SnT 2025
8 SEPTEMBER ONLINE DAY
9 TO 12 SEPTEMBER AT HOFBURG PALACE, VIENNA & ONLINE
CTBT: SCIENCE AND TECHNOLOGY CONFERENCE
LIGHTNING TALK
P4.3-576

Evaluating AI Code Assistants for Enhanced Software Development in CTBTO: A Modular Approach

Evangelos Dellis
Comprehensive Nuclear-Test-Ban Treaty Organization (CTBTO)

➢ We evaluate **AI Code Assistants and Agents** that can be deployed in our GPU infrastructure

➢ You will learn that an AI coding Assistant/Agent:
  o is a **tool** that uses a LLM that is fine-tuned for coding tasks and for "tool-use"
  o predictability and agency **affects** the coding task
  o augmentation vs automation ➔ when to use which option

➢ We explore various types of AI Coding tools connected to our CTBTO codebase:
  o what is an AI IDE & Copilot?
  o how to use a CLI-based agent?
  o what is a Teammate or SWE-agent?
  o learn about SWE-bench ➔ the de-facto coding benchmark

➢ Learn how to use an **autonomous SWE-Agent**:

  o that is open-source (including the LLM)
  o running locally in your GPU infrastructure
  o and solves real issues!

CTBTO
PREPARATORY COMMISSION