- Personalization of CAs is limited to food preference and pronoun selection.

Can we Train now?

- End-to-End generative models are known for generic and inappropriate response generation
    - We are experimenting the grounded-response generation



[6] Ribeiro, L. F., Schmitt, M., Schütze, H., & Gurevych, I. (2021, November). ***Investigating Pretrained Language Models for Graph-to-Text Generation.*** In Proceedings of the 3rd Workshop on Natural Language Processing for Conversational AI (pp. 211-227).

Thank You.