WP3: Data–driven approach to irradiation experiments proposals management and evaluation

Jaroslaw Szumega, CERN and Mines Paris, PSL University
jaroslaw.szumega@cern.ch
RADNEXT 2nd Annual Meeting – 9-10 May 2023
https://indico.cern.ch/e/radnext-2023

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No **101008126**

Context of WP03

Salvatore Fiore

(CERN)

WP leader

Jochen Kuhnhenn

(Fraunhofer INT)

Deputy WP leader

Jaroslaw Mikolaj Szumega

(Mines Paris, PSL University, CERN)

PhD student

Blerina Gkotse

(CERN)

Postdoc researcher

Pierre Jouvelot

(Mines Paris, PSL University)

Senior researcher

Context of WP03

Transnational Access management and harmonization

- **Transnational Access activities definition**
  - Application for beam time
  - Coordination of user selection and support
  - Optimised management of tests in facilities

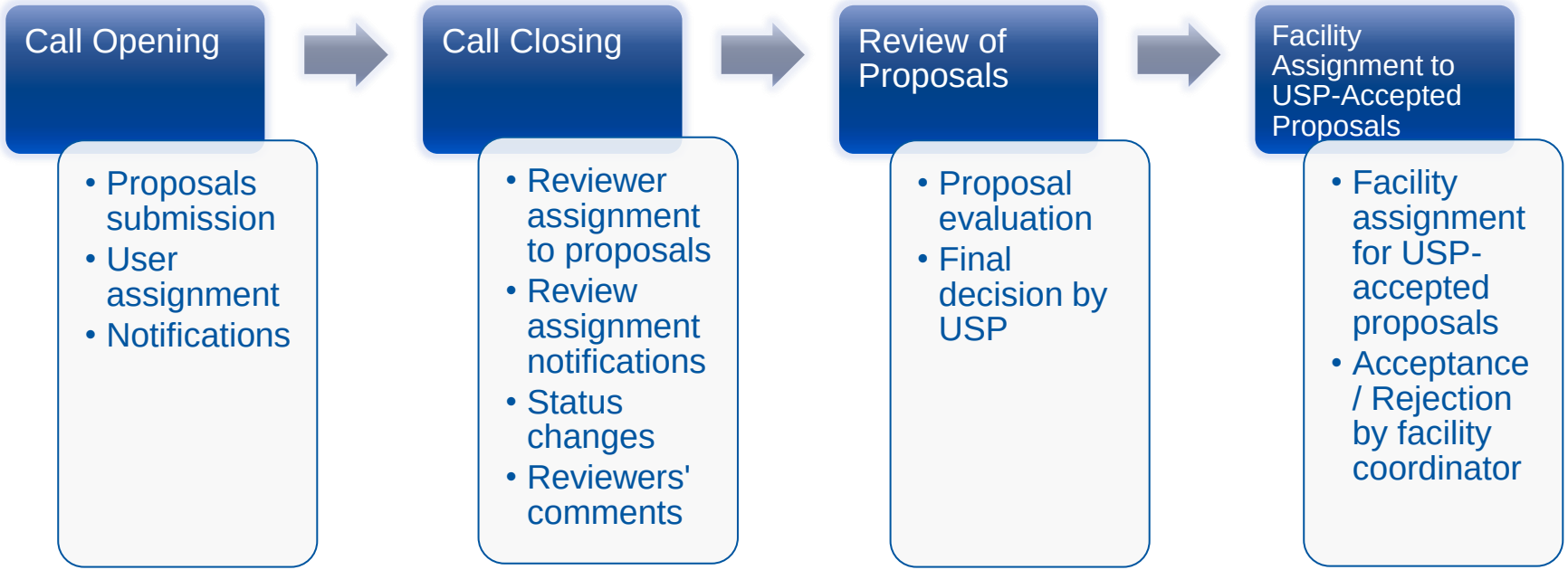
- **Tasks in WP03**
  - TA submission management via Online Portal
  - Data-model definition for users and experiments management
  - **Natural Language Processing techniques to support TA activities**
  - Definition of best practices and procedures

Content

- Context of WP03
- **Transnational Access submissions management**
  - Beamtime application procedure
  - Radnext TA Online Portal
  - Review of best practices and procedures
- Research on Natural Language Processing techniques
  - Submissions and reviews as data
  - Knowledge extraction
  - Towards automatic assessment
- Conclusions

Transnational Access submissions management

Beamtime application procedure

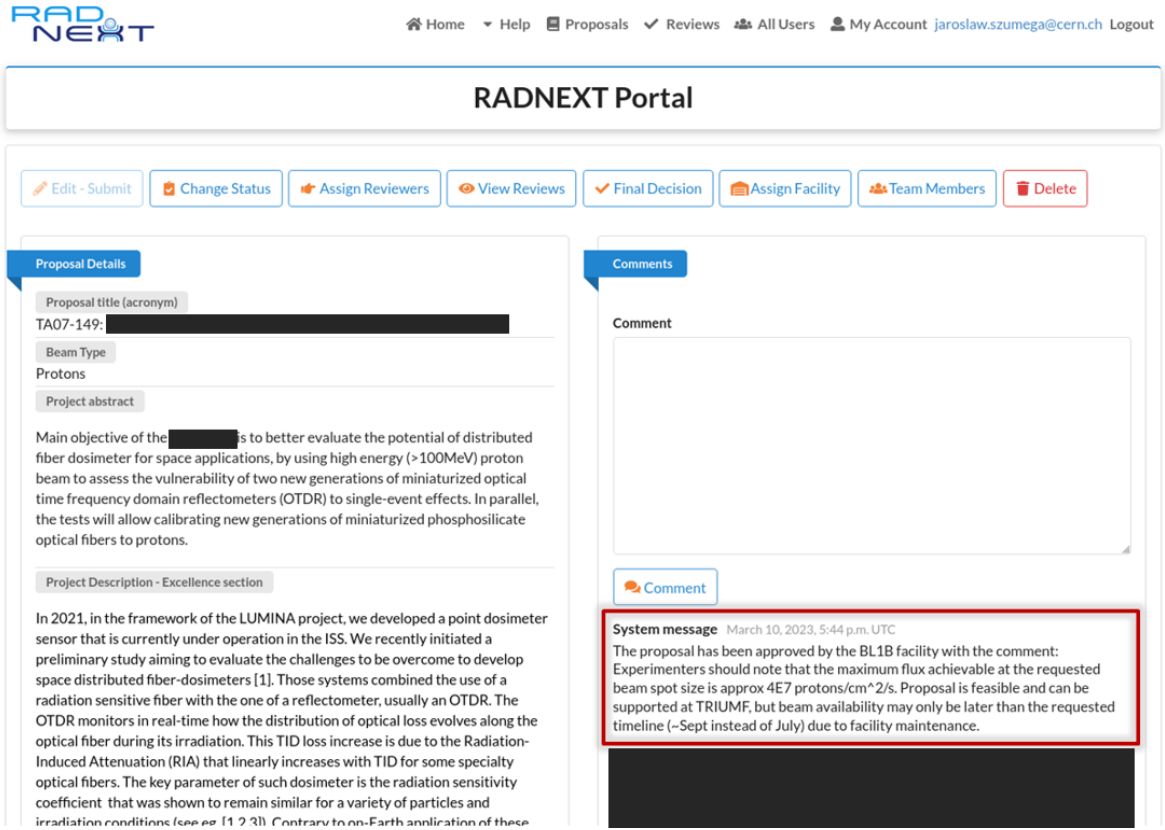


Transnational Access submissions management

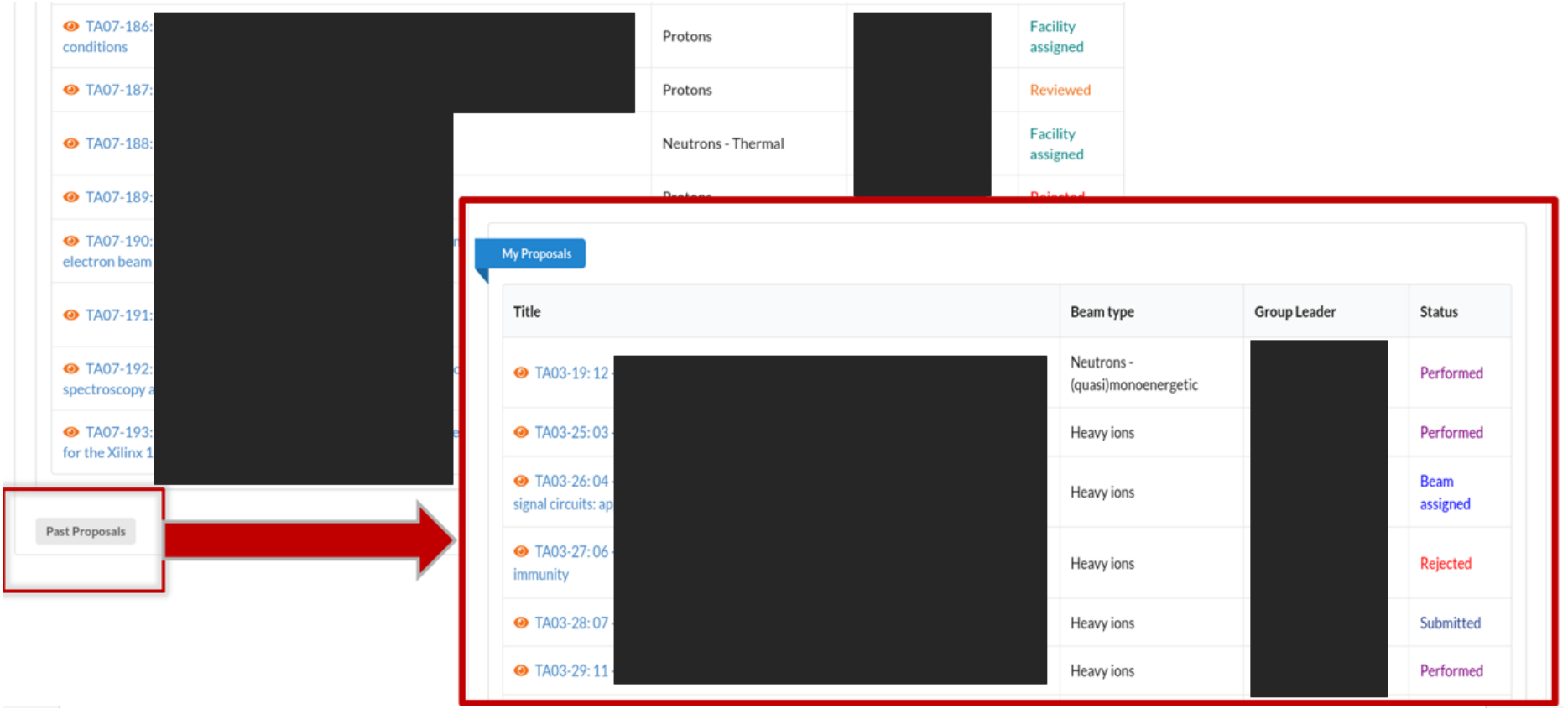
Radnext TA Online Portal – changes & improvements

- Automatic notifications for facility assignments
- Separation between closed and active proposals
- Additionally, separation of closed and active reviews
- Introduced details of requested beam properties
- Possibility to write comments
- Automated call opening
- Proposal „last modified” to simplify management

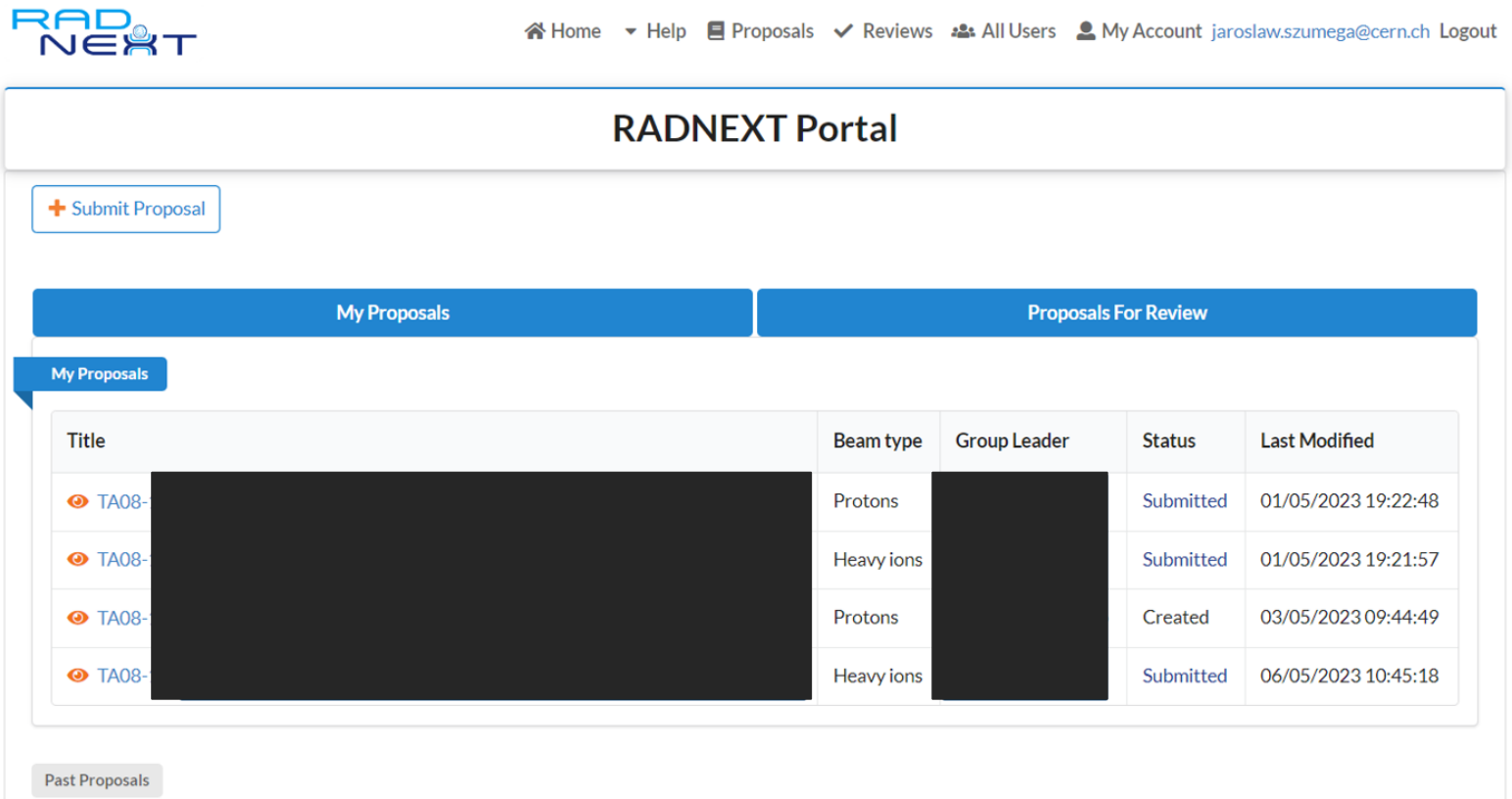
- # Automatic notifications for facility assignments



- # Separated closed and active proposals



- „Last modified” column added



Transnational Access submissions management

Review of practices and procedures in facilities

- Scheduled interviews with Facilities Coordinators to *extract knowledge* of best procedures in facilities
  (performed by Gerd Datzmann, RADNEXT partner)

- Creating a meaningful guideline and set of rules for RADNEXT beneficiaries

- Sharing the information to unify the best practices

Content

- Context of WP03
- Transnational Access submissions management
  - Beamtime application procedure
  - Radnext TA Online Portal
  - Review of best practices and procedures
- **Research on Natural Language Processing techniques**
  - Submissions and reviews as data
  - Knowledge extraction
  - Towards automatic assessment
- Conclusions

Research on Natural Language Processing

Submissions and reviews as data

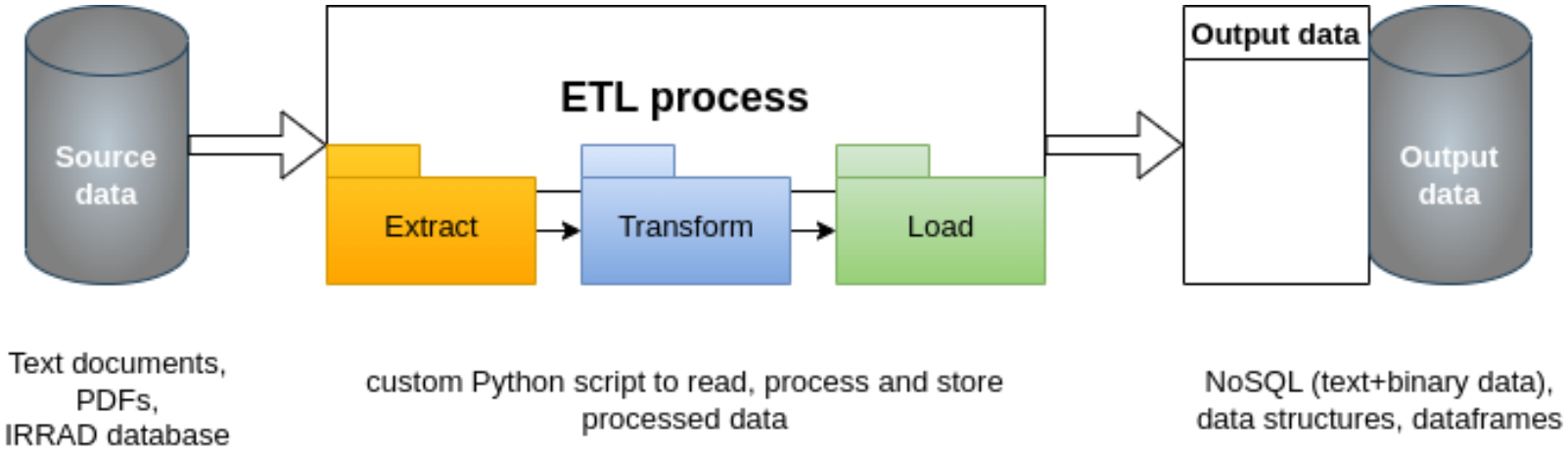
- Data extracted from the portal's database

	id	project_title	project_acronym	beam_type	project_abstract	project_description_excellence	project_description_impact	project_description_implem
0	19	12 - SEE evaluation on RFID tags under fast ne...	TagSEEn	Neutrons - (quasi)monoenergetic	Proper management of irradiated samples and sy...	Irradiation of material samples, devices and p...	Previous knowledge about RFID tags radiation h...	Beam energy range: 14 I range: u
1	25	03 - Integrated Sensor Interface for Harsh Rad...	ISIHR	Heavy ions	This work presents a capacitance-to-digital co...	Capacitance sensors are increasingly being use...	The performance of the CDC prototype under hea...	The test setup is comp sensor I
2	26	04 - Statistical approach to defect simulation...	Statset	Heavy ions	This project draws a parallelism between fabri...	The demonstrator that has been designed for th...	- Validation of the simulation methodology: If...	The main goal of this experir
3	27	06 - Qualification of the ULTRASAT sensor for ...	ULTRASAT sensor SEE test	Heavy ions	The ULtraviolet TRansient Astronomical SATelli...	The ULTRASAT is an astronomical satellite miss...	The UV camera is the only detector on-board of...	It is planned to test the sens

Research on Natural Language Processing

Submissions and reviews as data

- Extract Transform Load (ETL) process schema



Research on Natural Language Processing

Submissions and reviews as data

```
In [3]: etl_client.transform();
        etl_client.transformed_df
```

Out[3]:

	proposal_id	project_abstract	project_description_impact	project_description_excellence	proposal_status	excellence_vote	impact_vote	implementation_vote
0	19	Proper management of irradiated samples and sy...	Previous knowledge about RFID tags radiation h...	Irradiation of material samples, devices and p...	Performed	0.700000	0.766667	0.733333
1	25	This work presents a capacitance-to-digital co...	The performance of the CDC prototype under hea...	Capacitance sensors are increasingly being use...	Performed	0.700000	0.633333	0.800000
2	26	This project draws a parallelism between fabri...	- Validation of the simulation methodology: If...	The demonstrator that has been designed for th...	Facility assigned	0.766667	0.766667	0.800000
3	27	The ULtraviolet TRansient Astronomical SATelli...	The UV camera is the only detector on-board of...	The ULTRASAT is an astronomical satellite miss...	Rejected	0.633333	0.633333	0.433333
4	28	Current space-grade processors cannot satisfy ...	ESA has recognised that COTS SoCs will be game...	The z7000 has received increasing attention fr...	Submitted	0.766667	0.766667	0.633333

Research on Natural Language Processing

Knowledge extraction

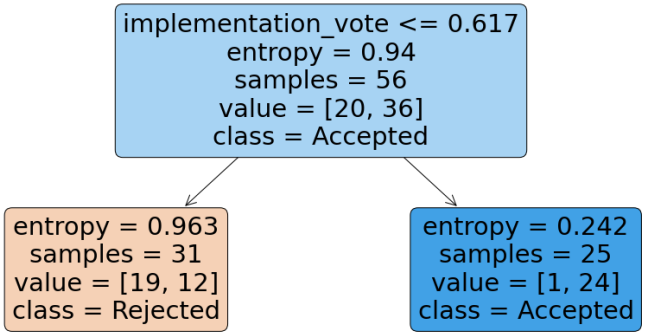
- Facilities interviews to be anonimized and analyzed by RADNEXT partners, specifically WP3 collaborators and Datzmann company
- Informations from the USP and reviewers to determine the most important features and values of each proposal

The implementation section of each proposal – key information towards acceptance…

Research on Natural Language Processing

Knowledge extraction - experiments

- Decision trees learn the rules on how to „divide” the data



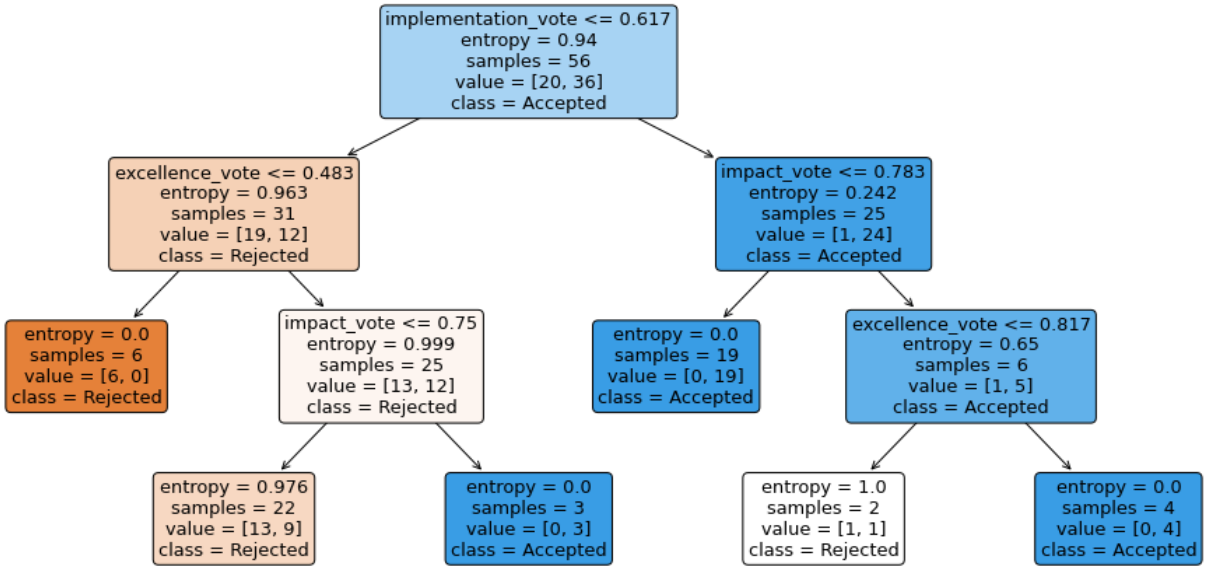


Fig. 1 Implementation vote at the root of decision trees

Research on Natural Language Processing

Towards automatic assessment

- Use advanced NLP techniques to support **users**
  1. Pre-evaluation assessment of proposals
  2. Generated advices regarding the quality of a proposal

- Use advanced NLP techniques to support **reviewers**
  1. Hints about proposals rejection/acceptance
  2. Pre-reviews generated to support internal work and assessment

Research on Natural Language Processing

Towards automatic assessment

1. Representation of text in numerical form – „embeddings”
2. Output generation with Large Language Models (LLMs)

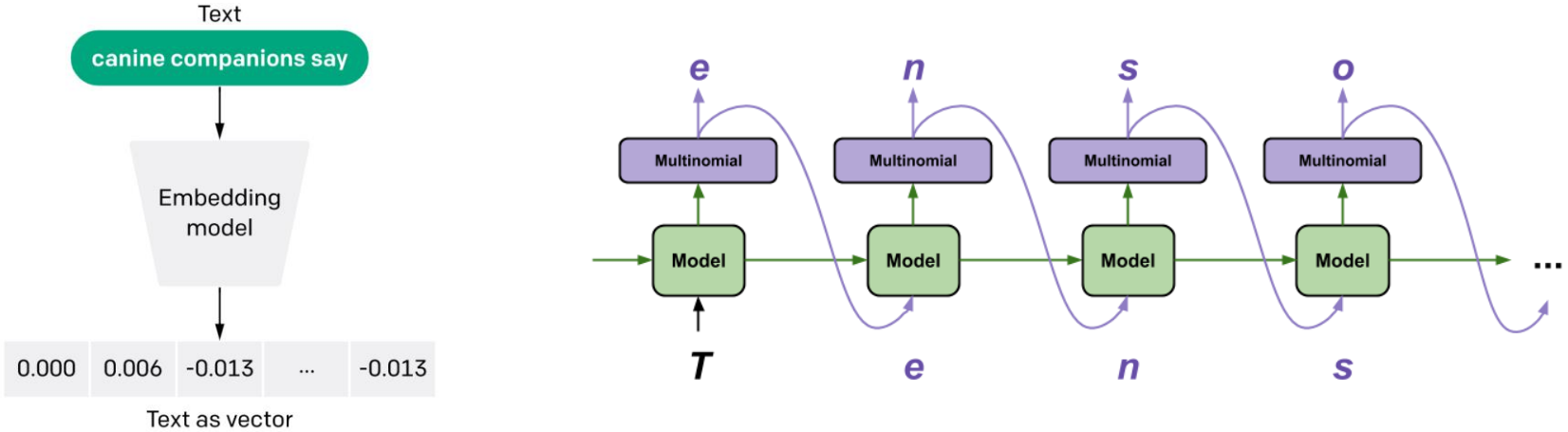


Fig. 2 Visual representation of NLP techniques (i) text embedding model, (ii) sequence generator
Source: openai.com

Conclusion

- Discussed context of WP3

- Transnational access activities
  - Focus on user Online Portal
  - Processing interviews and collecting data

- Natural Language Processing
  - Submissions and reviews as data
  - Ideas and first experiments to support the RADNEXT roadmap.

- **Summary : not only management but also engineering and research tasks to provide suport for all members in RADNEXT**

WP3: Data–driven approach to irradiation experiments proposals management and evaluation

Jaroslaw Szumega, CERN and Mines Paris, PSL University
jaroslaw.szumega@cern.ch
RADNEXT 2nd Annual Meeting – 9-10 May 2023
https://indico.cern.ch/e/radnext-2023

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No **101008126**