Quiz 7 - Enterprise Workflows w/ Nicolas Chapados & Alexandre Drouin (10/21)

Total points 5/5

INSTRUCTIONS:
Each of these quizzes is completion based, however we encourage you to try your best for your own education! These quizzes are a great way to check that you are understanding the course material. You can attempt this quiz as many times as you wish. You only need to complete the quizzes if you wish to earn a completion certificate. More information at the bottom of the course website.

IMPORTANT:
In order to receive credit, use the same email address as the one used to sign up for the course. If you are not sure which email you used, just complete the sign up form again with your preferred email.

PROBLEMS?
If you have any technical difficulties about this quiz, please ask course staff in our LLM Agents Discord.

Email *

christopherbare@gmail.com

✓ **What is the difference between API agents and web agents?** * 1/1

- ( ) API agents are inherently riskier than web agents, which can be solved by providing appropriate APIs
- ( ) API agents are best used for human-visualization tasks whereas web agents specialize in search history and chat history retrieval
- (•) API agents handle tasks by leveraging predefined programmatic interfaces, while web agents directly navigate and interact with web pages in a manner similar to how a user would ✓
- ( ) API agents are more versatile but have higher latency, while web agents are restricted to specific use cases

✓ **What are TapeAgents?** * 1/1

- ( ) A new type of hardware agent that interacts with tapes to store data more efficiently than traditional storage systems
- (•) A framework for building agents that uses a structured log (tape) to track, resume, and audit agent sessions, facilitating step-by-step debugging and session persistence ✓
- ( ) A software agent that uses tapes to archive historical data, providing access to previous records for training and analysis
- ( ) An API agent that manages data streams by compressing logs into a tape format for better performance and lower latenc

✓ **How do we evaluate web agents?** * 1/1

- (●) By testing their performance on real-world scenarios and assessing both their actions and final outcomes, with setups that can be hosted locally or run remotely ✓
- ( ) By measuring their ability to remember and replicate exact sequences of actions
- ( ) By exclusively running tests on local servers to control for external variables like network latency
- ( ) By focusing on their performance with simple, open-source software to avoid complex configurations

✓ **How did WorkArena++ improve upon WorkArena?** * 1/1

- ( ) By reducing the complexity of tasks to make it easier for agents to achieve full task automation
- ( ) By limiting the benchmark to only include open-source large language models to address performance disparities
- (●) By expanding the number of tasks, focusing on a broader range of skills such as planning, problem-solving, and reasoning, to better assess the capabilities of autonomous agents in enterprise settings ✓
- ( ) By removing the multimodal observation feature to simplify the environment for agent evaluation

✓ **What is the difference between top-down and bottom-up assessment?** * 1/1

- ○ Top-down assessment focuses on specific tasks, while bottom-up assessment looks at the overall performance across multiple jobs
- ○ Top-down assessment is more accurate but limited in coverage, while bottom-up assessment is less detailed but covers more areas
- ○ Top-down assessment is used exclusively for manual evaluation, whereas bottom-up assessment relies solely on AI-based automation
- ◉ Top-down assessment estimates task automation broadly across jobs, while bottom-up assessment evaluates specific tasks in detail and maps them back to job automation probability ✓