LINKSFOUNDATION.COM

FONDAZIONE links
PASSION FOR INNOVATION

Applied Data Science Project

Politecnico di Torino

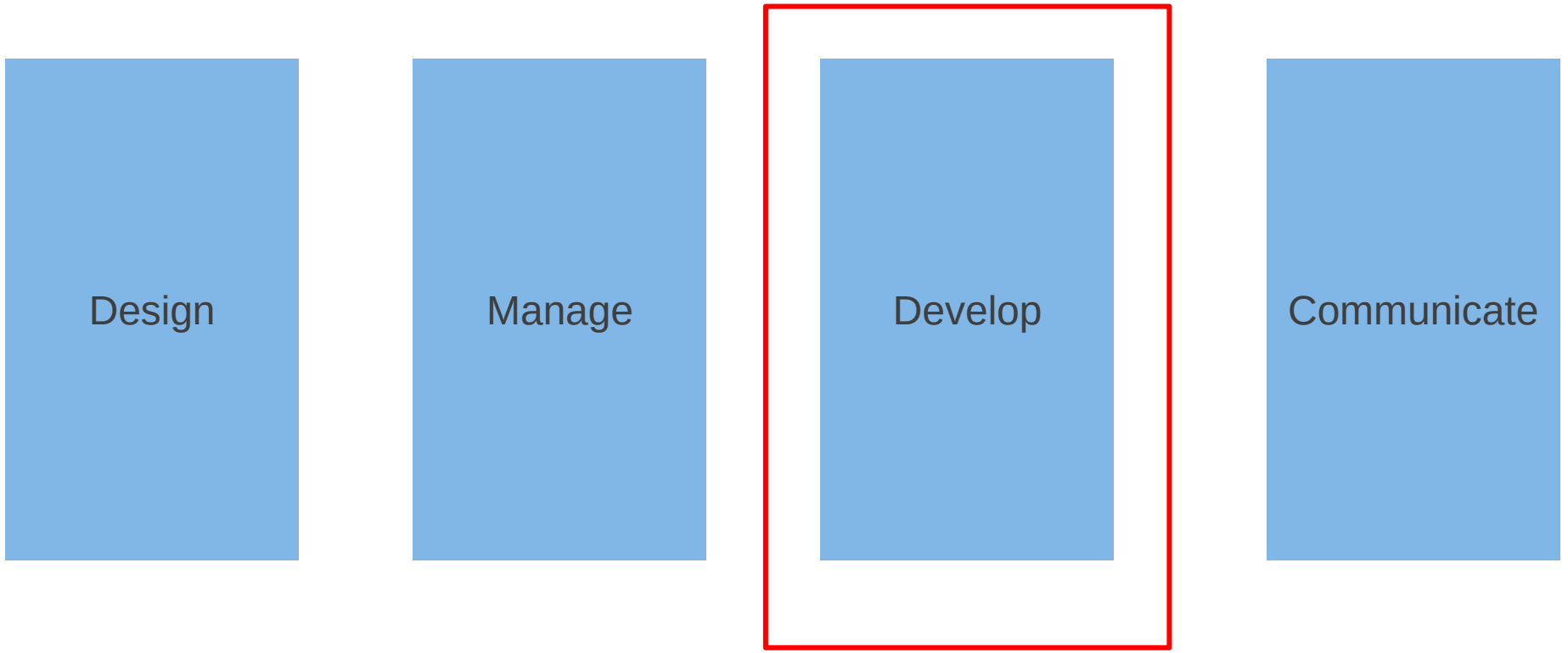
L13 – Foundation models: hub and examples

ellis
European Laboratory for Learning and Intelligent Systems

Giuseppe Rizzo

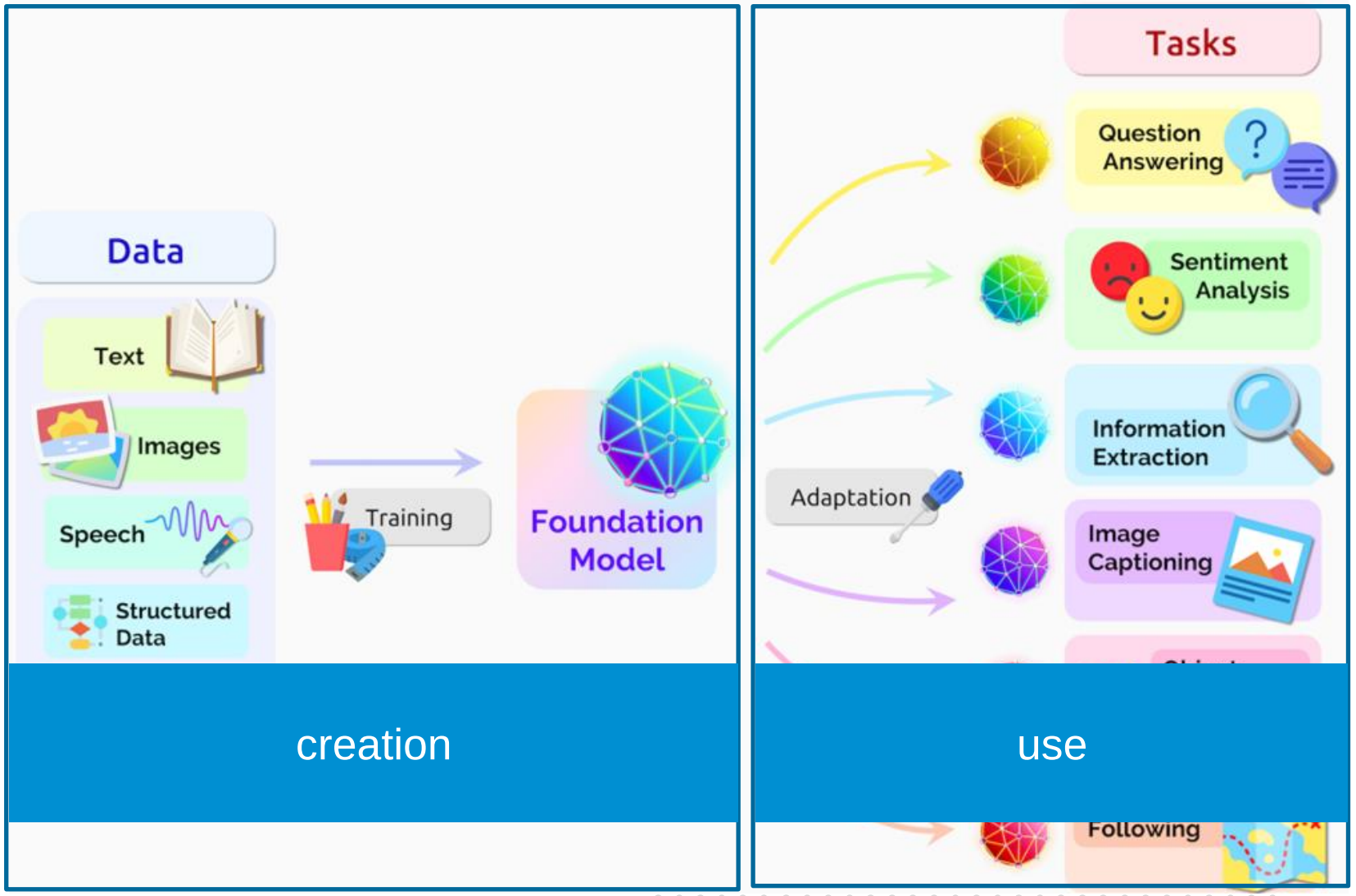
Turin, October 11, 2022

Pillars



linksfoundation.com
COPYRIGHT ©2021 LINKS

Foundation models



linksfoundation.com
COPYRIGHT ©2021 LINKS

Which one to use?

?

linksfoundation.com
COPYRIGHT ©2021 LINKS

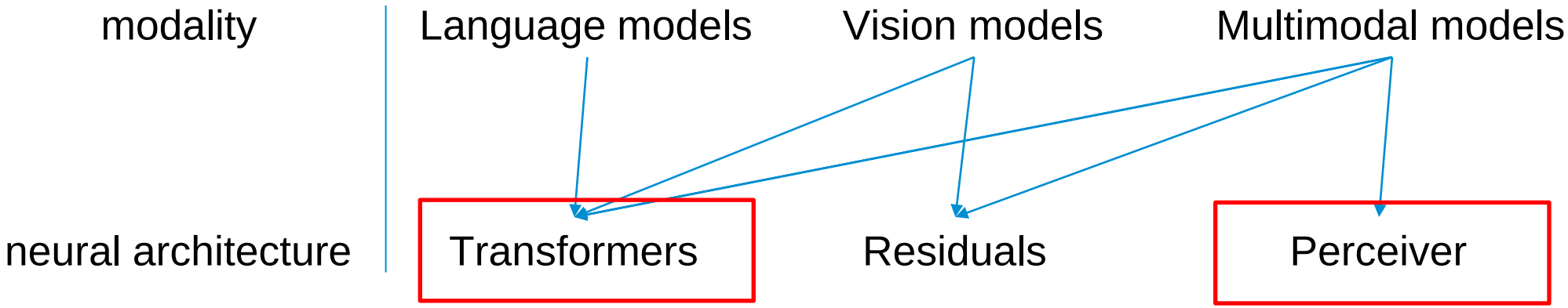
Foundation models featured by modality

Name	Paper	Modality
BERT	Devlin et al., “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding” paper	Language
XLNET	Yang et al, “XLNet: Generalized Autoregressive Pretraining for Language Understanding” paper	Language
ResNet	He et al., “Deep Residual Learning for Image Recognition” paper	Vision
ViT	Dosovitskiy et al., “An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale” paper	Vision
DALL-E	Ramesh et al., “Zero-Shot Text-to-Image Generation” paper	Multimodal
CLIP	Radford et al., “Learning Transferable Visual Models From Natural Language Supervision” paper	Multimodal

Notice: some those currently under study and use for being considered foundation models

linksfoundation.com
COPYRIGHT ©2021 LINKS

Modality and architecture



- As of today, they are considered state of the art for encoding the modality of use

linksfoundation.com
COPYRIGHT ©2021 LINKS

Hub of models

https://huggingface.co

It archives

- **foundation models** such as BERT or XLNet and
- **task-specific models** created by the community of Hugging face to solve downstream tasks such as token classification, object recognition

linksfoundation.com
COPYRIGHT ©2021 LINKS

Use of BERT for Masking and NER

Check L13 - ADSP - Hugging face.ipynb

linksfoundation.com
COPYRIGHT ©2021 LINKS

Finetuning of BERT for NER

Colab Notebook

Token Classification

File Edit View Insert Runtime Tools Help Cannot save changes

Share

+ Code + Text Copy to Drive

RAM Disk

Editing

If you're opening this Notebook on colab, you will probably need to install 🤗 Transformers and 🤗 Datasets. Uncomment the followin and run it.

```
[ ]  #! pip install datasets transformers seqeval
```

If you're opening this notebook locally, make sure your environment has an install from the last version of those libraries.

To be able to share your model with the community and generate results like the one shown in the picture below via the inference API, are a few more steps to follow.

First you have to store your authentication token from the Hugging Face website (sign up here if you haven't already!) then execute the following cell and input your username and password:

```
[3]  !pip install huggingface_hub

     from huggingface_hub import notebook_login

     notebook_login()
```

```
Login successful
Your token has been saved to /root/.huggingface/token
```

Then you need to install Git-LFS. Uncomment the following instructions:

```
[ ]  # !apt install git-lfs
```

Make sure your version of Transformers is at least 4.11.0 since the functionality was introduced in that version:

```
[ ]  import transformers

     print(transformers.__version__)
```

You can find a script version of this notebook to fine-tune your model in a distributed fashion using multiple GPUs or TPUs here.

Fine-tuning a model on a token classification task

In this notebook, we will see how to fine-tune one of the 🤗 Transformers model to a token classification task, which is the task of predicting a label for each token.

Widget inference representing the NER task

The most common token classification tasks are:

- NER (Named-entity recognition) Classify the entities in the text (person, organization, location...).
- POS (Part-of-speech tagging) Grammatically classify the tokens (noun, verb, adjective...)
- Chunk (Chunking) Grammatically classify the tokens and group them into "chunks" that go together

We will see how to easily load a dataset for these kinds of tasks and use the `Trainer` API to fine-tune a model on it.

This notebook is built to run on any token classification task, with any model checkpoint from the Model Hub as long as that model ha version with a token classification head and a fast tokenizer (check on this table if this is the case). It might just need some small adjustments if you decide to use a different dataset than the one used here. Depending on you model and the GPU you are using, you r

COPYRIGHT ©2021 LINKS

FONDAZIONE links

PASSION FOR INNOVATION

Thank you for your attention.

Questions?

CONTACTS

Giuseppe Rizzo

Program Manager (LINKS Foundation) and
Adjunct Professor (Politecnico di Torino)

giuseppe.rizzo@polito.it

FONDAZIONE LINKS
Via Pier Carlo Boggio 61 | 10138 Torino
P. +39 011 22 76 150
LINKSFOUNDATION.COM