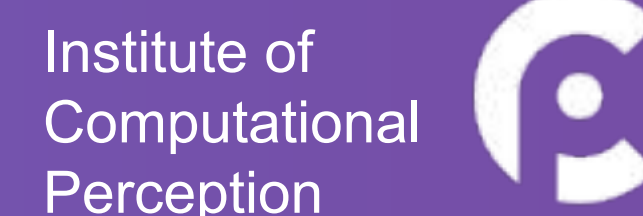




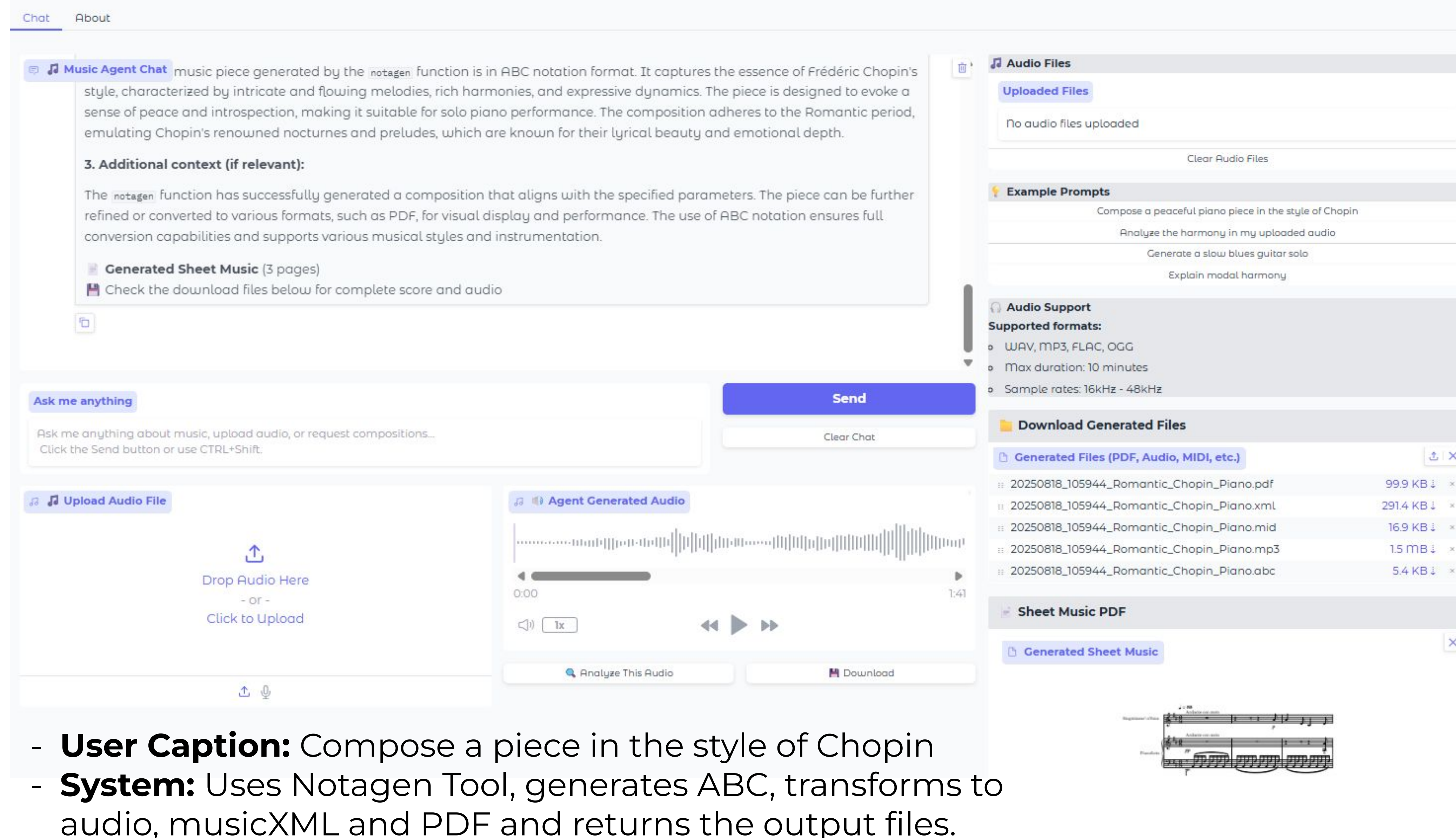
**Emmanouil Karystinaios<sup>1,2</sup>**  
 {firstname}.{lastname}@jku.at

**JKU**  
JOHANNES KEPLER  
UNIVERSITY LINZ



WeaveMuse is an open-source agentic system for music understanding & generation:

- ## System Preview



## Functional View

- Accepts multimodal inputs: text, symbolic scores (abc for now), audio
- Understands music through analysis tools (harmony, rhythm, structure)
- Generates new material: symbolic notation, audio, hybrid transformations
- Transforms across modalities (e.g., score  $\rightarrow$  audio, text  $\rightarrow$  score)

## Architectural View

- User Input Layer: text prompts, uploaded scores, audio files
- Manager Agent: plans, selects tools, and verifies results (perceive–plan–act loop)
- Specialized Tools:
  - *AudioFlamingo*, *NotaGen*, *StableAudio*
  - *Score renderers* → PDF/visual outputs
- Output Layer: multimodal results (scores, audio, structured metadata)

## Deployment / Resource View

(Where it Runs)

- Lightweight local setup: efficient execution on CPU/GPU laptops
- Hybrid mode: local orchestration + hosted APIs for heavy models
- Efficiency optimizations: quantization, caching, memory offloading, dynamic precision, lazy tool loading.
- Scalable setup: reproducible deployment across hardware tiers

## Implemented Tools and Models

- ❑ **ChatMusician** – text-to-symbolic music generation
- ❑ **NotaGen** – symbolic generation in ABC notation
- ❑ **Stable Audio Open** – text-to-audio synthesis
- ❑ sympy / **Partitura** / Parsers – parsing and symbolic feature analysis
- ❑ Score Visualization – symbolic-to-PDF rendering - **EngravingGNN**
- ❑ Audio Understanding – spectral, rhythmic, context and harmonic analysis (**QwenAudio** and **Audio Flamingo**)
- ❑ Gradio Interface

## Resources

