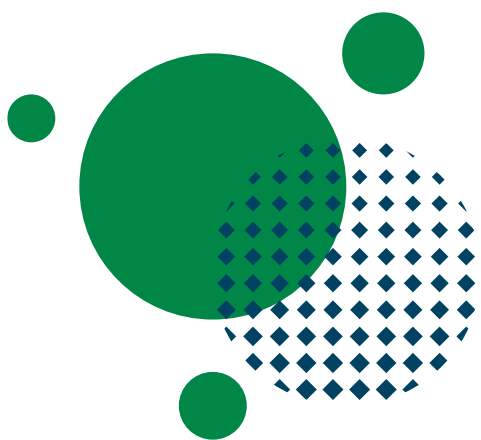




Chester County
**Economic
Development
Council**



TECH360°

Get MSTY With Me,

Using Local LLM for Security and Privacy

October 15, 2025

Ben Greisler
ben@kadimac.com



***My* Intro to Local AI**



Goals

- Why use local versus hosted AI (Pros and cons)
- What is needed to run local AI
- Caveats regarding local AI
- How to experiment with local AI
- Demonstration
- Q&A

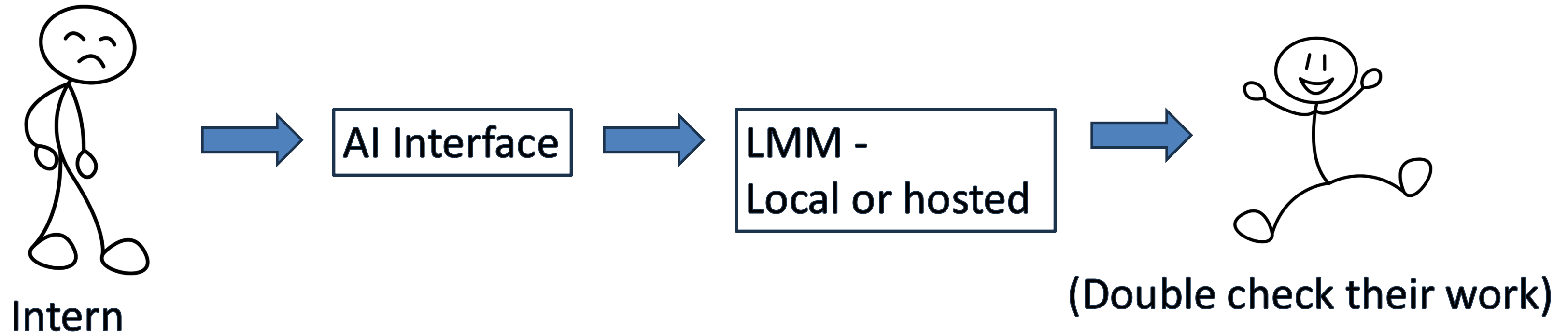
Note: This is a primer, not a deep dive



Why Local AI?

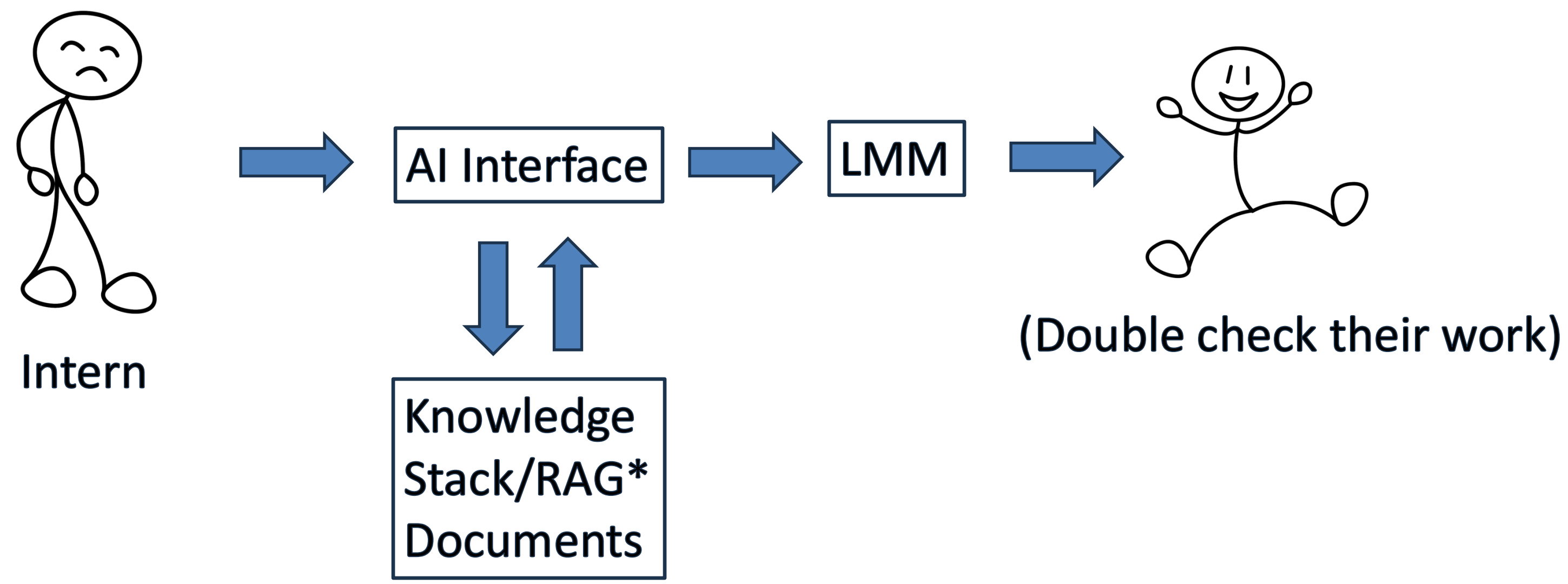
- Control
 - Environment is completely internal
 - No reliance on hosted systems
- Security and privacy
 - No data leaves your systems
- Caveats
 - Not all roses at the moment.
 - Can be experimental
- Cost
 - Hardware limited
 - Licensing for features

AI Workflow: Standard

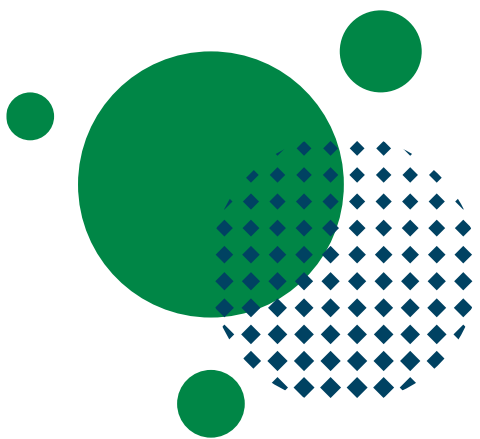




AI Workflow: Local w/RAG

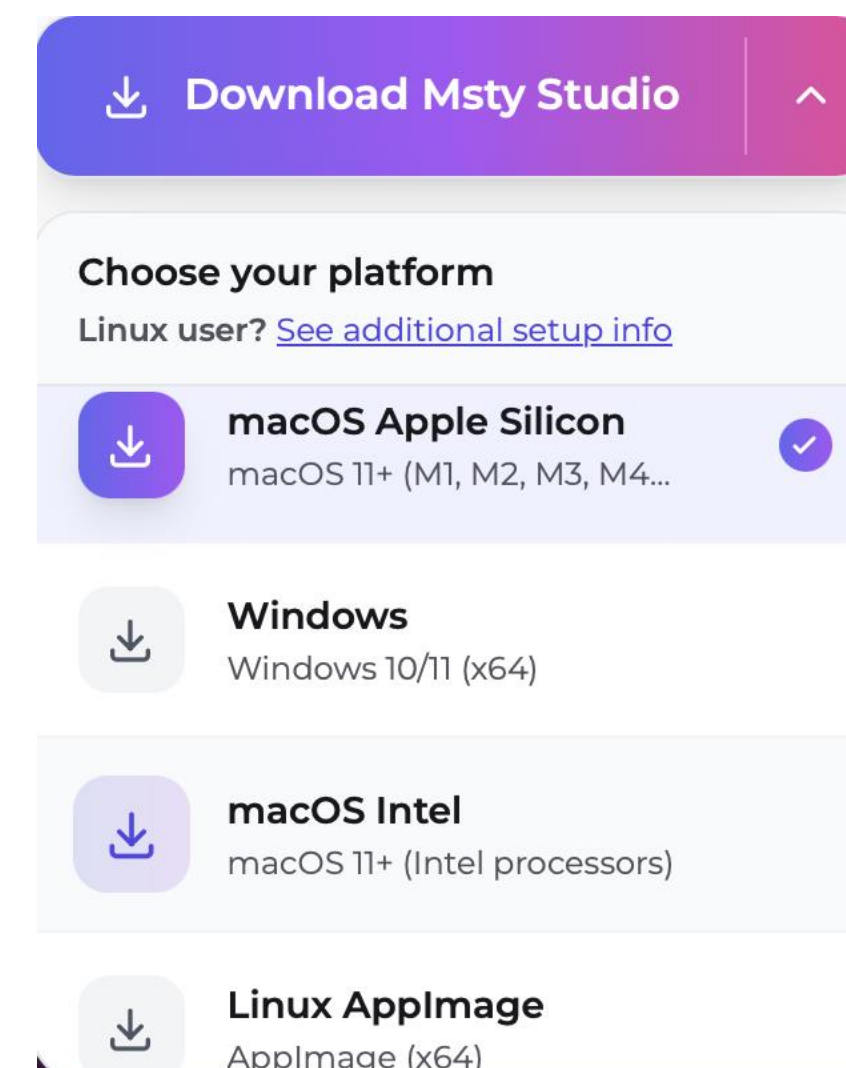


* RAG = Retrieval-Augmented Generation



Setting up local AI

MSTY Studio:
<https://msty.ai>





MSTY Components

- Hardware choices
- MSTY Studio app
- Models
- Embedding models
- Knowledge Stacks aka: Retrieval-Augmented Generation (RAG)
 - <https://msty.ai/blog/knowledge-stacks>



Tuning: KS/RAG

- What type of documents
 - .pdf, .csv, .md, .json, .jsonl, .xlsx, .docx, .rtf, .txt, .pptx or any other supported file types
- Compose settings
 - Embedding model
 - Chunking method, Sentence or Recursive Character
 - Chunk size
- Query settings
 - Number of chunks
 - Similarity
 - Search type



Demo

Time to tempt the demo gods



Q&A

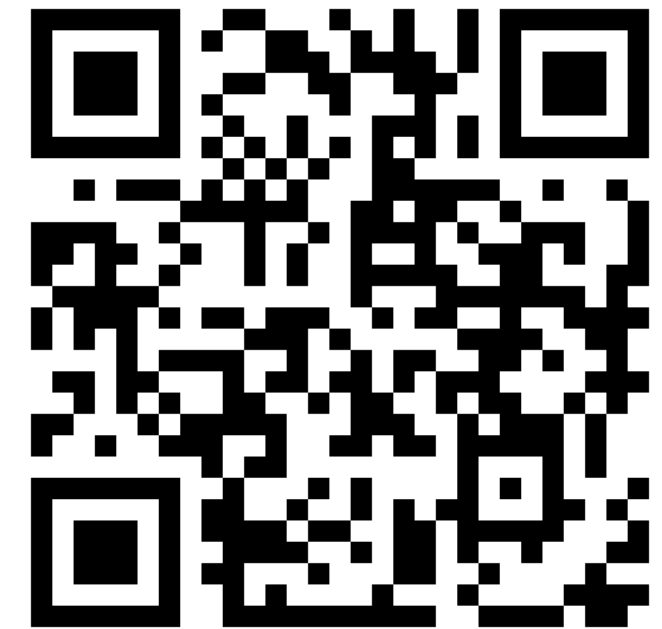




Contact Info

Ben Greisler

ben@kadimac.com



Feedback please!

