

# DeepFakeChain

*A chain of trust against fake multimedia content!*

## Presenters:

*Symeon Papadopoulos, Nikos Giatsoglou (CERTH)*



**CERTH**  
CENTRE FOR  
RESEARCH & TECHNOLOGY  
HELLAS

**Zelus**



# The challenge

Entering the era of hyper-realistic synthetic media



<https://twitter.com/EliotHiggins/status/1637927681734987777>

# The need

- Novel ways for algorithmic detection based on increased data for forgeries
- Collaborative human evaluation
- Auditable access to deepfake evaluations stored in blockchain



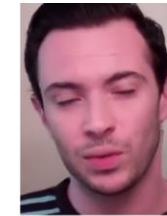
Deepfakes



FaceSwap



Face2Face



NeuralTextures



Faceshifter



ATVG-Net



StarGAN2



BlendFace



SC-FEGAN



MMreplacement



DALL-E 2



MidJourney



Stable Diffusion 1.5

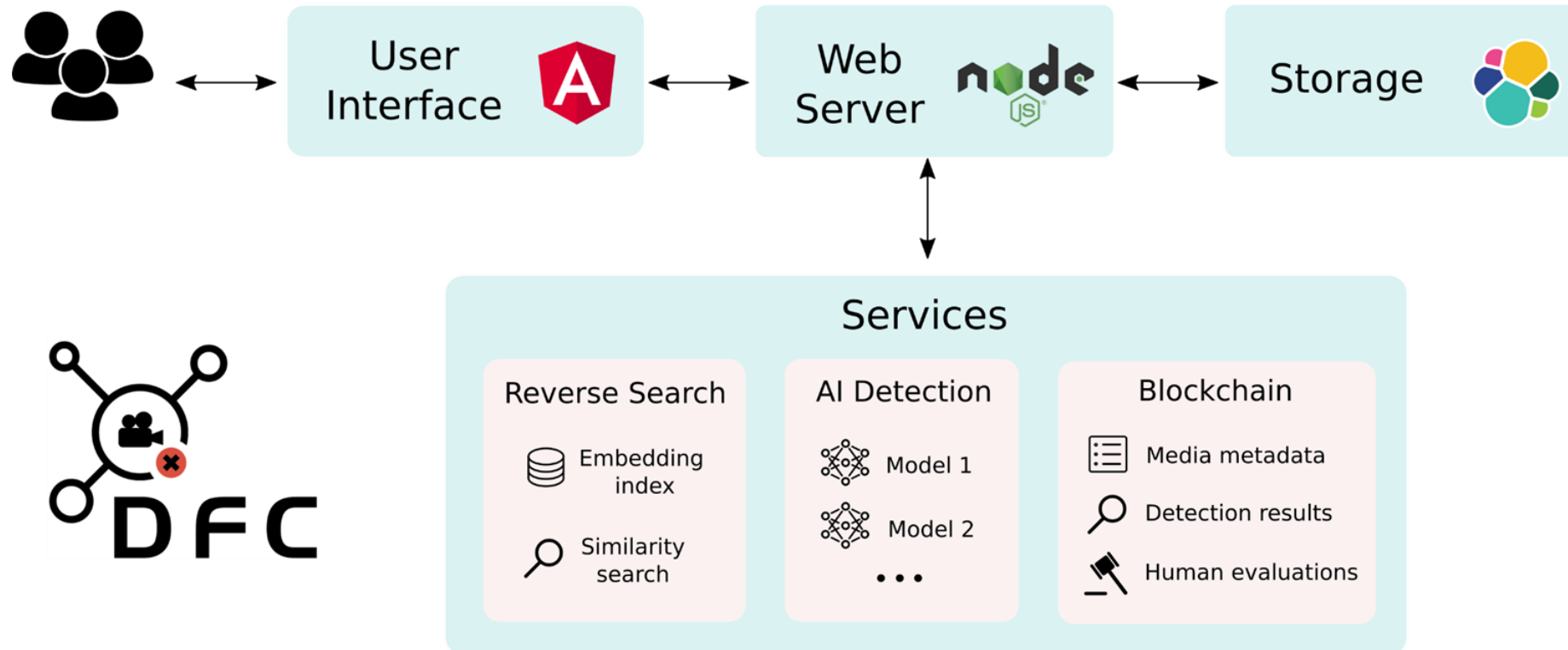


Stable Diffusion 2.1

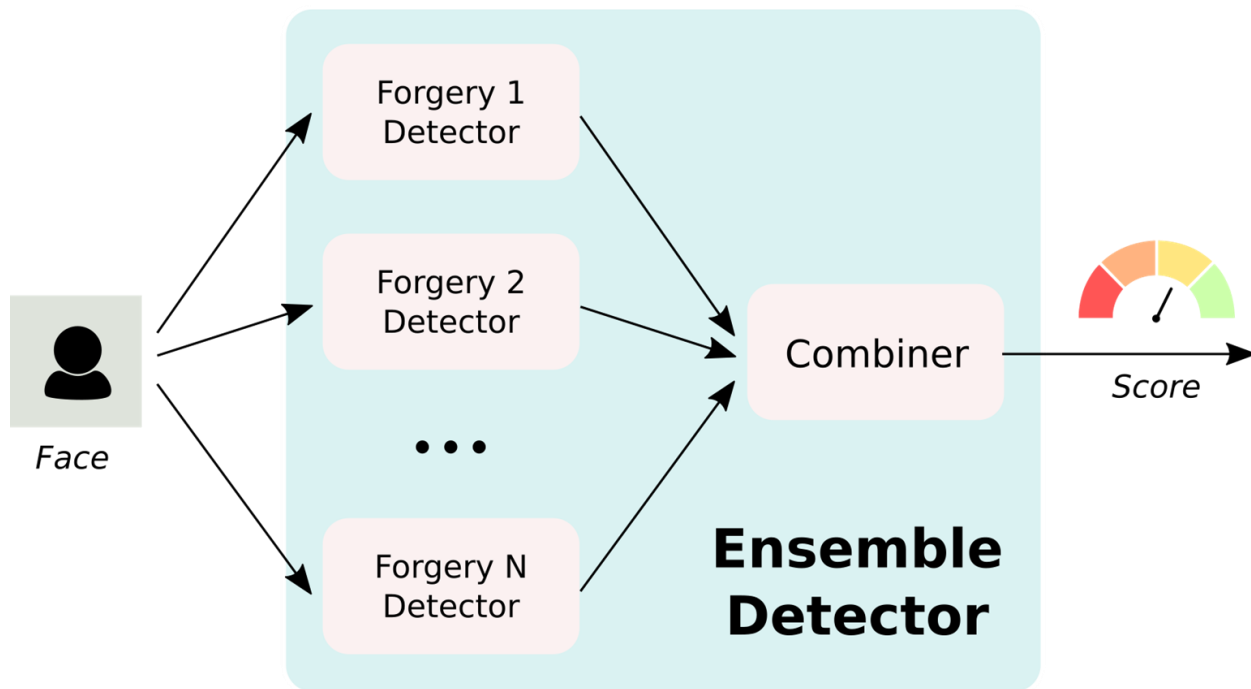


Stable Diffusion XL

# DeepFakeChain overview



# Scientific method



Ensemble of models specializing in different forgeries

- Extensibility
- Generalization
- Explainability
- Reliability

*Giatsoglou, N., Papadopoulos, S. and Kompatsiaris, I. (2023). Investigation of ensemble methods for the detection of deepfake face manipulations. arXiv preprint arXiv:2304.07395*

# Demonstration

# Relevance to TruBlo

## Criterion #1



### Detecting deepfake videos on the Russia-Ukraine war

2022-03-21 12:30:49

Luca Maiano

Validate Hash ?

Validated

Overall DeepFake confidence: 89%

Highest DeepFake confidence: 94%

Media file hash: 4660b85bc085432b108d142458a6fdca4a19b8c9cf013ea60a3b9f0726eca7c9

0x9fd9889b579e5aed2a4c7c11ff7cfe27607129a3c494c3a1306842b0ba70513b

Deepfakes have become a real threat. In recent days, fake videos have been released targeting #Zelensky and #Putin and the ...



Collaborative evaluation of multimedia  
beyond fact checking

Public trustworthy evaluations

**Check**



**Reality Defender<sup>®</sup>**

***Combining SotA deepfake detection and blockchain  
to create an easy and trustworthy process for  
recording expert opinions!***

# Originality & Innovation

Criterion #2



# Extensibility

## Criterion #11



Content moderation of social media platforms  
projected to 20B € in the next decade!

# The team



Dr. Symeon Papadopoulos  
Senior Researcher  
MeVer, CERTH



Nikos Giatsoglou  
Researcher  
MeVer, CERTH

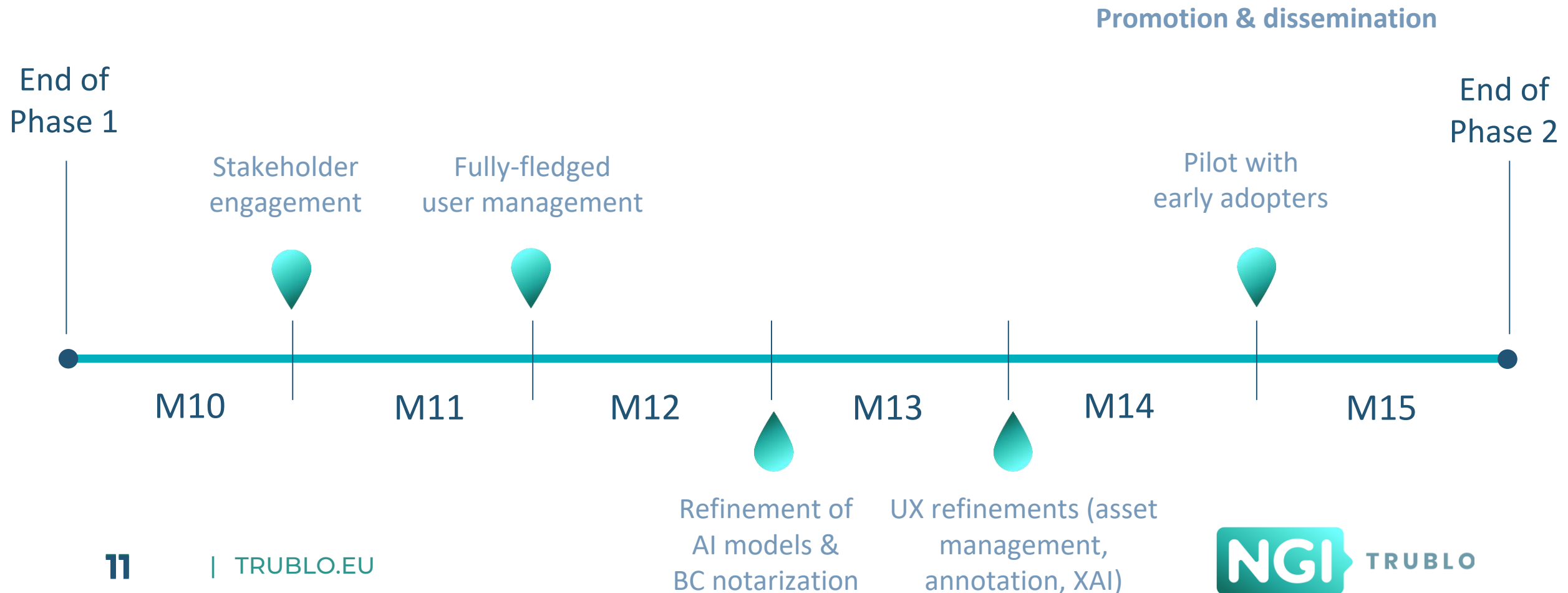


Dora Kallipolitou  
Delivery Manager  
Zelus



Stella Markopoulou  
Software Engineer  
Zelus

# Phase 2 roadmap



# Thank you!



<https://www.trublo.eu/deepfakechain/>



<https://www.certh.gr/>



<https://www.zelus.gr/>



Zelus

