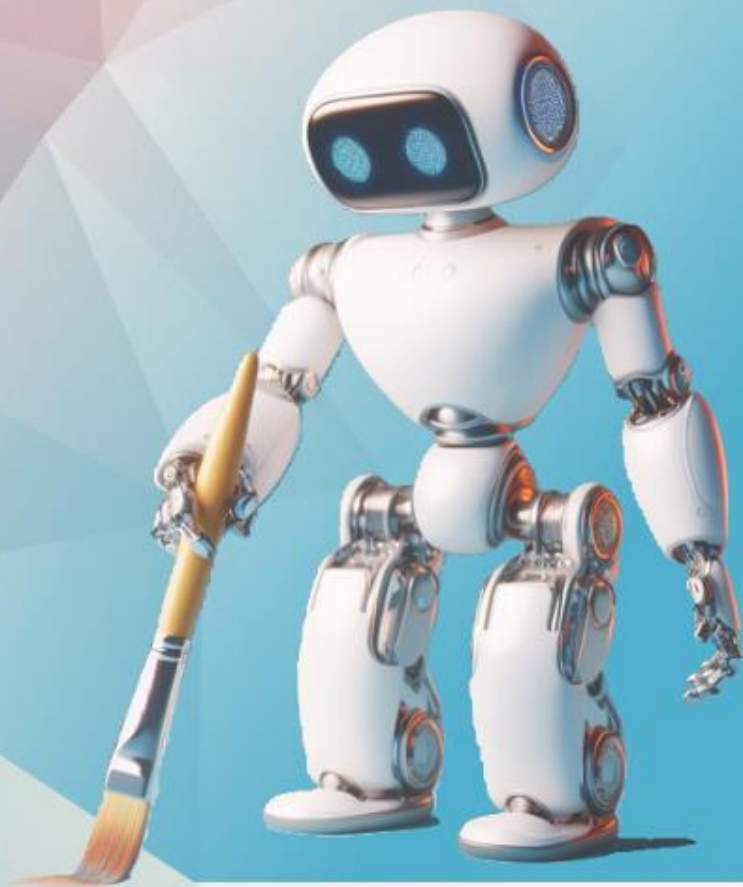


AI, CULTURAL HERITAGE AND ART

Between Research and Creativity



Workshop Proceedings
9-10 February, 2024

Edited by
Antonella Guidazzoli
Maria Chiara Liguori

AI, Cultural Heritage, and Art
Between Research and Creativity
Workshop proceedings – February 9-10, 2024

Edited by Antonella Guidazzoli and Maria Chiara Liguori

Cineca

AI, Cultural Heritage, and Art. Between Research and Creativity

Workshop proceedings – February 9-10, 2024

Edited by Antonella Guidazzoli and Maria Chiara Liguori

@ Cineca, 2024

www.cineca.it

visitlab@cinca.it

Copyright

I contenuti di questa pubblicazione sono protetti ai sensi della normativa in tema di opere dell'ingegno.

I diritti di questa pubblicazione appartengono a Cineca e ai rispettivi autori.

È consentito utilizzare, copiare e distribuire la presente pubblicazione e le sue parti senza fini di lucro.

Le note di copyright, gli autori ove indicati o la fonte stessa devono in tutti i casi essere citati nelle pubblicazioni in qualunque forma realizzate e diffuse.

Qualsiasi forma di link al presente sito, se inserita da soggetti terzi, non deve recare danno all'immagine ed alle attività di Cineca e degli autori della presente pubblicazione.

È vietato il cosiddetto deep linking ossia l'utilizzo non trasparente, su siti di soggetti terzi.

L'eventuale inosservanza delle presenti disposizioni, salvo esplicita autorizzazione scritta, sarà perseguita nelle competenti sedi giudiziarie civili e penali.

ISBN – 9788886037488

DOI: 10.1388/IIWORKSHOPAIBC

SUMMARY

Summary	p. 5
FOREWORD	
Antonella Guidazzoli , Maria Chiara Liguori	p. 7
AI, ART AND CULTURAL HERITAGE. BETWEEN IMAGE RECOGNITION, CONTEMPORARY ART, AND THE JURIDICAL PERSPECTIVE	
Alberto Salarelli, Francesca Zanella	p. 9
FROM EXPERIENCE COMPLEXITY ARISES	
Pierdomenico Memeo	p. 13
A RESEARCH METHODOLOGY FOR THE EVALUATION OF OPEN SCIENCE SOLUTIONS FOR CULTURAL HERITAGE	
Antonella Guidazzoli, Simona Caraceni, Rossella Pansini, Maria Chiara Liguori	p. 15
NEW PERSPECTIVES IN MANAGING HERITAGE DOCUMENTS	
Silvano Imboden, Giorgia Cardano, Corrado Consiglio	p. 29
ADVANCED SOLUTIONS FOR AUTOMATIC IMAGE AND TEXT ANALYSIS AND ENRICHMENT	
Donatella Sforzini, Gabriele Fatigati, Luca Mattei	p. 53
°°KOBİ: A KNOWLEDGE ECOSYSTEM FOR RESEARCH AND EDUCATION	
Franco Ripa di Meana, Alberto Giretti, Andrea Guidi, Massimo Vaccarini, Matteo Zambelli	p. 69
COPYRIGHT AND WORKS CREATED BY ARTIFICIAL INTELLIGENCE, TECHNICAL TRIALS OF PROTECTION: APPROACHES AND POSSIBLE SOLUTIONS	
Valentina Gastaldo	p. 85
L'ARTE ALGORITMICA DEI ROBOT	
Paolo Gallina	p. 103
THE ENCHANTMENT OF TRUTH: A CALL FOR ARTISTS TO TACKLE CONSPIRACY WITH WONDER	
Sofia Marasca, Federico Bomba	p. 121
OperAI RECORDS – ITS EXPERIENCE WITH AI, MUSIC AND NFTS'	
Remo De Vico, Niccolò Clemente	p. 129
DESTINATION EARTH. THE OCEAN BREATH	
Salomé Bazin	p. 137
TALES FROM THE RECEDING EDGE	
Calin Segal	p. 147

FOREWORD

Antonella Guidazzoli, Maria Chiara Liguori – VisitLab, Cineca

With this third workshop, organised by VisitLab Cineca, we continue our exploration into the possibilities offered by artificial intelligence, particularly in the field of digital humanities and the arts.

Through VisitLab, Cineca has supported cultural heritage, with what for a long time were considered non-traditional applications, and has always made available computing, computer graphics and extended reality resources in order to help researchers with their work and to communicate the complexity of their research.

Today, this activity is even more important.

Artificial intelligence can play a significant role in the relationship between Science and Art. In other words, between our collective intelligence, scientific knowledge and our imagination.

In 1987, a department dedicated to the study of artificial intelligence and neural network technologies already existed at Cineca, but the time was not yet ripe. Back then, we did not yet have enough digital data and computing power to train the networks.

Now, however, both advanced computing resources, such as Leonardo, and more and more digital resources are available. Thanks to the National Recovery and Resilience Plan, for example, the availability of digital content will be greatly increased, fuelling new research, new digital cultures and new artistic and creative productions, as well as stimulating real investments.

The workshop contributions collected in this volume ponder how AI can help researchers analyse documents of various kinds on a large scale, enabling them to ask new questions and find new connections between data and concepts. LLM, for instance, can contribute to a deeper knowledge in texts, acting as a complement to human input rather than its replacement. The same thing happens in the relationship between artist and artificial intelligence. The ideal author, therefore, will not be the AI, but the person who uses it competently. However, we will also read about the difficulties that arise in balancing neural network training and the defence of copyright

over images and intellectual production in general. This relevant session of the workshop has been coordinated by Professor Francesca Zanella, University of Modena and Reggio Emilia, and Professor Alberto Salarelli, University of Parma.

Finally, we will see the outcome of the complex interaction between art and science through the works of the three Emilia-Romagna resident artists of the GRIN - S+T+ARTS project.

This third workshop is just one step before the next, in the long train of considerations about a relationship that will never leave us from now on.

AI, ART AND CULTURAL HERITAGE. BETWEEN IMAGE RECOGNITION, CONTEMPORARY ART, AND THE JURIDICAL PERSPECTIVE

Alberto Salarelli - Università di Parma, Francesca Zanella - Università di Modena e Reggio Emilia

The University of Parma, specifically through its Centre [CSAC \(Study Centre and Communication Archive\)](#), started a fruitful collaboration with VisitLab Cineca since 2016, a collaboration that is still going on with the PhD Course Philological-literary, historical-philosophical and artistic sciences of the University of Parma. At that time our challenge was the valorization of visual and design archives through a multiplicity of actions: studies on design histories through the archives, cataloguing and digitization of archival fonds, valorization and dissemination through exhibition activities and digital environments. We verified that through the project *Sottsass Virtual Exhibitor*, a project started on 2016 which keywords were Accessibility, Narratives and Digital Archives, with a focus on the enhancement of cultural heritage, tangible and intangible, through the research, the cataloguing, the digitizing and the design of a digital environment. Thanks to the 3D representation of the 2017 exhibition *Ettore Sottsass. Oltre il design*, held on the Cistercian church of the Valsarena Abbey, we defined with Visitlab a gateway to the University of Parma online catalogue and to external resources.

Since then, the horizon of possibilities and challenges on the field of Digital Humanities, Cultural Heritage and Digital Technologies has drastically changed. Nowadays Culture Heritage can be valorized, but we can also say enhanced, through a multiplicity of technologies and digital instruments. As Pierdomenico Memeo wrote in the proceedings of the first workshop organized by VisitLab Cineca, *AI and Cultural Heritage. Between Research and Creativity*, today we are facing a revolution due to two factors: “systems such as neural networks, fuzzy control systems, and tools for mathematical optimization; on the other, the tremendous amount of data created and accumulated by the progressive expansion of digital systems. The former provided the seeds for this new paradigm; the latter, the soil in which it could grow” [1, p. 10].

That’s why the topics of the Digital Humanities seminar for the PhD Course at the University of Parma, are projects with which researchers and designers verify how to work with a great amount of data, digitized collections, digital artworks and consequently new kind of digital archives, or on image recognition (as in the speech of Lorenzo Baraldi, *Multimodal and self-*

supervised AI for Digital Humanities, for the project ArchiNervi, a collaboration between University of Parma, University of Modena and Reggio Emilia and Fondazione MAXXI) [2].

If we look at the papers presented at the prior AICH workshops, we can read interesting and stimulating interventions on creativity and AI, on one hand, and on the research and production of new knowledge in the field of archaeology, or history of arts.

The questions posed are how the definition of art and the creativity process are changing; how artists experiment with the new neural networks, or with a big amount of data differently produced and stored.

For the third AICH workshop we decided to face another problem that is strongly related to a crucial question in the artistic debate since the beginning of the XX century: the role of the artist in relation to the changing definition of art. We decided to face that question from the juridical perspective, somehow assuming an historiographical proposal of Denis Riout [3], who identified as a starting point for Modernity the 26th of November 1928 sentence against the tax applied to Constantin Brancusi statues, that the New York douane agents didn't recognized as works of art. The sentence was based on the recognition of the changing nature of contemporary art; now we need to understand the juridical perspective related to the presence of a non-human entity, too. If we assign a juridical status to robots and AI, have we to presume rights but even duties? Is it ethical to assign to a non-human subject rights and duties without its consent? [4].

All these questions, together with those related to NFT and blockchain techniques, are being addressed in this workshop by scientists from a research team at the University of Parma, with the aim of broadening the debate and opening up new research perspectives.

References

- [1] Memeo, P., *Introduction – The Greatest Enemy of Knowledge*, in AI and Cultural Heritage. Between Research and Creativity. Workshop proceedings, edited by Antonella Guidazzoli and Maria Chiara Liguori, Cineca, 2023.
- [2] Buda, C. Z., Zanella, F., *Pier Luigi Nervi: due fondi storici*, Milano, Electa, 2024.
- [3] Riout, D., *L'arte del ventesimo secolo: protagonisti, temi, correnti*, Torino, Einaudi, 2002.
- [4] Gastaldo, V., *L'arte algoritmica: il nuovo mondo che ci aspetta*, 2019, in “federalismi.i”, 30 giugno 2021

FROM EXPERIENCE COMPLEXITY ARISES

Pierdomenico Memeo - Science writer, outreach expert

«You're gonna need a bigger boat.»

Chief Martin Brody (Roy Scheider), *Jaws* (1975)

Never has it been clearer to me than in this third iteration of the workshop on "AI, Cultural Heritage and the Arts" that we will need many more perspectives and expertise to address the fundamental changes that artificial intelligence is bringing to our well-established world of cultural endeavour.

As we wade through the expanding applications of AI in a widening range of intellectual activities, and as we delve deeper into understanding its mathematical and technological underpinnings, we realise that we cannot confine ourselves to scientific debates and cultural discourses, but need to include in our conversation a much larger collection of actors from many different disciplines and circumstances.

For it is from experience that complexity emerges, and in order to address this complexity we have found ourselves looking at it from different experiences.

To better understand the inner workings of AI systems applied to cultural heritage, we again called on the expertise of mathematicians and programmers to explain their mechanics and machinery; to better understand how to use them to ignite our creativity, we asked artists to show us their inspirations and guide us through their process; to understand the impact of their use on art collections and material archives, we asked historians and critics to shed light on different media, from painted canvas to written parchment; to address pressing issues of rights and intellectual property, we invited legal experts and legislators to unravel the tangled web of conflicting interests.

On their own, each of the pieces you are about to read would have been a rewarding look at a fascinating topic in artificial intelligence and cultural heritage; but together they paint a sweeping view of a vast landscape that is both rapidly changing and constantly evolving.

To create this new perspective, we build on the previous workshops, and once again we move forward in the confidence that this is just one of many.

Again, my gratitude goes to all contributors for sharing their knowledge and giving us their unique perspectives on the intersection between artificial intelligence, cultural heritage and artistic expression. I am grateful also to Antonella Guidazzoli, Head of CINECA VisIT Lab, and Maria Chiara Liguori, digital historian, for inviting me again to host and chair the workshop, and for giving me the opportunity to do my part in the realization of these proceedings.

Now, with me— look forward.

A RESEARCH METHODOLOGY FOR THE EVALUATION OF OPEN SCIENCE SOLUTIONS FOR CULTURAL HERITAGE

Antonella Guidazzoli, Simona Caraceni, Rossella Pansini, Maria Chiara Liguori – VisiLab Cineca

Introduction

Among the fields of application of artificial intelligence tools, digital humanities constitute a particularly interesting field for experimentation, given the variety of our cultural heritage and its manifestations. Printed documents, manuscripts, paintings, archaeological finds are just some of the classes of objects that are part of our heritage, each with peculiarities to be valorised and made available to the public.

When dealing with such a rich variety of data, mapping the most advanced AI technologies in open science is a necessary activity to make the national heritage more accessible to all audiences, from specialized researchers to the general public.

An important aspect of these activities is the research on the state of the art, with the survey of the technical and scientific literature, with particular attention to the applications of AI technologies to the digital humanities, if available.

The immediate consequence of the census and analysis is the evaluation of the viability of the tools identified, which takes into account both the level of training of the models, their adaptability to culture-related contexts, and their usability within the High Performance Computing (HPC) infrastructure. Cineca in fact, thanks to its supercomputing machines, provides an ideal environment for training neural networks.

The research was mainly carried out by mining the main repositories for open-source tools: [Github](#) and [Hugging Face](#). Hugging Face, in particular, is one of the most popular platforms for learning and developing AI tools, with a large community of developers and researchers, and a collection of pre-trained models, datasets, tutorials and documentation.

Technology scouting must take into account the continuous progress related to AI technologies and is, therefore, carried out on an ongoing basis to ensure that technical documents are up-to-date and the most suitable tools are identified. The tools listed in the present text are those scouted up to February 2024.

Scouting activities

The areas of research taken into account are:

Object Detection, NER (Named Entity Recognition), OCR (Optical Character Recognition), HCR (Handwritten Character Recognition), Speech to Text, Text to Speech, Geocoding, Language Translation and AI Image Enhancement. These areas are described below to outline the area of activity and lay a common ground for understanding the technologies.

1) *Object Detection*

Object detection tools enable the detection of specific elements within a visual digital resource (images, video) and automatic metadata collection by means of AI. AI object detection models detect elements, delimiting them with bounding boxes and proposing a classification for them. The digital resources processed with Object Detection are then enriched with a series of tags that can be called up during the search.

Table n. 1 – A selection of Object detection tools

Name of the tool	Product site link	Brief description from web sites
CLIP	https://openai.com/research/clip	CLIP is a multimodal zero-shot model, which is able, given an image and textual descriptions, to predict the most relevant description for that image, without specific training.
X-CLIP	https://github.com/xuguohai/X-CLIP	X-CLIP is a multimodal zero-shot model, which given a video and text descriptions, is able to predict the most relevant description for that image, without specific training.
RegionCLIP	https://github.com/microsoft/RegionCLIP	RegionCLIP is an extension of CLIP that provides the spatial location of a textual description. This makes it possible to create Bounding Boxes.
OWL-ViT	https://github.com/google-research/scenic/tree/main/scenic/projects/owl_vit	Owl-Vit, starting with keywords, finds matches in images by drawing Bounding Boxes. No pre-training required.

vild	https://github.com/tensorflow/tpu/tree/master/models/official/detection/projects/vild	VILD is another extension of CLIP for locating the spatial location of a textual description.
Detectron2	https://ai.facebook.com/tools/detectron2/	Detectron 2 detects the classes it is trained for by drawing Bounding Boxes.
SegmentAnything	https://ai.meta.com/blog/segment-anything-foundation-model-image-segmentation/	SegmentAnything draws the contours of detected objects in an image (possibly also selected in bounding box).
YOLO v5	https://www.ultralytics.com/yolo	Deep learning framework, based on PyTorch, very popular for its versatility, ease of use and high performance
YOLO X	https://www.ultralytics.com/yolo	Further version of Yolo, with simpler design and better performance. The aim of the application is to bridge the gap between the research and industrial communities.

2) *NER - Named Entity Recognition*

NER (Named Entity Recognition) is a NLP (Natural Language Processing) method that enables the classification of text contents (words or sets of words) on the basis of categories, such as names of people, organizations and places. The recognition of categories, and thus of named entities, takes place by means of specially trained AI models and results in the semantic enrichment of digital resources with metadata that are searchable and relatable to each other.

Name of the tool	Product site link	Brief description from web sites
WikiNEuRal	https://github.com/Babelscape/wikineural	Model derived from the fine-tuning of a transformer (Bert-base-multilingual-cased) developed ad hoc for an NER task.
it_core_news_lg	https://spacy.io/models/it_core_news_lg	General purpose model for NLP tasks (including NER) based on the SpaCy library.

spacy-dbpedia-spotlight	https://spacy.io/universe/project/spacy-dbpedia-spotlight	Model that uses the SpaCy library to locate entities that are then searched within the Wikipedia DB.
Tint	https://dh.fbk.eu/research/tint/	Taking a text input via the dedicated site or locally using the prompt and indicating the task (NER), a txt file containing the predictions is generated. Based on Stanford CoreNLP.

Particularly in this field, we would like to mention [I-Media-Cities](#), a Horizon 2020 project, involving 9 of the major European film libraries, defined by the [EU Innovation Radar](#) as "market ready", which was also tested and used in other projects, such as the urban redevelopment of Ravenna's Dockyard (Project DARE, UIA, Urban Innovative Actions of the European Community). Never has it been clearer to me than in this third iteration of the workshop on "AI, Cultural Heritage and the Arts" that we will need many more perspectives and expertise to address the fundamental changes that artificial intelligence is bringing to our well-established world of cultural endeavour.

In order to be able to find elements within a text, it is necessary to read it from beginning to end. The same thing happens with old film footage, which has to be seen from beginning to end in order to be able to research or locate objects. Special technologies were applied in I-Media-Cities, in collaboration with the Fraunhofer Institute of Technology, that made it possible to run digitized and normalized film footage on the database with regard to the frame-per-second characteristic. In addition to this, there is another innovative operation that we carried out from the beginning in the I-Media Cities project that directly influenced the object recognition process, which was to create a 'reverse story-boarding' that made it possible to subdivide the footage into small portions that could be likened to shots, instead of the canonical scenes that screenwriters work on. This allowed the algorithm to work with more manageable portions than entire audiovisuals, in terms of length and complexity.

The Fraunhofer algorithm made it possible to detect various elements such as people, animals, means of transport, objects and various other elements. These, as can be seen in the image reproducing the interface (fig.1), have been 'tagged' as automatically detected, and marked by a unique number that refers to the same object within the film, and identified by a color (red, yellow or green) that, like a traffic light, indicates the degree of confidence of the algorithm in having correctly detected what is shown.

Ocropus	https://github.com/ocropus-us-archive/DUP-ocropy	OCROPUS is an OCR system that combines classical OCR approaches with new machine-learning-based methods. It includes a complete OCR workflow from pre-processing to analysis and layout recognition and can be used with Tesseract and other OCR engines.
Kraken	https://kraken.re/main/index.html	Kraken is an OCR engine that uses deep learning techniques designed to handle complex OCR tasks, such as text recognition in historical documents. It can be trained on new fonts and styles and supports multiple languages.
Calamari	https://github.com/Calamari-OCR/calamari	OCR engine based on OCROPUS and Kraken using python3. It is designed to be easy to use from the command line but also to be modular in order to be integrated and customized by other Python scripts.

4) HCR - Handwritten Character Recognition

HCR (Handwritten Character Recognition) tools enable automatic transcription of handwritten and typewritten texts, both alphabetic and numeric, using specially trained models.

Name of the tool	Product site link	Brief description from web sites
-------------------------	--------------------------	---

Transkribus	https://readcoop.eu/it/transkribus	Transkribus is an AI-based platform for text recognition, transcription and research of historical documents, from anywhere, anytime and in any language.
TrOCR	https://github.com/microsoft/unilm/tree/master/trocr	TrOCR is an end-to-end text recognition approach with pre-trained image and text Transformer models, which exploits the Transformer architecture for both image understanding and text generation.

5) *Speech to Text*

Speech to Text tools enable the automatic and simultaneous transcription of human speech into written text. These tools are now increasingly used and common in our daily lives, e.g. in subtitling videos.

Name of the tool	Product site link	Brief description from web sites
Pyannote	https://github.com/pyannote/pyannote-audio	It is an open-source toolkit written in Python for speech diarization. Based on the PyTorch machine learning framework, it comes with state-of-the-art pre-trained models and pipelines, which can be further optimized for personal data for even better performance. Speaker Diarization: splits audio based on people speaking
Whisper	https://openai.com/research/whisper	Automatic speech recognition system trained on 680,000 hours of supervised multilingual and multitasking data collected from the web.

SpeechBrain	https://speechbrain.github.io/	Open Source conversational artificial intelligence toolkit. Designed to be simple, flexible and well documented. It covers various technologies to deal with audio and achieves competitive performance in various fields.
Whisper speaker diarization	https://huggingface.co/spaces/vumichien/Whisper-speaker-diarization	Combination of Whisper and SpeechBrain, useful for using OpenAIs Whisper to transcribe and diarise audio files.

6) Text to Speech

Text to Speech (TTS) tools enable the artificial reproduction of the human voice for reading a written text.

Name of the tool	Product site link	Brief description from web sites
OpenTTS	https://github.com/synesthesiam/opentts	Tool that unifies access to multiple open source speech synthesis systems and voices for different languages. Supports a subset of SSML that can use multiple voices, speech synthesis systems and languages.
Whisperspeech	https://github.com/collabora/WhisperSpeech	An Open Source speech synthesis system built by inverting Whisper, formerly known as spear-tts-pytorch. The aim of the model is for it to be similar to Stable Diffusion, but for speech, powerful and easily customisable. Currently only used with appropriately licensed voice recordings. The code is Open Source.
Glow TTS	https://github.com/revsic/tf-glow-tts	Generative flow for speech synthesis via monotone alignment search.

Deep Voice 3	https://github.com/r9y9/deepvoice3_pytorch	PyTorch implementation of speech synthesis models based on convolutional neural networks.
Tacotron 2	https://github.com/NVIDIA/tacotron2	PyTorch implementation, which reports faster and more real time inference than the former.
WaveGlow	https://github.com/NVIDIA/waveglow	Flow-based network capable of generating high-quality speech from mel spectrograms. WaveGlow combines the information from Glow and WaveNet to provide fast, efficient and high-quality audio synthesis without the need for auto-regression. WaveGlow is implemented using a single network, trained via a single cost function to maximize the verisimilitude of the training data.
WaveNet	https://deepmind.google/technologies/wavenet/	Google Wavenet Text to Speech is a powerful and advanced text-to-speech (TTS) system developed by Google's DeepMind. It uses state-of-the-art machine learning and deep learning algorithms to synthesize high-quality, natural-sounding speech from text input in audio files.
Parallel WaveNet	https://github.com/kensun0/Parallel-Wavenet?tab=readme-ov-file	Wavenet implementation with other models
Parallel Tacotron 2	https://github.com/keonlee9420/Parallel-Tacotron2	PyTorch implementation of Google's Parallel Tacotron 2: a non-self-regressive neural TTS model with differentiated duration modeling.
WaveGrad 2	https://wavegrad.github.io/v2/	Google Brain WaveGrad 2 is an end-to-end non-autoregressive generative model for speech synthesis trained to estimate data density gradients.

ItoWave	https://wushoule.github.io/ItoAudio/	Vocoder based on a pair of forward- and reverse-time stochastic differential equations (SDE).
DiffGAN-TTS	https://github.com/keonlee9420/DiffGAN-TTS?tab=readme-ov-file	Efficient, high-fidelity voice synthesis with GAN denoising
SUNO-AI-BARK	https://github.com/suno-ai/bark	Bark is a transformer-based text-to-audio model created by Suno. Bark can generate highly realistic multilingual speech and other types of audio, including music, background noise and simple sound effects. The model can also produce non-verbal communications such as laughter, sighing and crying.
VALL-E-X	https://github.com/Plachtaa/VALL-E-X	An open source implementation of Microsoft's VALL-E X zero-shot TTS model.
StyleTTS2	https://github.com/yl4579/StyleTTS2	Text-to-speech (TTS) model that exploits style diffusion and adversarial training with large speech language models (SLM) to achieve human-level TTS synthesis
coqui TTS	https://github.com/coqui-ai/TTS	Library for the generation of advanced text-to-speech, with pre-trained models for more than 1100 languages; tools for training new models and perfecting existing models in all languages; tools for analyzing and processing datasets.
Tortoise-TTS	https://github.com/neonbjb/tortoise-tts	Tortoise is a text-to-speech tool built with the following priorities: <ul style="list-style-type: none"> -Strong multiple speech synthesis capabilities. -Highly realistic pronunciation and intonation.

7) Geocoding

Geocoding is the function that allows geographical information (in particular latitude and longitude coordinates) to be extracted from the textual description of a location, such as an address or place name.

Name of the tool	Product site link	Brief description from web sites
Google Maps	https://developers.google.com/maps/documentation/geocoding/overview?hl=it	An application that among other functions is useful for finding and reaching the desired location in just a few clicks, shows and analyzes geospatial data on a map to help the user make decisions, identifies places in the vicinity to simplify everyday decisions.
Map Box	https://docs.mapbox.com/api/search/geocoding/	Application for both forward geocoding, i.e. converting locations in text into geographical coordinates, and reverse geocoding, i.e. converting geographical coordinates into locations, names.
OpenStreetMap	https://wiki.openstreetmap.org/wiki/OSMPythonTools	Tool built with the interaction of several users (community of mappers) that helps to keep up-to-date data about roads, trains, cafes, railway stations and much more for the whole world.
Nominatim	https://nominatim.org/	Tool that uses data from OpenStreetMap to find geographical locations from addresses or cities (geocoding). It can also be used to find addresses from geographical coordinates.
Geopy	https://pypi.org/project/geopy/	Geopy allows Python developers to easily locate the coordinates of addresses, cities, countries and landmarks around the world using third-party geocoders and other data sources. The tool includes geocoding classes for OpenStreetMap Nominatim, Google Geocoding API (V3) and many other geocoding services.

8) Language Translation

Language Translation tools enable AI-based translation of textual metadata of a digital resource from one language to another.

Name of the tool	Product site link	Brief description from web sites
OpenNMT	https://opennmt.net/ https://github.com/OpenNMT	OpenNMT is an open source ecosystem for neural machine translation and neural sequence learning.
LibreTranslate	https://libretranslate.com/	API translation tool, free and Open Source. Self-hosted, offline and easy to configure. Unlike other APIs, it does not rely on proprietary providers such as Google or Azure to perform translations. Its translation engine is powered by the open source library Argos Translate.
OPUS MT	https://opus.nlpl.eu/	OPUS is an ever-growing collection of web-translated texts. The OPUS project seeks to convert and align free online data, add language annotations and provide the community with a publicly available parallel corpus. OPUS is based on open source products and the corpus is also provided as an open content package.

9) AI Image Enhancement

Image enhancement tools enable the improvement in resolution of digital images. Image enhancement is often preliminary to the application of other AI tools, such as Object Detection, OCR and HCR.

Name of the tool	Product site link	Brief description
Stable diffusion	https://stability.ai/	Multi-purpose CV model based on diffusion models
Mmagic	https://github.com/open-mmlab/mmagic	multi-purpose zoo model including image super resolution using both GAN and diffusion models

The models identified by the analysis of the technical/scientific literature in the field of open science reveal a very rich and varied panorama, potentially capable of satisfying sorting and classification needs. The assessment of the immediate operability of the tools cannot, however, disregard the awareness of the specificity of the field of application. Hardly a general-purpose model for Object Detection, NER or HCR will be able to interpret and return accurate metadata on cultural heritage, which includes within its various expressions, each with its own peculiar characteristics.

Conclusions

The work done so far allows us to reflect on the importance of adapting open source engines to the specific needs of cultural heritage and on the need to organize working environments that allow the creation of customized datasets, responding to the needs of different profiles, such as archivists and researchers in the field of digital humanities (DH) in Human - Ai collaboration environments.

The application of technological workflows, combining digitisation and artificial intelligence tools, aims to make the management and interrogation of cultural content much faster and more efficient.

The time needed by the scholar to evaluate the source, irrespective of whether it is iconographic or textual, is reduced to the essentials of examining what is really pertinent.

This genuinely computerized breakthrough finally makes it possible to strengthen research with quantitative studies that were previously only feasible at the cost of very long execution times and that, in any case, did not guarantee the collection and analysis of what was actually useful for furthering the research topic.

The leap enabled now by the application of these tools, therefore, guarantees a drastic improvement. In the days of the I-Media-Cities project, even the application of an algorithm capable of automatically segmenting audiovisual content into shots had already allowed tagging at a higher level of granularity.

During research on public transport in an urban context, for instance, the availability of the tag 'tram', referring to the precise shot where a tram is shown, drastically reduces the time needed in order to find pertinent content.

Similarly, by applying a pipeline such as the one hypothesized for newspapers and presented in these Proceedings by Silvano Imboden, Giulia Cardano and Corrado Consiglio, should a researcher wish to carry out an analysis on print advertisements, it would be possible to quickly isolate the desired material. A further drastic cut in time required would be prompted by the presence of recognisable text to be submitted to OCR systems.

This would allow the automatic selection of advertising campaigns by content type, as well as the possibility of collecting data on the size of advertisements and their distribution in the newspaper, both within the page and in the newspaper as a whole.

It is evident, therefore, how the well-controlled application of the right tools and the most appropriate pipelines, depending on the different types of content, will enable the humanities to progress, deepening and broadening its research, at an unprecedented rate.

NEW PERSPECTIVES IN MANAGING HERITAGE DOCUMENTS

Silvano Imboden, VisitLab - CINECA; Giorgia Cardano, Corrado Consiglio – UNIBO

Introduction¹

Italy is home to a wealth of institutions, both public and private, which are the custodians of an immense collection of documents. These records date from ancient times, when scribes meticulously wrote texts by hand, to the invention of the printing press, which revolutionised access to books and newspapers. The nature of the content varies widely, including legal documents, notarial acts, correspondence between nobles, as well as literature, music, poetry and information. This vast "big data" encapsulates our past, and it is the responsibility of the community to preserve and make the best use of it.

With the advent of digital technology and scanning, many of these documents have been digitised and made available online, leading to significant improvements in both preservation and accessibility. After scanning, the logical next step is to convert images into text using Optical Character Recognition (OCR) techniques.

The benefits of converting images into structured information are many. It opens possibilities for further processing, such as Named Entity Recognition (NER), allows documents to be searched by text or keywords, and facilitates data mining for statistical insights. Text-to-speech technologies can enable visually impaired people to 'listen' to documents, while Large Language Models (LLMs) can generate derivative documents such as summaries or translations.

However, a significant share of these documents has yet to be converted into text. While there are effective OCR tools, such as [Tesseract](#), that can convert printed text in different languages, these tools struggle with more complex documents, such as manuscripts, scores and periodicals.

Fortunately, advances in AI technologies are bringing new approaches to overcome these challenges. The present contribution introduces readers to the results of two studies carried out by Cineca on these innovative tools.

The first analysis focuses on the processing of newspaper and magazine images. Before proceeding with OCR, it is essential to recognise and split the page into its constituent elements:

¹ The following section has been written by Silvano Imboden.

titles, paragraphs, images and captions. This analysis must also reveal the hierarchical structure of the content, indicating, for example, that a caption refers to a particular image that is part of an article with a specific title. This process, known as "layout analysis", uses a tool called LayoutParser. It is an open-source tool that comes with several models tailored to specific use cases.

The results indicate that for certain document types, such as scientific publications, existing models already yield good quality results. However, for more complex layouts, like those found in newspapers, fine-tuning of models is necessary to address specific cases. In conclusion, we can say that layout analysis is a problem that is partially solved and continues to evolve.

The second study is in handwritten text recognition (HCR). We experimented with [TrOCR](#), an open-source tool from Microsoft based on the latest transformer architectures. Our aim was to use it to process the corpus of letters of Isabella d'Este, generously provided by the Mantua Archive and previously digitised in collaboration with the [IDEA project](#) (Isabella d'Este Archive, directed by Professor Deanna Shemek of the University of California, Irvine).

During this experiment, it became clear that the models provided with TrOCR, trained on modern English texts, were not suitable for our corpus and required specialization. Fine-tuning a model involves creating pairs of question/answer inputs, where each pair consists of an image of the text and its corresponding transcription. This data is then used to train the model starting from its previous version.

However, while the work required is substantial, preparing a few hundred documents can enable a model to process thousands. In the case of Isabella's letters, which comprise about 37,000 images, we achieved a commendable 87% in character recognition accuracy, which can be further improved with subsequent iterations.

In summary, thanks to modern AI solutions, many obstacles that have hindered the complete digitization of our cultural heritage are being addressed. The automated processing of large corpora will soon be feasible, making this data highly valuable for the entire community.

Research on analysis and layout detection of documentary resources leveraging open-source artificial intelligence software²

As part of a national endeavour to digitize the Italian cultural heritage and create a [digital library](#); thanks to an action promoted by the Italian Ministry of Culture and financed through [PNRR funds](#), Cineca was committed, along with other tasks, to the technical scouting activity. In particular, the task aims to look for and test open - source software capable of exploiting artificial intelligence models to perform specific actions on images of texts of different nature (printed texts, manuscripts, newspapers, microfilm, musical scores etc...). These tasks mainly involved character recognition activity of printed texts (OCR), i.e., scanning of images with a single area of text; character recognition of handwriting in manuscripts (HCR) through a detection of individual lines of text in the image; layout parsing, i.e., analysis of images of texts to be divided into several areas; NER (Named Entity Recognition) tasks to identify words and/or phrases and classify them.

In the next few paragraphs we are focusing specifically on research and testing of software features that can recognize the layout of text images using artificial intelligence models.

Layout parsing and recognition of documentary resources

The activity of layout parsing applied to document images consists in identifying and subdividing images reproducing formatted pages of texts, defining the different areas that compose the structure and arrangement of the elements of the document under analysis. This operation can be applied to resources of a varied nature: from the most common ones, such as the pages of a book, a newspaper or a periodical magazine, but also scientific publications (legal, medical etc...), historical documents, manuscripts, to documents that have an articulated structure, such as tables, receipts/invoices, musical scores, maps (Fig. 1).

Through a layout parsing process it is possible to locate on a page of a book the areas that identify the title of a chapter, paragraphs and sub-paragraphs, any headings and subheadings, images, tables, etc.

The tools must therefore recognize certain areas on the page, such as the header (name of the newspaper), titles and subtitles of articles, any image and the paragraph of the articles. Once these areas had been identified, it would then be possible to apply OCR and, thus, extract the textual content within the layout area.

² The following section has been written by Corrado Consiglio.



Fig. 1 - Samples of documents on which layout parsing can be performed.

The first part of the research focused on a scouting activity aimed at finding potential open-source software with related documentation on the network that could meet the Ministry of Culture's requirements. In this regard, the two main repositories consulted were [GitHub](#) and [Hugging Face](#). After careful research, the software chosen to achieve the goals set was [Layout Parser](#).

Scouting activity: Layout Parser software

The technology scouting activity highlighted that, at the moment, there is not a large diffusion of open-source tools equipped with artificial intelligence models that can effectively perform layout parsing of documents. The only tool identified capable of performing this task properly is the aforementioned Layout Parser.



Fig. 2 – Layout Parser applied to different type of printed contents.

Layout Parser is a software with a rich repository of deep learning models for layout detection and a set of unified and optimized APIs for their use. Due to this, layout analysis and detection of document images can be performed in only a few lines of code.

Layout Parser, through its APIs, provides several functions to perform layout recognition tasks to the fullest:

- Selecting layout/textual elements in the left column of a page.
- Performing OCR for each detected Layout Region.
- Flexible APIs for visualizing the detected layouts.
- Loading layout data stored in Json, csv, and even PDFs.

Furthermore, the tool comes with an [open platform](#) for sharing pre-trained layout detection models and pipelines for conducting image analysis.

After studying the related documentation, the choice of Layout Parser seemed the most suitable thanks to its rich repository of pre-trained and ready-to-use artificial intelligence models, an aspect that was in line with Ministry requirements as well. Once the validity of the tool was approved, the next step was to better understand what models were pre-trained and integrated into Layout Parser and how they work. Those models are presented in Fig. 3.

Model Catalog

Dataset	Model	Config Path
HJDataset	faster_rcnn_R_50_FPN_3x	lp://HJDataset/faster_rcnn_R_50
HJDataset	mask_rcnn_R_50_FPN_3x	lp://HJDataset/mask_rcnn_R_50C
HJDataset	retinanet_R_50_FPN_3x	lp://HJDataset/retinanet_R_50_
PubLayNet	faster_rcnn_R_50_FPN_3x	lp://PubLayNet/faster_rcnn_R_5
PubLayNet	mask_rcnn_R_50_FPN_3x	lp://PubLayNet/mask_rcnn_R_50
PubLayNet	mask_rcnn_X_101_32x8d_FPN_3x	lp://PubLayNet/mask_rcnn_X_10
PrimaLayout	mask_rcnn_R_50_FPN_3x	lp://PrimaLayout/mask_rcnn_R_
NewspaperNavigator	faster_rcnn_R_50_FPN_3x	lp://NewspaperNavigator/faster
TableBank	faster_rcnn_R_50_FPN_3x	lp://TableBank/faster_rcnn_R_50
TableBank	faster_rcnn_R_101_FPN_3x	lp://TableBank/faster_rcnn_R_10
Math Formula Detection(MFD)	faster_rcnn_R_50_FPN_3x	lp://MFD/faster_rcnn_R_50_FPN

Fig. 3 - Official Model Catalog of Layout Parser

As shown in the table, the available models are the *faster_rcnn*, *mask_rcnn* and *retinanet*. The *faster_rcnn* model and the more advanced *mask_rcnn* essentially are two neural networks

specialized in image recognition [1], while the *retinanet* model represents a benchmark in object detection because of its advanced features [2].

All three models were trained on several datasets to recognize the layouts of different types of documentary resources. Three datasets attracted our interest in the early stages of the research: HJDataset, PubLayNet, and TableBank (Fig. 4).

Layout Parser Main Datasets

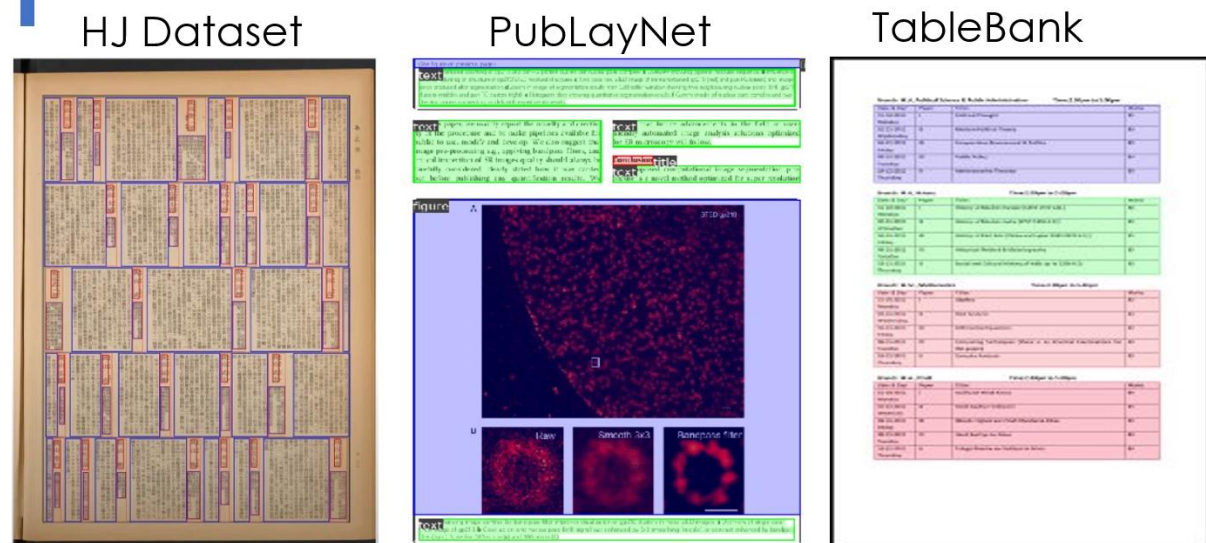


Fig. 4 - Three example images of main datasets in Layout Parser

HJDataset is a large collection of historical texts of Japanese culture characterized by a complex layout structure. There are in fact more than 250,000 layout elements identified through seven different types of annotations [3].

PubLayNet is a dataset for document layout analysis by automatically matching the XML representations and the content of over 1 million PDF articles that are publicly available on PubMed Central. The size of the dataset is comparable to established computer vision datasets, containing over 360 thousand document images, where typical document layout elements are annotated [4].

TableBank is a new image-based table detection and recognition dataset containing 417K high quality labeled tables and built with novel weak supervision from Word and Latex documents on the internet [5].

According to what the documentation illustrates, the training of the models was mainly focused on these three datasets; nevertheless, there are also three other datasets that were used for

training and hence they have been studied to have a comprehensive understanding of what the tool could globally offer. These additional datasets are presented below:

PrimaLayout: This dataset has been created primarily for the evaluation of layout analysis (physical and logical) methods. It contains realistic documents with a wide variety of layouts, reflecting the various challenges in layout analysis. Particular emphasis is placed on magazines and technical/scientific publications which are likely to be the focus of digitisation efforts [6].

Newspaper Navigator: This dataset consists of extracted visual content for 16,358,041 historic newspaper pages in [Chronicling America](#). The visual content was identified using an object detection model trained on annotations of World War 1-era *Chronicling America* pages, including annotations made by volunteers as part of the [Beyond Words](#) crowdsourcing project [7]. The resulting visual content recognition model detects the following types of visual content:

- Photograph
- Illustration
- Map
- Comics/Cartoon
- Editorial Cartoon
- Headline
- Advertisement

Math Formula Detection (MFD): FormulaNet is a new large-scale Mathematical Formula Detection dataset. It consists of 46'672 pages of STEM documents from [arXiv](#) and has 13 types of labels. The dataset is split into a train set of 44'338 pages and a validation set of 2'334 pages [8].

Once the main characteristics of each dataset were understood, the next step was to try the functioning of these models on a test dataset, i.e., a collection of newspapers, to evaluate the ability to recognize the layout of this type of documents as required by the MIC. To accomplish this task, while waiting for the datasets provided by the Ministry, a specific dataset was created *ad hoc* based on the newspapers digitised by the [Storia e Memoria di Bologna](#) project.

Newspapers layout detection on the Storia e Memoria di Bologna newspapers dataset

Storia e Memoria di Bologna is a project that intends to give voice to all the major and minor protagonists who have made the history of the city: the fallen during the Great War and the Resistance, the victims of the Monte Sole massacre, and the illustrious predecessors who rest in

the Monumental Certosa Cemetery. The biographies of these figures are strictly connected with events, places, works of art, monuments and archival documents that, through the different sections into which the portal is divided ("Scenari"), are explored in more detail.

In particular, the section used to create the dataset was the one dedicated to the city's newspapers ("[Sfoggia i giornali](#)") (Fig. 5).



Fig. 5 - *Storia e Memoria di Bologna* web page: the historical newspaper section.

This section contains a collection of newspapers in Italian language, *Il Resto del Carlino* and *Il Giornale dell'Emilia*, published during the World Wars years. From these newspapers a selection has been made by choosing for each year all the issues that came out within 3-4 specific months. This choice was made to cover possible changes in layout that characterized the newspapers over time. After collecting an adequate number of newspapers, the different models were tested by simply following the code in the software documentation (Fig. 6).

```
import layoutparser as lp
model = lp.Detector2LayoutModel(
    config_path = 'lp://PubLayNet/faster_rcnn_R_50_FPN_3x/config', # In model catalog
    label_map = {0: "Text", 1: "Title", 2: "List", 3: "Table", 4: "Figure"}, # In model`label_map
    extra_config=["MODEL.ROI_HEADS.SCORE_THRESH_TEST", 0.8] # Optional
)
model.detect(image)
```

Fig. 6 - The code used for testing each model with *Storia e Memoria di Bologna* dataset.

Dataset	Label Map
HJDataset	{1:"Page Frame", 2:"Row", 3:"Title Region", 4:"Text Region", 5:"Title", 6:"Subtitle",
PubLayNet	{0: "Text", 1: "Title", 2: "List", 3:"Table", 4:"Figure"}
PrimaLayout	{1:"TextRegion", 2:"ImageRegion", 3:"TableRegion", 4:"MathsRegion", 5:"SeparatorRegion
NewspaperNavigator	{0: "Photograph", 1: "Illustration", 2: "Map", 3: "Comics/Cartoon", 4: "Editorial Cart
TableBank	{0: "Table"}
MFD	{1: "Equation"}

Fig. 7 - The label maps defined for each dataset.

To test each model individually, it is enough to select the configuration path indicated in the models' catalogue (*config_path*) and the specific *label map* that allows the recognized layout areas to be named by labels (Fig. 7). After defining the path where the png images, converted from the original pdf files, are located, the model-specific function for detecting the chosen image layout (*model.detect(image)*) was applied and the code executed.

After several tests on selection of images using the three main Layout Parser datasets (HJDataset, PubLayNet, Tablebank) as a reference point, issues were encountered:

- On different occasions the areas in the newspaper page (header, titles, subtitles, images, text of the individual article, etc...) were not recognized correctly by the corresponding bounding boxes, i.e., rectangular figures that the model uses to delimit a specific identified area. Often the position of those areas was inaccurate with respect to its content.
- Overlapping bounding boxes belonging to areas of different types; this appears to be the consequence explained in the previous point, in fact many times the boundaries of one area, in addition to being misplaced with respect to its content, were also overlapped with those of another.
- No recognition of any area on the page under examination.
- Excessive noise found in the final result regardless of the recognition threshold value set. Within the code there is the possibility of setting the threshold value of the algorithm (*confidence threshold*): this value ranges from 0 to 1 and allows the model to produce a number of areas that it considers to be potentially more or less valid:

```
extra_config=["MODEL.ROI_HEADS.SCORE_THRESH_TEST", 0.8]
```

With values close to zero the model identifies on a page a greater number of areas of which it is not totally certain of their correctness, while values close to one the model will provide areas that are considered valid with a greater degree of confidence.

The *confidence threshold* acts to some extent as a filter in the number of recognized areas, since those with values below the threshold will not be considered by the model: the closer the threshold is to zero potentially the greater the number of identified areas, even though some of them may be incorrect; on the contrary, a threshold value close to one implies a potentially smaller number of areas on a page although theoretically more correct.

A non-optimal *confidence threshold* value (too low or too high) can lead on the one hand to many incorrectly identified areas, and on the other hand to a small number of correct areas but with the possibility of having excluded others that are potentially valid.

Nevertheless, in our tests setting the confidence threshold value close to 0 or 1 in several cases did not affect the number of areas recognized on the page because they were either too many and indistinguishable or completely absent.

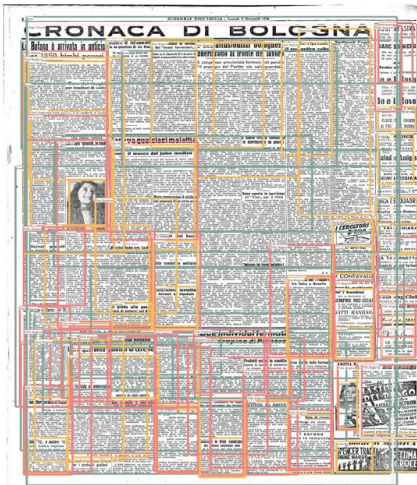
The main cause of the problems listed above was identified in the fact that the models were trained to recognize layout elements on different types of text compared to that of *Storia e Memoria di Bologna* newspapers, and this would explain the reason for largely unsatisfactory results.

More specifically, the historical Japanese-language texts in HJDataset have a layout structure (text arranged for vertical reading) and language characters (ideograms) that are completely different from the Italian language of the newspapers.

As for the PubLayNet and Tablebank datasets, these are born-digital scientific publications containing computer-generated writing, complex images and tables of various nature, mathematical formulas and text arranged mostly in one or two columns. In contrast, the resources of *Storia e Memoria di Bologna* are medium-to-high resolution scans of newspapers from the first half of the 20th century, with text that is often very dense and arranged in multiple columns.

Example of two complete opposite results

Dataset: HJDataset, Modello: retinanet_R_50_FPN_3x, Threshold: [0.10, 0.50, 0.80]



Dataset: PubLayNet, Modello: mask_rcnn_X_101_32x8d_FPN_3x, Threshold: [0.10, 0.50, 0.80]



Figg. 8-9 – Testing threshold values with different dataset and models.

In the image on the left it can be seen that, regardless of the threshold values set, the final result was a page in which no precise layout area could be accurately detected; on the other hand in the image on the right, using the same threshold values but different dataset and models, the final result was always a page in which no layout area could be detected. Further tests were done on other images trying the three models and corresponding datasets, but the results were still approximately the same (Figg. 8-9).

Newspaper Navigator Dataset and Faster_rcnn_R_50_FPN_3x model

Following the results previously obtained the test dataset was also tested with the *mask rcnn/faster_rcnn* models trained on the PrimaLayout, Math Formula Detection and Newspaper Navigator datasets. In the case of the PrimaLayout and Math Formula Detection models the results were very similar to those shown above for the same reasons presented before. The *Faster_rcnn_R_50_FPN_3x* model trained on the Newspapers Navigator Dataset, to the contrary, provided the best results, compared to HJ Dataset, Publaynet e TableBank datasets as well.

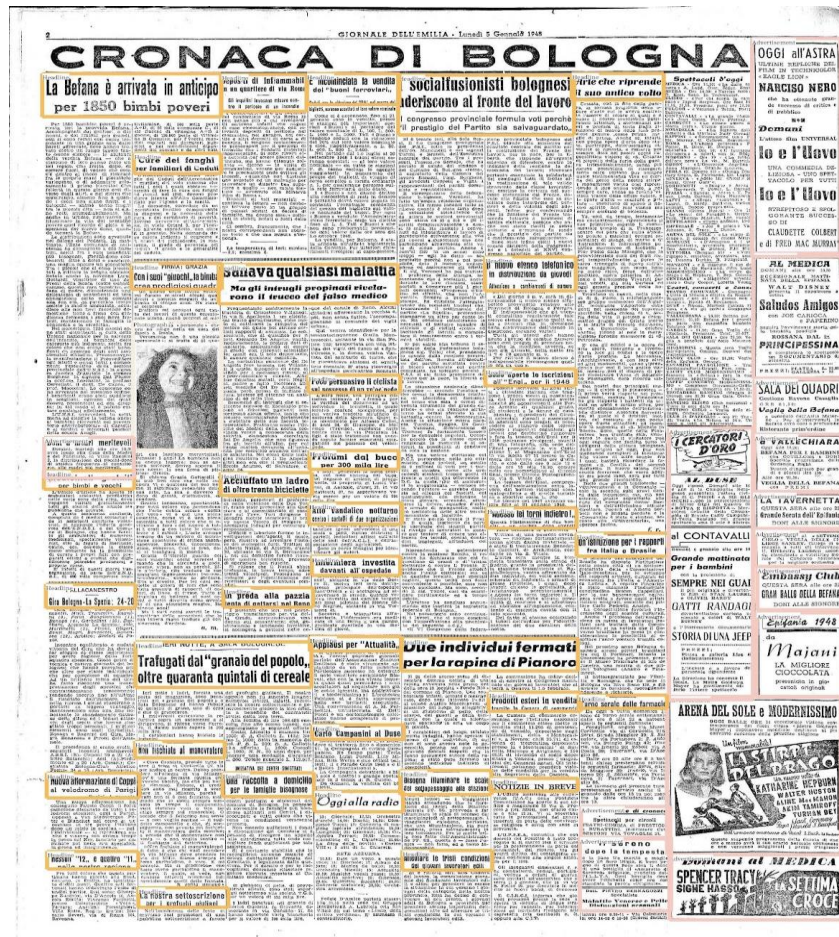


Fig. 10 - A result of the faster_rcnn model trained with Newspaper Navigator dataset.

As already mentioned, The Newspapers Navigator Dataset is a collection of 16,358,041 visual contents extracted from historical newspaper pages of *Chronicling America*. Such content was identified using an object detection model on *Chronicling America*'s page annotations related to World War I. Among the various contents were identified titles and subtitles of articles, advertisements, maps, illustrations, comics/cartoon, photographs, etc.

Figure n. 10 shows one of the best results obtained: in yellow are highlighted the headlines and subheadings of the articles, while in pink, on the right, are the advertisements. Even the photograph in the second column from the left has been properly recognized. On each bounding box there are labels for naming the detected area (Headline, Advertisement, Photograph). As can be seen most of the headlines, advertisements and the one photograph were identified and labelled correctly with a confidence score threshold set to 0.10.

The reason for the excellent results obtained is that the Newspaper Navigator dataset, on which the model was trained, presents a very similar type of resources in terms of layout with *Storia e Memoria di Bologna* dataset (they are both collections of newspapers from the first half of the

1900s) and the visual content identified in *Chronicle America* appears to be similar to what could be potentially extracted from the newspapers available in *Storia e Memoria* using the same object detection model of Newspaper Navigator dataset.

Unfortunately, the model was not trained to recognize paragraphs and textual areas under headlines, so it was not possible to identify the textual content of individual articles.

Nevertheless, an alternative way to solve this problem was attempted and the results are presented in more detail in another research conducted in relation to the Ministry project as well. The capability to recognize specific layout areas in a document such as a newspaper can open up opportunities to further disseminate the content of a resource in original manner as well. Tools such as Layout Parser highlight not only the great potential of software equipped with artificial intelligence models but also how much such instruments can improve the dissemination, study, and understanding of complex and distant documentary resources.

Creation and development of a dataset for automatic transcription of handwritten text (HCR)³

The second study presented in this paper is about handwritten character recognition (HCR). Italy's cultural heritage is characterised by a considerable archival and book heritage, like loose documents and codices, most of which are handwritten [9].

These codices and documents typically belong to the medieval age and are characterised by being difficult to read. This happens because in the course of history, book and document writing, in their typologies, have undergone exponential evolution.

An example is the change between the *Roman capital script* of the 3rd century A.D. and the *Oncial* script of the 4th century A.D.; about a century elapsed between these scripts and in this time the script completely changed its initial connotations.

In addition, in codices and manuscript documents, it is also possible to observe a large number of abbreviations of various forms and types. The practice of abbreviations was very much in vogue in the Middle Ages, writing materials such as papyrus and parchment were expensive and difficult to obtain, and the production, complex, was time-consuming. Consequently, it was essential to apply space-saving criteria to save as much as possible on the purchase of papyrus and parchment media.

Unfortunately, there are not many codices and manuscript documents transcribed, meant as given in a type of handwriting that we can easily understand [10], with these types of changes:

³ The following section has been written by Giorgia Cardano.

dissolving abbreviations, replacing ancient punctuation (where present) with a modern one, capitalising letters as in the manuscript, separating words to bring them back to the morphosyntactic norm in obvious cases, marking textual gaps along with line changes, and maintaining numbering as in the original manuscript [11].

Therefore, as there are not many transcriptions, research in this area appears more complex and requires more time for the analysis of the codex/document.

To fill this great gap, [Cineca VisitLab](#), thanks to the national project for the digitisation of cultural heritage, is working at a tool for the automatic transcription of manuscript text (HCR), which could significantly reduce the time needed to analyse the materials and at the same time enrich our cultural heritage by bringing people with no particular knowledge on this field closer to it, as well as making a major contribution to the forthcoming creation of the [Digital library](#) with PNRR funds.

This project was based on the digitisations of the *Copialettere of Isabella d'Este* by [Prof. Deanna Shemek](#) of the University of California Irvine, who kindly granted the images for neural network training.

Between the 15th and 16th centuries the passage from the medieval to the modern age was definitively established. The advent of humanism characterised the flourishing of classical language studies and classical literature, as they were considered a tool of spiritual elevation for mankind, and the recovery, study and publication of classical texts were placed at the centre, with the final goal to elaborate a more modern civilisation based on a new way of thinking.

It was during this period that the figure of Isabella d'Este emerged in Mantua, born in Ferrara in 1474 from Duke Ercole and Eleonora of Aragona and educated by Battista Guarini. At the age of sixteen, she was given in marriage to the Marquis of Mantua Francesco II Gonzaga and in this environment, she celebrated art in all its forms from the beginning, so much so that she gathered around her the *Accademia de Santo Pietro* in which Equicola (her secretary), Battista Spagnoli, Baldassarre Castiglione and Matteo Bandello took part.

Isabella entrusted high-caliber artists with constructions and decorations, bringing together masterpieces and precious collections in her *Studiolo*.

In addition, Leonardo da Vinci, il Francia, and Tiziano painted portraits of her, while Raffaello was asked to represent her eldest son, Federico. Isabella was a woman of great culture and a great patron of the arts with a truly superfine artistic taste, so much so that Ariosto described her

as “Isabella liberal and magnanimous”, while Matteo Bandello described her as “supreme among women' and Niccolò da Correggio reserved for her the title of “the first woman in the world”.

The Marchioness of Mantua proved to be politically astute as she had to take the regency of the Marquisate for a year because her husband, Francesco II Gonzaga, was imprisoned in Venice in 1509. Thanks to the diplomatic intervention of Isabella and her ambassadors: Ludovico Brognolo, Fra Anselmo da Bologna and Pope Julius II, she managed to free her husband from imprisonment, although she had to hand over her son Federico, who stayed at the papal court, as a pledge of loyalty [12]. Isabella also had a very dense correspondence, as she maintained friendships with men of letters and artists, as well as diplomatic and power relations with European sovereigns, emperors, princes and lords of the Italian courts.

This very complex network of relationships can be read through her more than considerable epistolary (one of the most important for quantity and quality of the Renaissance). Today, the Mantua State Archive preserves and safeguards this legacy of great importance, of which there are a total of 47 letters written by Isabella's hand, divided into the *Autografi* collection, constituted by the archivists in the 18th century, *Autografi Volta*, dating back to the first half of the 19th century, and *Tamassia*, a miscellany of epistles and decrees. In contrast, the F. II. 9 of the impressive Gonzaga Archive includes all the *Copialettere* of the Gonzaga household, the books in which the missives written by the chancellery were copied.

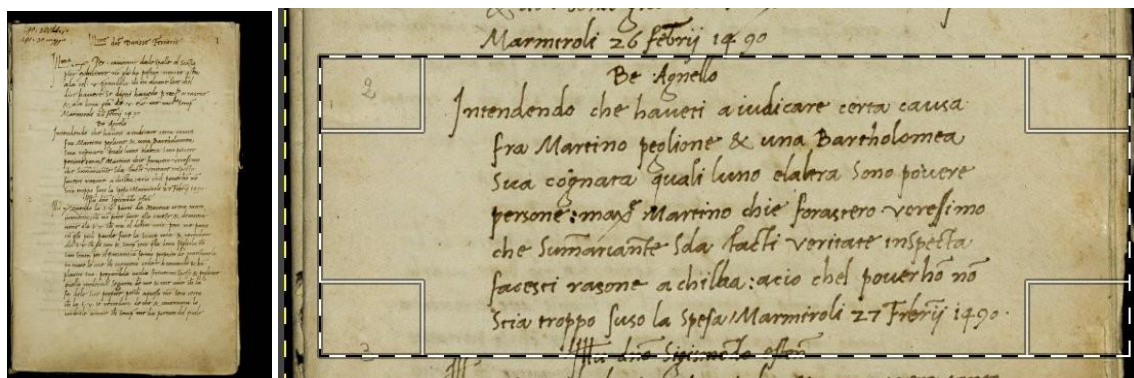
Isabella had 53 *Copialettere* reserved for her, divided into 10 envelopes (from 2991 to 3000) and there are also other registers scattered in the ordinary *Copialettere*, which collect the missives of the Lords of Mantua. Among the letters there is unfortunately a large gap from mid-November 1524 to mid-June 1527, the period when the Marchioness stayed in Rome; in fact, during the return journey to Mantua one of the two ships on which the missive registers were loaded was attacked and plundered by the Moors [13].

Part of these *Copialettere* was digitised and transcribed as a case study for the development of workflow using neural networks trained in the automatic transcription of manuscript text (HCR). These developments can change the fate of source research and be a turning point in the fields of paleography, codicology, archives, library science and philology.

Creating the neural network training dataset

Each sheet in the letter copies contains several letters within it (Fig. 11), so with the help of the [Gimp](#), the first step consisted of selecting the relevant portion of text concerned (Fig. 12) and the smoothing, through the gradient tool, of all those signs that are not writing, such as.: strokes

belonging to the previous or following letter, writing that leaked on the front side and stains on the paper. The aim was to eliminate all those handwriting marks and strokes that in some way could disturb and confound the neural networks in their training.



Figg. 11-12 – A page from a *Copialettere* pertaining to Isabella d'Este and one of the letters.

In the next step, the transcription of the manuscript text was carried out without following the canonical transcription criteria but by equating the carriage returns, retaining the ancient punctuation marks and numbering in Arabic or Roman numerals and dissolving the abbreviations in the text (Fig. 13).

```
C136_D2.py
1 [{"doc": 'C136_D2', 'image': 'MDP2904/0003_[136]_C_001_R.tif', 'top': 286, 'bottom': 487, 'transcription':
2     "" Be .Agnello
3     Intendendo che haveti a iudicare certa causa.
4     Fra Martino Peglione et una Bartholomea
5     sua cognata quali luno elaltra sono povere
6     persone: et maxime Martino chiè forastero. voressimo
7     che summariamente sola facti veritate inspecta
8     facesti rasone a chilha: acio chel pover homo non
9     stia troppo suso la spesa, Marmirolì 27 februarii 1490""}]
```

Fig. 13 – An example of transcription

The decision to deviate from the transcription norms stems from the need to match the transcription to the handwritten text and to make it easier for the neural network to associate the handwritten script with the given, manually done, transcription, so that the neural networks residing within the automatic transcription tool (HCR) can be trained.

In the course of this complex process, paleographic evaluations were made, which relate to the subject of paleography, the discipline that investigates writing and studies written documents of the past, of all types and epochs in socio-cultural contexts, reconstructing the history of their graphic forms. Paleography is also able to decode, date and locate writings through the study of execution techniques and the comparison of different examples [14].

Paleographic assessments cover the fundamental criteria underlying the discipline [15], namely:

- The form: that is, the figural aspect of individual letters, an example is the way a letter can appear in its different manifestations (Fig 14). In this case, we can see that although the letter *t* is recurrent in the words *nocte* and *partesti*, it's represented differently in the two terms. In the word *nocte* the *t* is higher, and a link connects it to the *c*, while in *partesti* we can observe two low *t* that are very different from the previous one.

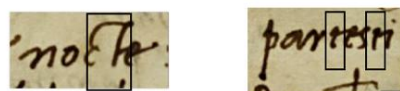


Fig. 14 – Some paleographic assessments on the ways of writing the letter “t” in the same document.

- The modulus: that is, the dimensions related to the height and width of the letters. In this case, the initial letters of the words *scordare* and *recordamovi* are not simple capital letters, but lowercase letters with a large modulus. If we were to draw a quadrilinear system, these two words would not fit into this scheme, but they do in a bilinear system and are therefore to be defined as lowercase letters (Fig. 15).

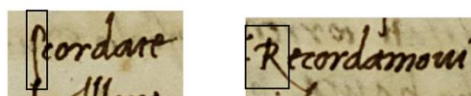
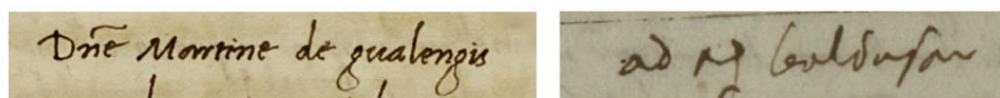


Fig. 15 – Some paleographic assessments on the ways of writing capital or lowercase letters.

- The ductus, meant as the course of writing. Two examples are given here: in the first, there is a laid writing in which the words are almost drawn rather than written (Fig. 16), in the second there is more cursive writing that unlike the first is full of ligatures between the letters and is slanted to the right; moreover, the need to write quickly is evident. It is precisely this celerity that prohibits the letter-painter from raising the writing instrument from the writing support between the letters of the same word, to keep the writing rhythm tight (Fig. 17).



Figg. 16-17 – Some paleographic assessments on the course of writing.

- The hatching: expresses the number, order and direction in which the individual strokes of the letter are executed, but also their thick or thin nature. In this case, the letter *r* has three strokes, the first corresponds to a vertical stroke, the second is a curved stroke and

the third stroke proceeds diagonally, so we can state that these three strokes head in different directions and have a thick nature (Fig. 18).

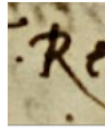


Fig. 18 – Some paleographic assessments on the way a letter “r” is written.

- The ligature: this is the link that binds two or more letters together without the writing instrument rising from the writing material. In this case, in the abbreviation “vro”, the letter v ends with a downward stroke and the letter r begins with an upward stroke; therefore, in order to connect these two letters, the writing instrument, not rising from the writing material, created a ligature that allowed the joining of the letters v and r (Fig. 19).

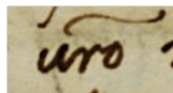
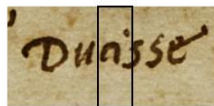


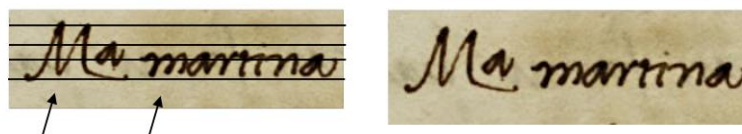
Fig. 19 – Some paleographic assessments on ligature between written letters.

- The nexus: it is the set of letters that share the same stroke. In Isabella d'Este's *Copialettere* there are not many nexuses, one of the few is in the word *ducis*, in which we can observe the union of the letters “c” and “i” (Fig. 20).



- Fig. 20 – One of the few words in the *Copialettere* presenting a nexus.

- Upper- and lower-case letters: these letters are determined according to whether they enter a quadrilinear or bilinear system. In this case, the first m entering a quadrilinear system can be defined as an uppercase letter, whereas the second occupies the two lines of the bilinear system perfectly and is therefore a lowercase letter (Figg. 21-22).



Figg. 21-22 – The quadrilinear system shows whether letters can be considered capital or lower-case.

Another important issue that the HCR tool had to take into account are the abbreviations. The abbreviations found in Isabella d'Este's *Copialettere* are mainly of two types, by truncation and contraction. Abbreviations by truncation are found when the initial part, where its radical alphabet is preserved. An example is in the case below, where *Ma* stands for *Madama* (Fig. 22).

On the other hand, abbreviations by contraction occur when a dot or other signs replaces the central letters of a word, as you can see in Fig. 23. In particular cases abbreviations for the same word occur in different ways in different letters.

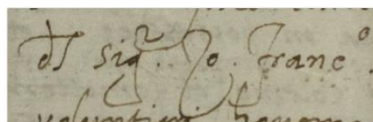
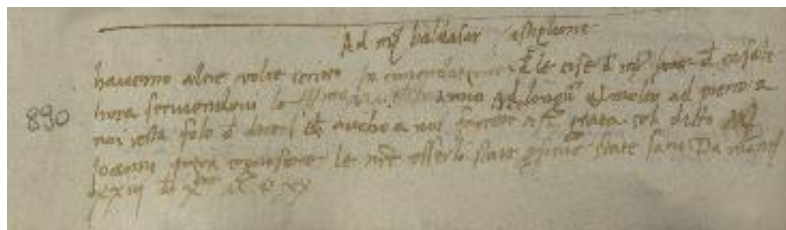
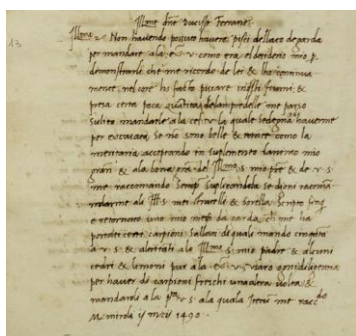


Fig. 23 – An example of abbreviations by contraction.

Other difficulties ensue from abbreviations applied to dates and the use of Arabic or Latin numerals.

Another relevant aspect in the training of neural networks was the heterogeneity of the letters in terms of both spelling and structure. Although the approximately two hundred letters were in humanistic script, as they do not belong to the same hand, they are very different from each other and even the way the words are abbreviated is affected by the subjectivity of the writer.

In addition, the layout of the missives is not homogenous; in some letters, the rifling of the paper and the definition of the writing mirror is clearly visible (Fig. 24). On the other hand, other letters, due to the absence of these operations, are more disordered and irregular, making them more difficult to understand (Fig. 25).



Figg. 24-25 – Disomogeneities in the layout of the letters.

To conclude, after the creation of a dataset consisting of carefully selected images with the corresponding transcriptions, performed manually, and a careful palaeographic and codicological

analysis of the missives, we now proceed with the training of neural networks and the automatic transcription of manuscript text (HCR).

Training of neural networks

The training of neural networks was initially carried out using [TrOcr](#) (Transformers - based Optical Character Recognition), an open-source solution based on the use of Transformers-type networks. TrOcr has ready-to-use models in English, but since it had to work on coplefts with a Renaissance Italian language, specific training had to be implemented. It was precisely to provide for this necessity that the dataset described above was created, comprising a corpus of approximately two hundred letters with the corresponding manually executed transcriptions. Since the network works with images containing a single line of text, a division of the digitisations of the missives by lines with the corresponding transcriptions was implemented. In support of this, the [P2PaLA](#) (Page to Page Layout Analysis) neural network was used. It identifies the individual lines of text in the image and return their position using a polyline formed by segments connecting points expressed as coordinates in pixels in the image (Fig. 26).

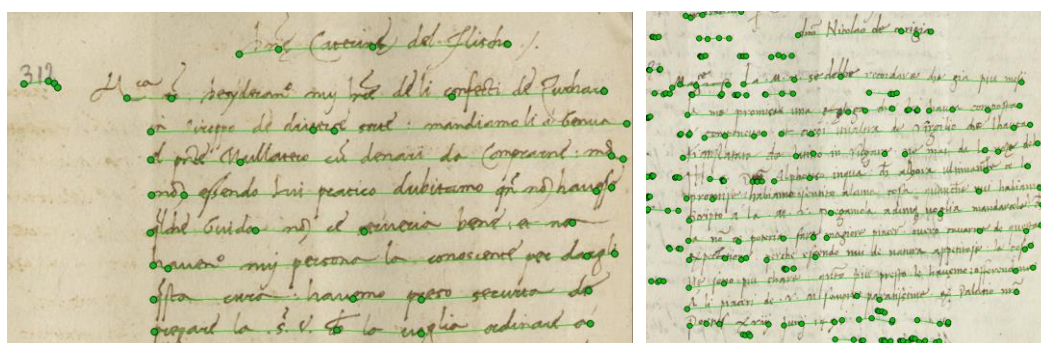
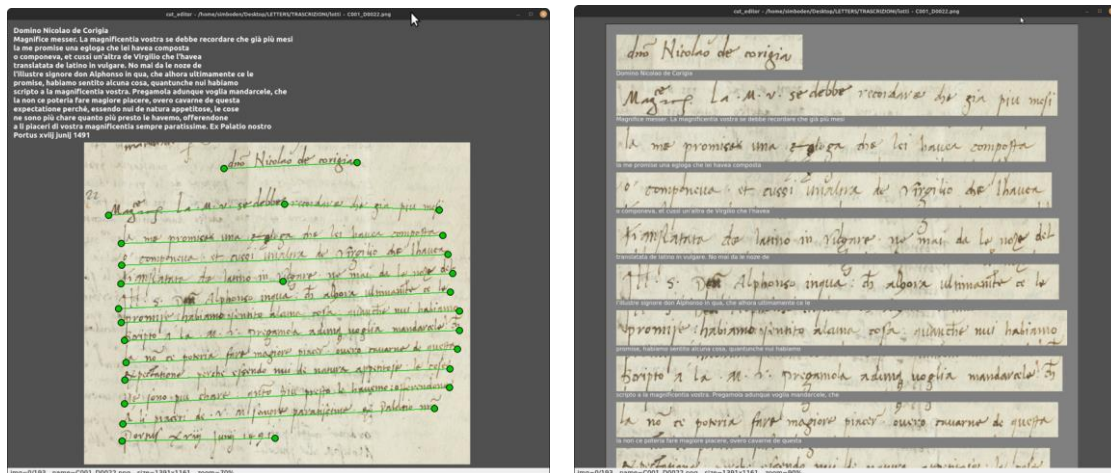


Fig. 26-27 – Selecting lines of text with P2PaLA. Errors can ensue due to irregular writing or the selection of characters and signs visible from the back of the page.

Sometimes, however, extraneous elements such as stamps or numbering are recognised between the lines of text. In addition, the lines may appear broken into several segments because the network is confused by various elements, such as: the beginning and end of words, abbreviations, capital letters, decorative flourishes and tildes. The transparency of the writing support also causes the unwanted presence of writing traits that belong on the back of the paper (Fig. 27).

To overcome the problem of inconsistent quality in line recognition a specific tool, the *Row_splitter*, was developed for the visualisation and possible correction of line recognition. In edit mode, it is possible to insert, move or delete dots and polylines, even on multiple selections,

and see the transcription text by moving quickly between the various missives in the corpus (Fig. 28). *Preview mode* shows a preview of the subdivision of the image into individual lines of text. Lines are not only isolated but also automatically rotated to bring them to the horizontal axis. In addition, this mode allows the text of the original manuscript to be checked for correspondence with the one of the digital transcriptions (Fig. 29).



Figg. 28-29 – Edit and Preview mode.

Once the above steps were completed, initial training was carried out, in which the transcribed lines had loose abbreviations, but it soon became clear that the network needed an exact correspondence between the characters of the image and those of the transcription. Therefore, it was necessary to contract the previously loose abbreviations to match the number of characters of the original manuscript (Fig. 30).

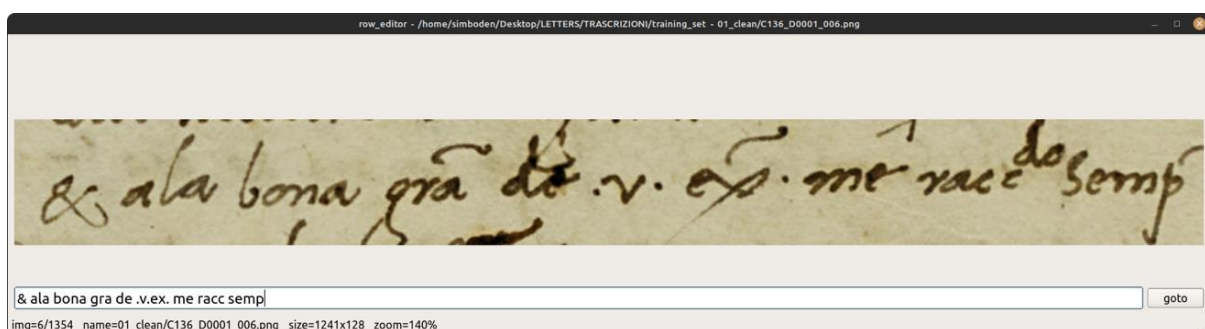


Fig. 30 – Matching the transcription with the original text.

A second training with contracted abbreviations was performed, leading to a correct automatic transcription of handwritten text by 86%, a very high percentage indeed. This result is an excellent starting point for easier research in the fields of paleography, archiving, library science and philology, as well as for a flourishing transcription activity in the future, with

the hope of possessing transcriptions of all Italian manuscripts and documents regardless of their difficulty or importance.

References

- [1] Elisha Odemakinde, *Everything About Mask R-CNN: A Beginner's Guide*, March 19, 2022, <https://viso.ai/deep-learning/mask-r-cnn/>
- [2] Everton Gomede, *RetinaNet: Advancing Object Detection in Computer Vision*, August 27, 2023, <https://medium.com/@evertongomede/retinanet-advancing-object-detection-in-computer-vision-719ceb744308>
- [3] Zejiang Shen, Kaixuan Zhang, Melissa Dell, *A Large Dataset of Historical Japanese Documents with Complex Layouts* (2020), <https://arxiv.org/abs/2004.08686>
- [4] Xu Zhong, Jianbin Tang, Antonio Jimeno Yepes, *PubLayNet: largest dataset ever for document layout analysis* (2019), <https://arxiv.org/abs/1908.07836>.
- [5] Minghao Li, Lei Cui, Shaohan Huang, Furu Wei, Ming Zhou, Zhoujun Li, *TableBank: A Benchmark Dataset for Table Detection and Recognition* (2020), <https://arxiv.org/abs/1903.01949>
- [6] A. Antonacopoulos, D. Bridson, C. Papadopoulos, S. Pletschacher, *A Realistic Dataset for Performance Evaluation of Document Layout Analysis*, Proceedings of the 10th International Conference on Document Analysis and Recognition (ICDAR2009), Barcelona, Spain, July 2009, pp. 296-300, https://www.primaresearch.org/datasets/Layout_Analysis.
- [7] Lee, B. C., Mears, J., Jakeway, E., Ferriter, M., Adams, C., Yarasavage, N., Thomas, D., Zwaard, K., & Weld, D. S., *The Newspaper Navigator Dataset: Extracting and Analyzing Visual Content from 16 Million Historic Newspaper Pages* in *Chronicling America* (2020), <https://arxiv.org/abs/2005.01583>.
- [8] F. M. Schmitt-Koopmann, E. M. Huang, H. -P. Hutter, T. Stadelmann and A. Darvishy, *FormulaNet: A Benchmark Dataset for Mathematical Formula Detection*, in *IEEE Access*, vol. 10, pp. 91588-91596, 2022, doi: 10.1109/ACCESS.2022.3202639.
- [9] A. Pratesi, *Una questione di metodo: l'edizione delle fonti documentarie*, «Rassegna degli Archivi di Stato», 17 (sett.-dic. 1957), pp. 312-333.
- [10] C. Paoli, *Diplomatica*, ed. aggiornata da C. Bascapè, Firenze, Le Lettere, 1942.
- [11] G. Tognetti, *Criteri per la trascrizione di testi medievali latini e italiani*, Roma, s.e., 1982.
- [12] R. Tamalio, *Isabella d'Este, marchesa di Mantova*, in *Dizionario Biografico degli Italiani - Volume 62*, Istituto della Enciclopedia Italiana Treccani, 2004, [https://www.treccani.it/enciclopedia/isabella-d-este-marchesa-di-mantova_\(Dizionario-Biografico\)/](https://www.treccani.it/enciclopedia/isabella-d-este-marchesa-di-mantova_(Dizionario-Biografico)/)
- [13] M. Basora, *L'epistolario di Isabella d'Este: dai libri dei copialettere alla piattaforma IDEA*, in *Natura Società Letteratura*, Atti del XXII Congresso dell'ADI, Associazione degli Italianisti, Bologna, 13-15 settembre 2018, a cura di A. CAMPANA e F. GIUNTA (edd.), Roma, Adi editore, 2020.
- [14] *Paleografia*, in *Enciclopedia Treccani online*, <https://www.treccani.it/enciclopedia/paleografia>
- [15] A. Petrucci, *Breve storia della scrittura latina*, Milano, 1992.

ADVANCED SOLUTIONS FOR AUTOMATIC IMAGE AND TEXT ANALYSIS AND ENRICHMENT

Donatella Sforzini, Gabriele Fatigati, Luca Mattei – Cineca

Introduction

2023 will be remembered as the year that marked the watershed between before and after in reference to AI, especially Generative AI (GenAI). In just 12 months, many of the image/text/language processing technologies, that were considered state of the art at the beginning of the year, have become obsolete. Not only has the number of tools that deal with different tasks exploded, but also the main players have released evolutions of their models within a few months of each other.

In this regard, in this paper we will discuss the development and updating of two tasks in the Cultural Heritage field that are particularly useful in the Digitalization and Cataloguing Process of Cultural Heritage: Object Detection, with reference to computer vision, and Named-entity recognition (NER), with regards to Natural language processing (NLP).

What can AI do for art?

In January 2023 we posed the question "What can AI do for art?" to ChatGPT (v.3.0) and received the following answer:

"AI can be used in various ways to create, generate or analyse art. For example, it can be used to generate images, music or videos autonomously using machine learning algorithms. In addition, it can be used to analyse large amounts of data for research purposes or to help create new artistic techniques. In addition, it can be used to create interactive user interfaces for art and to generate new forms of entertainment."

However, 2023 will be remembered as the year that marked the watershed between before and after in reference to the explosion of Generative AI (GenAI). In 12 months, it has made many of the previous image/text/language processing technologies obsolete, and in addition, a new evolutionary step in its models is released by the main players every 6-9 months.

A year later then, we asked ChatGPT (v.3.5) the same question and the answer turned out to be much more complex:

«Artificial intelligence (AI) can play different roles in the field of art, contributing in innovative and influential ways. Some of the most significant applications include:

1. **Art Creation:** Automatically generating algorithms can create works of art, including paintings, music, poems, and even sculptures. Artists and developers often collaborate to create unique works using AI models, contributing to new forms of artistic expression.
2. **Creative assistance:** AI-powered tools can assist artists in the creative phase, offering suggestions, generating ideas or helping to choose colors and styles. This can be especially useful for artists looking for inspiration or wanting to experiment with new directions.
3. **Analysis and criticism:** AI can be used to analyze and criticize works of art. Some algorithms are able to evaluate aesthetics, originality and other artistic aspects, providing feedback that can be useful for artists in refining their works.
4. **Restoration and conservation:** AI can be used in the restoration and conservation of works of art. Templates can help repair damage digitally, improve image quality, and digitally preserve fragile or damaged works in the real world.
5. **Immersive experiences:** AI can be integrated into immersive art experiences, such as interactive installations or virtual experiences. This can broaden creative possibilities and engage audiences in innovative ways.
6. **Art Market Analysis:** AI can be used to analyze art market trends, predict valuations of artworks, and assist collectors in buying or selling based on historical data and market trends.
7. **Personalization of art:** AI-based technologies can be used to create personalized works of art based on individuals' tastes and preferences, offering a unique artistic experience for each person.

It is important to note that the use of AI in art also raises ethical and philosophical questions, such as the definition of creativity, artistic authenticity, and the human role in artistic creation. However, the collaboration between artists and artificial intelligence continues to evolve, offering new possibilities and challenges in the art world."

In addition to being much more complex, the answer we were given also raises ethical and philosophical questions that are at the origin of the need to provide AI developers and operators with clear requirements and obligations regarding the specific uses of AI through a specific law (AI act).

Besides the 7 points listed, in the broader field of Cultural Heritage, at Cineca we are building workflows based on AI algorithms to power collaboration between human and artificial intelligence in processes such as Cataloguing Cultural Heritage. In this article we will discuss the research done in two specific tasks such as Object Detection (OD) and Named Entity Recognition (NER).

Overview of tested tools in Cineca workflow

[Object Detection](#) is a computer technology in the field of computer vision. The tools we use provide indications about the objects in the image being analysed such as: what type of object it is, the confidence measure of the algorithm and its location in the image.

As we have already seen for GenAI, the speed with which new tools or their updates are released is impressive. In 2023 we had to change one of the Object Detection tools precisely because of the continuous evolution in this field, moving from [Detectron2](#) to [Yolo5](#).

NER is a Natural Language Processing task. Starting from a text file, it identifies entities within it and classifies them into a given set of labels (e.g., PER-people, ORG-organizations, LOC-locations, MISC-miscellaneous).

For both types of analysis, we will see how it is possible, indeed necessary, to adapt the tool to the specific context of Cultural Heritage to improve the performance of the AI algorithms which, in this context, integrates, supports and collaborates with the human and does not completely replace it.

Object detection

Object Detection - computer technology related to computer vision and image processing - detects instances of semantic objects of a certain class (such as humans, buildings, or cars) in digital images and videos. More in details it is asked to find the object and the rectangle that delimits the boundary of the located object. If the problem requires finding different classes of object, then the programme must also explain which class of object is inside the rectangle.

Sometimes this process is confused with image classification or image recognition, which aims to predict the class of an image or an object within an image into one of the categories or classes.

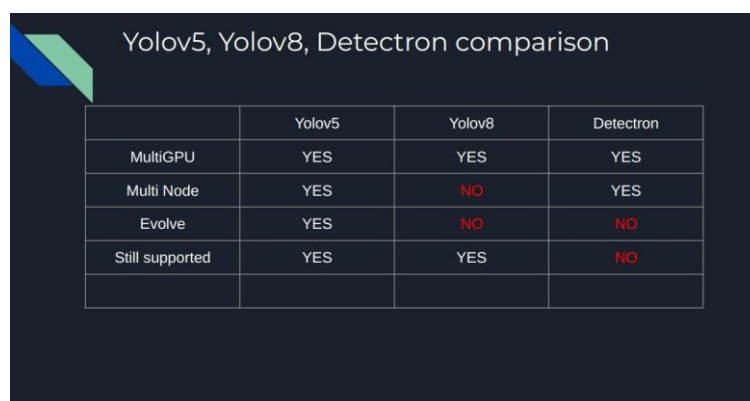
The Object Detection task in Cineca workflow for Cultural Heritage began in 2023 with the use of Detectron2, CLIP and [OWL-ViT](#) on the SGoaB Datasets [1].

In 2024 Object Detection technical activity saw two important innovations: the transition from Detectron2 to Yolo and the implementation of the entire pipeline from Marconi supercomputer to the Booster partition of Leonardo HPC cluster. Leonardo is equipped with 4 Nvidia A100 GPU per node. Each GPU has 64 Gb of RAM. The total of nodes is 3456 nodes, so the total amount of GPUs is 13824. Porting the pipeline on Leonardo required some updates, such as modules version and source code upgrade.

The choice of Yolo was due to Detectron2 no being maintained by the community any longer, with consequent stop to new features, bug fixes etc. Yolo, on the other hand, presented in 2015 by Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi in their famous research paper “You Only Look Once: Unified, Real-Time Object Detection” [2] has been the state of the art for Object Detection for years now and enjoys a large community that constantly updates it.

The chosen version of Yolo is v5:

- Simple to use, both in training and validation phase;
- MultiGPU and Multinode running mode;
- By default, it generates a variety of visual and textual report statistics useful for the researcher;
- Evolve phase: possibility to tune training hyperparameters (can take a long time).

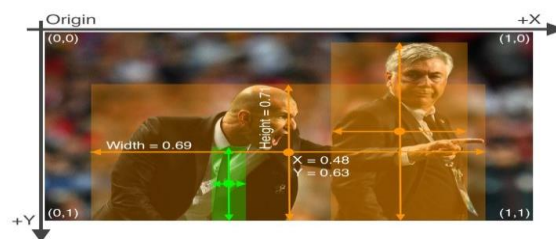


	Yolov5	Yolov8	Detectron
MultiGPU	YES	YES	YES
Multi Node	YES	NO	YES
Evolve	YES	NO	NO
Still supported	YES	YES	NO

Table n. 1 - A comparison between Yolov5, Yolov8 and Detectron.

The first two characteristics, in particular, are essential if you plan to do massive training on large datasets (300,000 images and more). Yolo has a different data format than Detectron, so this new format had to be implemented in the processing pipeline.

In Fig. 1 it is possible to see an example of an annotated image and the file of its annotations. Bounding boxes in Yolo are described starting from the origin at the top left. For each image a .txt file is created in which the first column is the class-id and the coordinates of the bounding box are represented according to the diagram centre of the image (second and third columns), then width (fourth column) and height (fifth column).



The annotation file for the image above looks like the following:

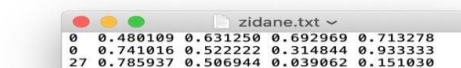


Fig. 1 – Annotated image.

As already introduced in the previous work [1], there are not many annotated datasets useful for YOLO training and those present have few images for each of the classes we want to analyze. Furthermore, these classes are very heterogeneous so there are some that are very general (crucifixion, landscape...) while others are very specific (skullcap, trumpet, sheep...) so we chose to continue using zero-shot tools for automatic annotation without class limits via Owl-Vit on whose results Yolo is trained.

Despite SGoaB is annotated dataset, Owl-Vit is still necessary to have many more images available.

The Owl-vit version used is a branch of original code, with some additional features developed internally:

- Images with bounding box;
- Xml annotation files. Very useful to the entire pipeline;
- CSV file with confidence, class detected and bounding box coordinates.

The typical workflow to prepare and run Yolo v5 on Leonardo HPC system foresees:

- Autolabeling of non-annotated images with Owl-Vit.
- Split of the global dataset into training datasets and validation datasets via Multilabel Stratification10 algorithm. This method allows for an optimal division, guaranteeing the best possible coverage for all classes, both in training and validation. In fact, by dividing randomly, there is the risk that some classes are not well represented (few images in training or validation).
- Converting the annotation format to Yolo format (as shown above).
- Training a Yolo network on an HPC machine.
- Network fine tuning.

These steps are very time consuming, especially step 5 can require days, but allow you to have the best possible network with the dataset available. To improve the quality of the prediction, according to the mAP50-95 reference metric [3], the following experiments were carried out:

- Increased Owl-Vit threshold, from the original one of 0.25 to 0.80
- Merge of the SGoaB dataset with the images of MiC repository (MiC dataset)
- Test with Yolov5, v8, v9 [4] - very recent
- Test with Owl-Vit v2

MiC dataset is a sample of about 30,000 heterogeneous images provided by the Italian Ministry of Cultural Heritage on which we applied our pipeline. For the task of Object Detection for a vector of about 70 classes it is a very poor datasets and it is very unbalanced in favour of some categories (posters are two-thirds of the global data; see Table n. 2).

	N. of Images	
painting/sacred_family	2,767	9%
painting/madonna_with_child	1,010	3%
painting /dead_nature	1,875	6%
painting/portrait_category	3,043	10%
icon	679	2%
posters	19,722	66%
dressings	830	3%
	29,926	100%

Table n. 2 – Content of a sample dataset provided by the Ministry of Italian Culture for the workflow studies

Once it was verified that Owl-Vit works (compared to SGoaB SGoaB annotations), its use was extended to the dataset provided by MiC.

Increasing the threshold images from Owl-Vit allows for better accuracy in the bounding boxes for the training dataset, but with fewer images elected and usually fewer classes.

The performance of the Object Detection and localization algorithm is evaluated by a metric called Average Precision (AP)[5]. To evaluate the performance of the different tests and compare them with each other, we calculate the AP indicator.

Owl-Vit threshold	AP	Num Images	Num classes
0.25 (rif)	0.27	26400	74
0.30	0.30	24483	74
0.35	0.28	21633	73
0.40	0.30	16951	70
0.45	0.33	11472	66
0.50	0.33	6922	54
0.60	0.41	1719	33
0.70	0.51	106	14
0.80	-	3	3
Merge (Sgoab 0.50)	0.37	22024	85

Table n. 3 - List of various tests for calculating the Average Precision

Note that when the threshold is too high, the results are useless, due to small amount of data filtered.

Yolov5, the official version chosen, achieves an AP level of 0.37 on the SGoaB -MiC merge dataset. This result is considered acceptable but certainly improvable. According to the guidelines that Yolo provides, you should have at least 1,500 images per class. Unfortunately, now, the available MiC dataset is very far from this requirement, consequently it is difficult to have significant improvements in prediction.

In order to improve the results, it is therefore necessary to have a dataset at least 10 times larger than the current one, thus guaranteeing good training feeding.

In Figures 2 and 3; 4 and 5 some examples of the improvements obtained using Yolo5 instead of Detectron2 for the 0.25 threshold can be appreciated.

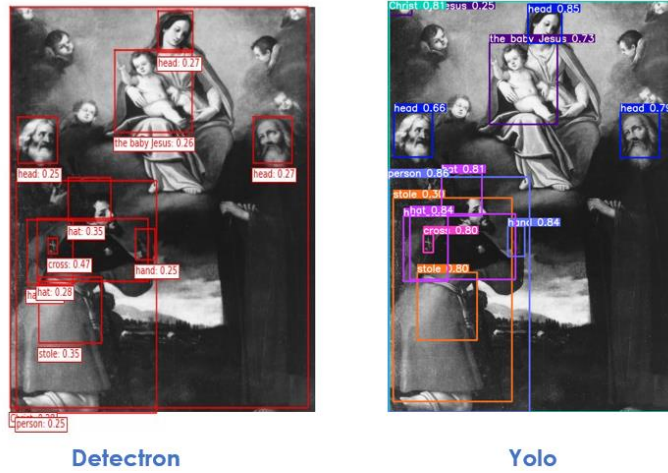


Fig. 2-3 – A comparison between the different performance of Detectron and Yolo

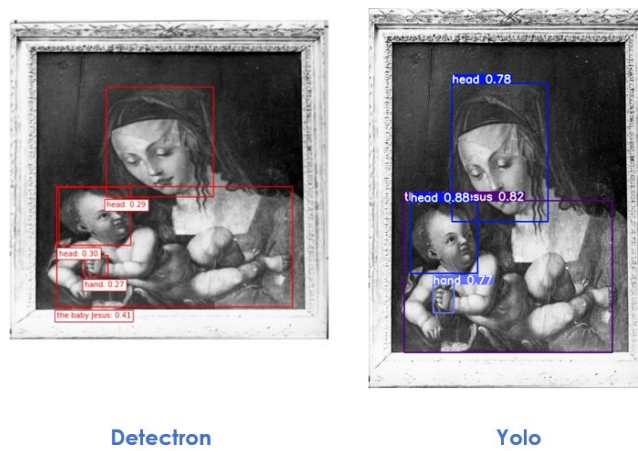


Fig. 4-5 – A comparison between the different performance of Detectron and Yolo

Next test is about the newest version of Yolo, v9 [4]. This version implements new network paradigms, allowing better object detection with fewer parameters, therefore less memory is required. The memory used is a very important factor to take into consideration when working on GPUs, since the memory available is in the order of 16-32-64 Gb max.

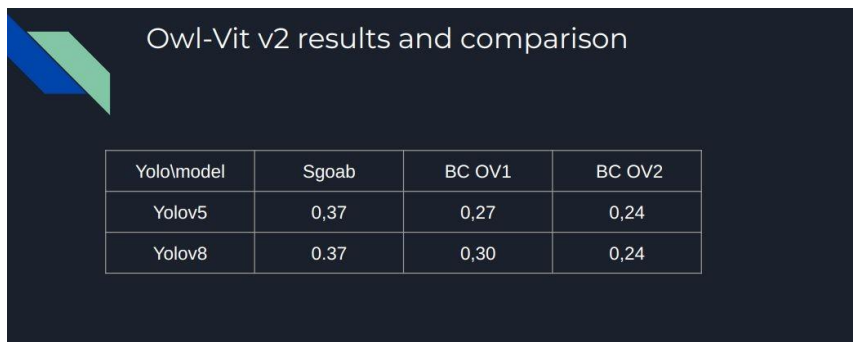
Yolov9

	Scratch				Pretrained			
	Sgoab	BC	BC-0.45	Merge	Sgoab	BC	BC-0.45	Merge
9c	0.33	0.29	0.32	0.33	0.34	0.29	0.34	0.36
9e	0.33	0.33	0.32	0.33	0.37	0.30	0.34	0.39

Table n. 4 – Results of tests conducted on Yolov9

In this test the peak of the AP is reached: 0.39. However, please remember that the official version is 5, since 9 lacks some essential features therefore, at the moment, it is used as a comparison (Table n. 4).

The last test in chronological order saw the test of Owl-Vit v2. The newest version introduces some internal variations about how to manage bounding box coordinates. Some source code modification was developed in order to align our version to the newest one. The entire pipeline was running starting from new annotations, but no final improvement was reached (Table n. 5).



Owl-Vit v2 results and comparison

Yolo\model	Sgoab	BC OV1	BC OV2
Yolov5	0,37	0,27	0,24
Yolov8	0,37	0,30	0,24

Table n. 5 – Results of tests conducted on Owl-Vit v2

Named Entity Recognition (NER)

Nowadays, text tagging, also known as text classification or text labelling, is considered a very important technique due to the many applications it could be used for:

- **Organizing the information:** tagging helps organize large volumes of text data by assigning relevant labels or categories to each document or piece of text. This organization facilitates efficient retrieval and management of information, making it easier to navigate and analyse textual data.
- **Search and Retrieval:** tagging enables users to search for specific topics, themes, or entities within a corpus of text. By associating relevant tags with document, users can quickly retrieve the information they need without having to go through irrelevant content.
- **Recommendation systems:** tagging can be used to personalize content recommendations for users based on their interests and preferences. By analysing the tags associated with previously consulted content, recommendation systems can suggest similar or related items that are likely to be of interest to the user.
- **Text Summarization:** tagging can assist in text summarization by identifying the main themes or topics covered in a document. By tagging key concepts or entities,

summarization algorithms can prioritize important information and generate concise summaries that capture the essence of the text.

Overall, text tagging plays a fundamental role in structuring, analysing, and extracting insights from textual data across a wide range of applications and domains. It enhances the usability, relevance, and efficiency of text-based systems and contributes to a better understanding of the content and context of textual information.

One of the most popular and used techniques for tagging text is NER, short for Named Entity Recognition, which is a subtask of natural language processing (NLP).

The aim of NER is to locate and classify named entities inside a text into pre-defined categories such as names, organizations, locations, dates, monetary values and more, depending on the needs.

The variability and ambiguity of natural language represent some of the most challenging characteristics of this task: named entities can appear in various forms, including different spellings, abbreviations, and linguistic variations.

The importance of NER was already recognized back in the 20th century, when various methods and algorithms began to be studied to perform this task; moreover, since they have been used for so long, the various techniques reflect the technological evolution.

The three main ones are:

- Rules-based Methods.
- Statistical Methods.
- Deep Learning Methods.

Rules-based NER was the first method adopted and it uses patterns and rules predefined by experts in order to retrieve the named entities in the text.

Later on, statistical methods became more and more popular in the field of NLP and they have been exploited for NER as well, especially with methods like Hidden Markov Models, Maximum Entropy models and Conditional Random Fields.

Finally, much more recently, with the advent of Deep Learning, the named entity recognition moved to this class of models, just like the majority of AI tasks. Once again, it followed the

researches in this field, by exploiting Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) at first, moving then to Long Short-Term Memory (LSTM) and finally Transformers Networks, which represent the State of the Art.

Considering the domain of cultural heritage, Named Entity Recognition plays a crucial role due to the necessity of organizing data, retrieving relevant information associated to the entities and most important, making all the heritage more accessible, searchable and analysable, not only for us but also for the next generations.

We experimented a few NER tools applying them on cultural data, more precisely, some descriptive cards retrieved from [Storia e Memoria di Bologna](#) (Fig. 6).

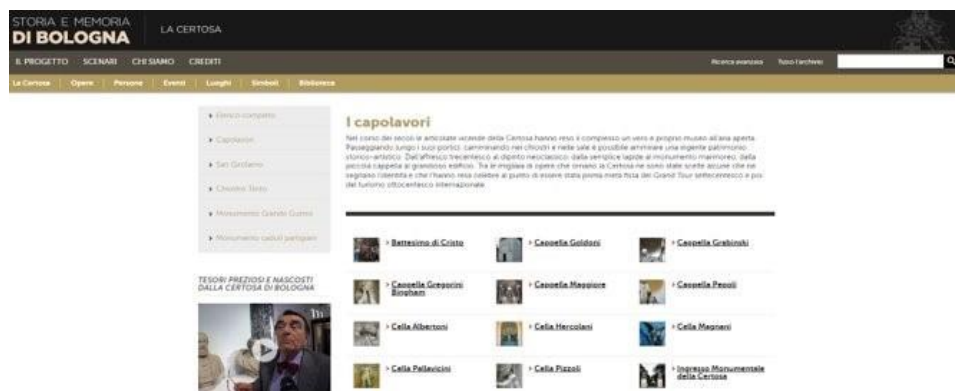


Fig. 6 - Storia e Memoria di Bologna web site

Storia e Memoria di Bologna is a municipal entity which deals with the coordination and valorisation of the heritage related to the history and memory of the homonymous city. The data used are all about the Certosa: a monumental cemetery located in Bologna. The sample was made of 1270 different textual data, but only 1130 were usable.

Unluckily, there was no labelled data to train a proper model with, which would have been a better option due to the specific domain; it follows that we had to use models already trained on other data, with a single constraint on the language (Italian), leading us to test the following tools:

[WikiNeuRal](#) is a Deep Learning model created by Babel by fine-tuning a multilingual BERT (Bidirectional Encoder Representations from Transformers) on the homonymous dataset.

[it-core-news-lg](#) is a model available on the *SpaCy*, an extremely popular open-source natural language processing (NLP) library, which relies on a combination of Rule-based and Machine Learning techniques.

Finally, [*spacy-dbpedia-spotlight*](#) uses the SpaCy model to find the entities, but instead of using the SpaCy labels it searches for a match inside of the Wikipedia Data-Base.

On the other hand, both *WikiNEuRal* and *it-core-news-lg* use the same set of labels to categorize the entities:

- PER, indicating a person;
- LOC, indicating a location;
- ORG, indicating an organization;
- MISC, for all the entities that the model finds but cannot associate with one of the other labels.

Regarding *spacy-dbpedia-spotlight*, after experimenting on it and seeing the extreme specificity of the labels and the high rate of wrong categories, we decided to drop it in favour of the other two models.

Going back on the unavailability of data, in order to improve our confidence on the results, we decided to take as true results only the entities retrieved by both *WikiNEuRal* and *It-core-news-lg* with the same label, making our inference pipeline a very simple ensemble: a popular technique used in the context of machine learning which consists in combining multiple models to improve the overall predictive performance.

On the total of 1130 documents *WikiNEuRal* found a total of 27602 entities, while the SpaCy model 42805; combining the two models without taking into consideration the labels, we ended up with a total of 21706 entities, dropping to 15148 if we get rid of the duplicates; in the first case the percentage of entities sharing the same label is equal to 83.6%, while it reaches 84.3% not considering the duplicates.

In Table n. 6 an example of the output of the inference model. The absence of an annotated dataset also makes difficult evaluating the performance of the model since there is no ground truth to compare the inference results with. However, we noticed that the number of unique entities found by both the models but with different labels were only 2385, so we moved on by manually labelling these entities so that we could have a rough estimate of the performance of the two single models. We proceeded by hypothesizing that the entities retrieved by both the models and with the same label were all correct and then we calculated our metrics on the rest of the entities.

Finally, using a weighted average we ended up with our esteem:

DUPLICATES	Accuracy	F1	Precision	Recall
It-core-news-lg	0.88	0.88	0.88	0.88
WikiNEuRal	0.92	0.90	0.90	0.90

NODUPLICATES	Accuracy	F1	Precision	Recall
It-core-news-lg	0.89	0.89	0.89	0.89
WikiNEuRal	0.91	0.89	0.89	0.90

Tables n. 6-7 - An example of the output of the inference model with *WikiNEuRal* and *It-core-news-lg*

NOTE: these metrics are the results of extremely strong hypothesis and don't take into consideration the entities not found by the models and the not common entities as well. Said that, it seems that *WikiNEuRal* might be a better choice between the two models, however, the ensemble is a safer choice due to the high uncertainty regarding the metrics.

Conclusions

AI tools appear to be a valid support for the Cataloguing Process of Cultural Heritage through the automatic extraction of information from different types of files: images, videos, texts.

This additional information allows enriching the metadating of cultural heritage through, for example, more complete and detailed descriptions. It also allows to improve the search for catalogued resources and the identification of similar assets using new search keys never used before. However, AI algorithms do not always perform well. For example, they can provide (clearly) incorrect results (to the human eye) even on simple tasks when they apply to contexts totally different from those they learned in the training phase (unforeseen situations). They do not admit that they do not know how to answer, so they tend to invent when they do not find a plausible answer (a phenomenon called hallucination).

Finally, they do not know the common sense typical of human beings, understood as the combination of causal thinking, intuitive psychology, physics and intuitive sociality.

Gerd Gigerenzer¹⁵ in his book "Why human intelligence still beats algorithms" compares the behaviour of AI to that of small children and underlines how, unlike the latter, it is not able to independently build its causal model of the world, not exploits the feelings and intentions of people and is unable to acquire the basic information typical of experience if it is not taught in detail that, for example, time cannot be reversed or that a solid cannot pass through another solid (a lesson which children learn every time they throw a toy against a wall...).

However, even though there is no perfect tool, no substitute and solution tool, AI tools are a sensational help and we can consider them as a superpower at our service!

Bibliography

- [1] Giorgio Pedrazzi, Donatella Sforzini, Gabriele Fatigati (2023), *AI for Art: A Look at Object Detection and Transformers*, in Antonella Guidazzoli and Maria Chiara Liguori (Eds.), *AI and Cultural Heritage. Between Research and Creativity Workshop proceedings*, Cineca, DOI: 10.1388/WORKSHOP-AICH-01.
- [2] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi, *You Only Look Once: Unified, Real-Time Object Detection*, arXiv:1506.02640v5, 2016.
- [3] <https://docs.ultralytics.com/guides/yolo-performance-metrics/>
- [4] Chien-Yao Wang, I-Hau Yeh, & Hong-Yuan Mark Liao, *Yolov9: Learning what you want to learn using programmable gradient information*, arXiv preprint arXiv:2402.13616, 2024.
- [5] Aqeel Anvar, *What is Average Precision in Object Detection & Localization Algorithms and how to calculate it?*, in *Towards Data Science*, May 13, 2022, <https://towardsdatascience.com/what-is-average-precision-in-object-detection-localization-algorithms-and-how-to-calculate-it-3f330efe697b>
- [6] Gerd Gingerenzer (2023) *Perché l'intelligenza umana batte ancora gli algoritmi*, pp.120-121

Further readings

Dasiopoulou, Stamatia, et al. "Knowledge-assisted semantic video object detection." *IEEE Transactions on Circuits and Systems for Video Technology* 15.10 (2005): 1210–1224.

Yuxin Wu, Alexander Kirillov, Francisco Massa, Wan-Yen Lo, & Ross Girshick. (2019). Detectron2.

Glenn Jocher, Ayush Chaurasia, Alex Stoken, Jirka Borovec, NanoCode012, Yonghye Kwon, Kalen Michael, TaoXie, Jiacong Fang, imyhxy, Lorna, 曾逸夫(Zeng Yifu), Colin Wong, Abhiram V, Diego Montes, Zhiqiang Wang, Cristi Fati, Jebastin Nadar, Laughing, Mrinal Jain. (2022). ultralytics/yolov5: v7.0 - YOLOv5 SOTA Realtime Instance Segmentation (v7.0). Zenodo. <https://doi.org/10.5281/zenodo.7347926>

Matthias Minderer, N. (2022). *Simple Open-Vocabulary Object Detection with Vision Transformers*. ECCV.

Sechidis, K., Tsoumakas, G., Vlahavas, I. (2011). *On the Stratification of Multi-label Data*. In: Gunopulos, D., Hofmann, T., Malerba, D., Vazirgiannis, M. (eds) Machine Learning and Knowledge Discovery in Databases. ECML PKDD 2011. Lecture Notes in Computer Science(), vol 6913. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-23808-6_10

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. & Polosukhin, I. (2017). *Attention is All you Need*. In I. Guyon, U. von Luxburg, S. Bengio, H. M. Wallach, R. Fergus, S. V. N. Vishwanathan & R. Garnett (eds.), Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA (p./pp. 5998--6008) .

Simone Tedeschi, Valentino Maiorca, Niccolò Campolungo, Francesco Ceconi, and Roberto Navigli. 2021. *WikiNEuRal: Combined Neural and Knowledge-based Silver Data Creation for Multilingual NER*. In Findings of the Association for Computational Linguistics: EMNLP 2021, pages 2521–2533, Punta Cana, Dominican Republic. Association for Computational Linguistics.

Honnibal, M., & Montani, I. 2017. *spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing*.

Daiber, Joachim, et al. 2013. *Improving Efficiency and Accuracy in Multilingual Entity Extraction*. Zenodo (CERN European Organization for Nuclear Research), 4 Sept. 2013.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. 2019. *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. arXiv [Cs.CL].

°°KOBİ: A KNOWLEDGE ECOSYSTEM FOR RESEARCH AND EDUCATION

Franco Ripa di Meana¹, Alberto Giretti², Andrea Guidi¹, Massimo Vaccarini², Matteo Zambelli³

¹*Accademia delle Belle Arti di Roma, Via di Ripetta, 222 – 00186 Roma, Italy*

²*Università Politecnica delle Marche, Via Brecce Bianche, 60131 Ancona, Italy*

³*Università degli Studi di Firenze – Dipartimento di Architettura – DiDA, Via della Mattonaia, 8, 50121 Firenze, Italy*

Abstract

This paper presents an overview of the latest developments in °°KOBİ, a platform aimed at supporting creativity in art and design education. °°KOBİ utilises Large Language Models (LLMs) to facilitate an advanced exploration of collective knowledge, offering users a nuanced, contextual comprehension of artistic concepts derived from vast databases of artistic literature. This integration is indicative of the wider academic movement towards employing AI technologies to enhance educational paradigms. Significant features to °°KOBİ encompass the use of AI-powered voice interactions, AI-assisted conceptualization, and the three-dimensional visualisation of knowledge organised by semantic proximity. A further use case of °°KOBİ takes advantage of augmented reality technology, which provides a dynamic, multimodal experience within a mixed-reality environment. This form of interaction is grounded in the principles of experiential and embodied learning theories, advocating that learning achieves its highest efficacy when it engages multiple senses and physical interaction.

Introduction - What is °°KOBİ?

We are happy to be guests of the CINECA VISIT LAB and thank Antonella Guidazzoli and Maria Chiara Liguori for this opportunity to present °°Kobi. But what is °°Kobi?

°°Kobi is a prototype of a digital knowledge ecosystem for research and education in the field of arts and design. °°Kobi was born within a research project, which for a place like CINECA is the bread and butter, but the research project within which °°Kobi was born had a particularity: it was the only HORIZON project dedicated to artistic research, and in particular to the assertion of the role and status of artistic research as a specific way of creating knowledge: [EU4ART-differences](#).

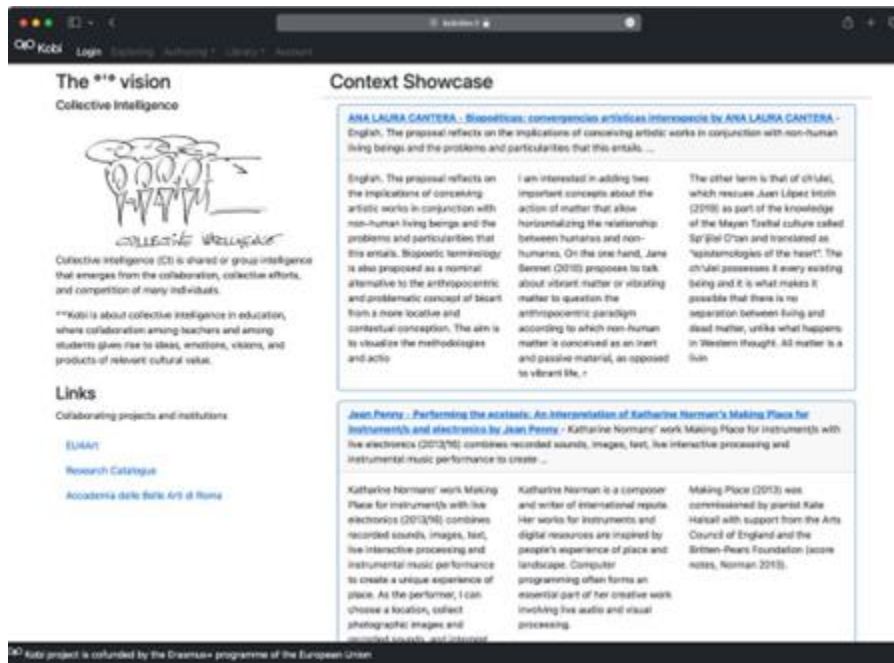


Fig. 1 – A view of °Kobi web portal.

The field of artistic research, which differs from artist research or from stylistic research, is a recent disciplinary field [1], which started to develop around the 2000s, and is still in great ferment [2], far from stabilisation and normalisation. In recent years there have been significant institutional developments - such as the Vienna Declaration on Artistic Research [3] or the changes proposed to the Frascati Manual [4] - and the emergence of discipline-specific tools, such as the [Research Catalogue](#), the most important digital platform for artistic research.

What we are interested in saying in this context is that it is only when artistic research claims its own autonomy, that transdisciplinary research becomes possible, i.e. research in which all components have equal dignity as bearers of their own specific methodology.

°Kobi is transdisciplinary research.

We had to discover a common ground, from which to develop a shared methodology, thus achieving results that none of the single participants would have reached on their own. Often in transdisciplinary confrontation, each player tends to have an outdated image of the others. Scientists have an idealistic conception of art, anchored in the 19th century, conceiving artistic process as intuitive, linked to individual genius and inspiration. In turn, artists have a 19th century idea of science, for example a stubborn Newtonian image of physics.

But when we give each discipline equal dignity and bring into play current conceptions of making art and doing science, then we can find a common methodology, select points of contact and, perhaps even more interestingly, points of crisis, where the disciplines are irreducible.

Like all true research, then, our research can fail. In the meantime, we have come a long way. And to make this journey we have chosen as our starting point, as our common centre, a specifically and uniquely human faculty: creativity. A faculty that defines the human being, and to the human being specifically belongs.

°°*Kobi is a research project on this common centre, on creativity.*

But research is not an automatic, disembodied process, research is done by individuals, and so we also present ourselves to you as a group of individuals who have found a possible form of collaboration, starting from far apart. It is fair to say that our collaboration was paradoxically facilitated by the fact that each of us, within our own discipline, has a heterodox, borderline position.

We are a director, as such an artist with a question mark, who additionally looks at art as process and method; an engineer who looks at engineering as a possibility to investigate creative mechanisms; an architect who is more interested in the methods of designing than in the result; and finally a musician and a second engineer who both work on boundaries, on the interfaces that make dialogue and exchange possible.

We pooled our knowledge, which became the collective knowledge of our group, greater than that of the individual members. This enabled us to achieve something that none of the components could have achieved alone.

And this is precisely the second axis of °°Kobi: collective intelligence, the intelligence of knowledge communities.

When we set to work to build the first demo of our system, which we presented for the first time during the European Researchers' Night at the Academy of Fine Arts in Rome on 29 September 2023, and which we are presenting to you today, we needed a qualified, rich knowledge base, and we were lucky to establish a collaboration with the Research Catalogue.

The Research Catalogue is itself a research project [5], born out of an investigation into what it means to publish artistic research, often made up of composite materials, many of which are non-verbal. A publication tool, which created its own specific practice of the *exposition* (as opposed to the exhibition) [6], investigating the differences between documenting the artistic process and

publishing research. The Research Catalogue offered itself as a tool at the service of individual artistic researchers and the institutions where artistic research takes place, the Academies, Conservatories, Academies of Dance and Performing Arts, accumulating over time a significant knowledge base in the discipline of artistic research.

More so, the Research Catalogue hosts a series of Journals, similar to those you are familiar with, namely peer-reviewed journals [7], accompanied by rigorous reflection on the significance of this practice in the field of artistic research, such as the Journal for Artistic Research. °°Kobi has been structured thanks to the access to the Research Catalogue, developed and managed by the [Society for Artistic Research \(SAR\)](#), which partnered with our research.

The realisation of the demo closed the circle, developing the third axis of °°Kobi: the relationship with technology, in particular with Artificial Intelligence and with embodiment linked to augmented reality.

AI in Artistic Research

Through experimental studies, our research investigates how the interactions with AI can trigger divergent thinking and foster knowledge engagement [8] in artists' creative practice. By analysing participants' engagement with AI within a structured knowledge ecosystem, our current research aims to assess the potential of AI to understand and enhance creativity in art and design. This research also contributes to the broader discourse on the implications of AI for educational practices, highlighting the need for innovative pedagogical models that embrace technological advancements in artificial intelligence while preserving the core values of artistic education.

In the era of Artificial Intelligence, art and design education faces transformative opportunities and challenges. Education in such areas traditionally relies on experiential learning and the development of creative skills through hands-on practice and heuristic exploration [9]. Learning in art and design not only fosters imaginative capabilities such as visualisation and originality but also immerses students in the rich tapestry of cultural history [10] [11]. The pedagogy of art education, deeply rooted in craftsmanship and apprenticeship, balances the mastery of conventions with the need for creative expression and innovation [12].

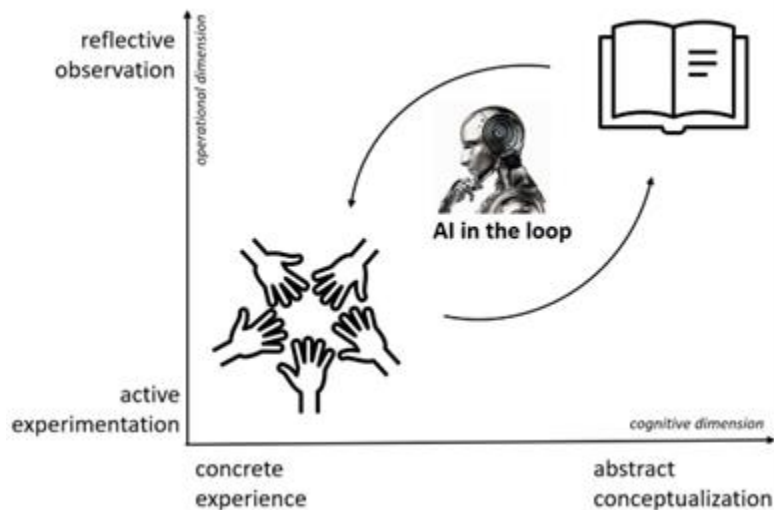


Fig. 2 – Placing AI in the creative loop, °°Kobi aims to harmonise the cognitive and operational dimensions in the context of artistic and design practices

Recent advancements in AI, particularly in Generative AI and Large Language Models (LLMs), present novel opportunities for enhancing the creative process in education. These technologies, capable of generating contextualised text and supporting idea generation, offer potential benefits for creative educational environments [13] [14] [15] [16] [17]. Integrating AI into art and design education could foster imaginative thinking and expand interpretative skills, grounded in a multicultural perspective. However, effectively incorporating AI technologies poses significant pedagogical challenges, requiring innovative approaches to teaching and learning that complement traditional artistic practices [18].

Divergent Thinking in Art Education

The By integrating °°KOBİ and AI into the creative process, we specifically aim to investigate how such technological tools can stimulate divergent thinking among art students. Our focus on divergent thinking is related to its crucial role in enabling individuals to generate a multitude of solutions and ideas, thereby expanding the boundaries of creative expression. Through °°KOBİ, we explore the potential of AI to act as a catalyst for divergent thinking, providing a rich, interactive environment that encourages users to explore a wide range of creative possibilities and conceptual frameworks. Divergent thinking is a pivotal aspect of creativity, particularly within the realm of art and design education [9]. This form of thinking is characterised by the ability to generate multiple unique solutions to a given problem, emphasising non-linear reasoning and the exploration of

many possible solutions. It stands in contrast to convergent thinking, which focuses on finding a single, correct answer to a problem. Divergent thinking is essential for the creative process as it fosters originality, flexibility, and innovation. Education in art and design, with its inherent emphasis on experiential learning and the development of creative skills, cultivates an environment where divergent thinking can thrive [10] [11].

To stimulate divergent thinking, °°KOBİ integrates a knowledge ecosystem. In such ecosystems, the dynamic interplay between knowledge, creativity, imagination, and cultural insights forms the basis for educational practices that promote divergent thinking [19]. This approach emphasises the role of AI in facilitating a dialogical and reflective learning process, where students engage critically with diverse cultural and knowledge contexts. Faced with the challenge of constructing a knowledge-based foundation for °°KOBİ, we leveraged the Research Catalogue repository.

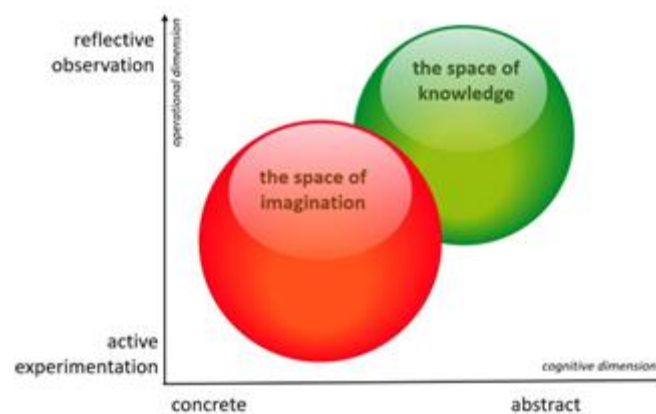


Fig. 3 – °°KOBİ -leverages AI to facilitate the interplay between imagination and knowledge exploration fostering divergent thinking for art and design.

The Research Catalogue (RC) provides an online platform for both individual researchers and institutions to publish artistic research contributing to a growing body of qualified knowledge. Drawing from this rich collection of art-focused articles, enhanced by the incorporation of peer-reviewed journals and rigorous academic discourse, our project has crafted a system that embodies the convergence of artistic research and technology.

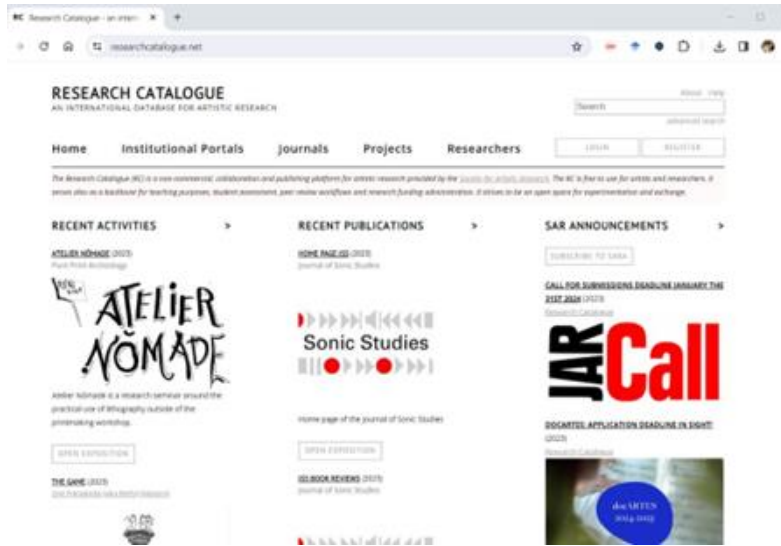


Fig. 4 – Research Catalogue web portal.

The °°KOBİ System

The name °°KOBİ derives from the Greek term "koinòs bios," which translates to a communal or shared life, symbolising the system's objective to harness AI in fostering collective intelligence within communities of practice. This iteration of °°KOBİ is specifically engineered to construct and facilitate the navigation of semantic spaces formed from a wide array of knowledge sources. This development leverages content from the Research Catalogue (RC) that is an online database designed specifically for the documentation, publication, and archiving of artistic research. It serves as a digital platform where artists, researchers, and academics can create 'expositions' to showcase their artistic research projects. These expositions allow for the integration of text, images, audio, and video elements, offering a rich multimedia experience that traditional academic publications might not accommodate.

°°KOBİ generates its semantic space by transforming data from Research Catalogue into a proprietary database schema, categorising information into contexts, media (text, images, video, music), concepts, and keywords. These elements are linked through contextual belonging and semantic closeness, essentially crafting an "imagination" or "idea" space where potential model outcomes are spatially organised. The integration of new knowledge sources is facilitated by specialised crawlers, for instance, the RC crawler, which maps web pages to semantic contexts,

further enriching the system's semantic architecture. Notably, this process abstracts the original structure, presenting content based on contextual association and semantic links to foster more flexible AI interactions.

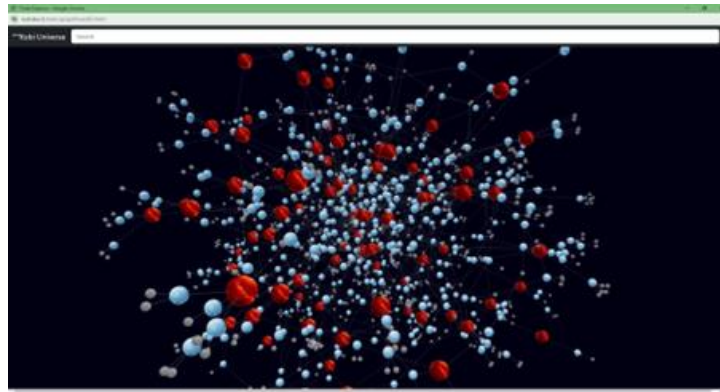


Fig. 5 – A view of the semantic space. Each sphere represents a publication (red), a concept (light blue) or a keyword (gray).

At the heart of °°KOBİ lies three servers: a knowledge server hosting a graph database via Neo4j, a language server for performing translations and encoding textual data through open-source LLMs, and a multimedia server dedicated to video and music content management. A Communication Manager, through an AI Gateway, manages the system's interactions, employing the OpenAI GPT4 engine for nuanced AI queries. This setup supports two user interfaces: a conventional web interface and an augmented reality interface, both designed to offer a uniform user experience across varying levels of immersion.



Fig. 6 – A view of the augmented reality experience °°KOBİ. Media, such as images and texts, resulting from queries to the AI are visualised as interactable holograms.

Users engage with ^oKOBI through a straightforward chat interface, initiating dialogues that lead the AI to scout the semantic space for closely related nodes to the discussion topic. These nodes not only provide direct responses but also shape the conversational context with the AI. Leveraging the artistic contents from the Research Catalogue, the system often yields insights with unique semantic interpretations, potentially stimulating creative thought. Users can explore this enriched semantic landscape visually in the web interface or through voice and gesture commands in an augmented reality setup, offering an immersive and interactive knowledge discovery journey. Future enhancements may explore AI's capacity to create multimedia content tailored to the user's semantic positioning, potentially deepening the system's creative utility and user engagement.

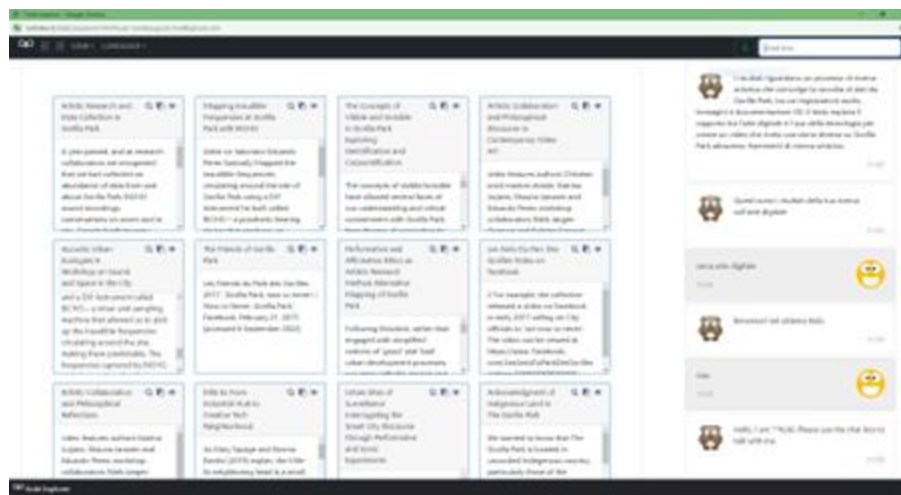


Fig. 7 – KOBI AI conversational agent chat.

Research Studies

The research undertaken thus far includes an initial demonstration and two subsequent experiments. The demonstration unfolded at the European Researcher's Night in 2023, followed by two studies executed at the Academy of Fine Arts in Rome. In a preliminary phase, encapsulating the demonstration and the first experimental study, we concentrated on evaluating the system's usability and refining both the experimental setup and protocol. Subsequently, the second phase, incorporating the latter study and ongoing investigations, seeks to examine a foundational hypothesis: the presence of divergent thinking during creative activities can be inferred and observed through the analysis of textual interactions with AI. Early results, which

will be discussed in more detail in this document, offer provisional yet optimistic evidence in support of the hypothesis.



Fig. 8 – A participant using °°KOBi during the Artistic Research Days in Rome.

1) The demo at the Artistic Research Day

The first evaluation of our system was conducted during the Artistic Research Days at the Rome Fine Art Academy in late September 2023. This stage aimed to explore the initial reactions and adaptability of users to the system, utilising a mixed reality interface for its innovative appeal. Participants, a varied group of educators, researchers, students, and artists received an introduction to the system's main features, such as search, chat, and the knowledge graph. They interacted with the system in ten-minute sessions, during which we collected data on their use, gathered their personal assessments, and conducted informal interviews.

Feedback indicated that participants quickly became comfortable with the system, taking between three to six minutes to familiarise themselves with its functionalities. While the usability of command, navigation, and search features was rated highly, the chat function was found to be less intuitive. This highlighted the need for enhancements in chat interactions, particularly in improving the interface's voice recognition capabilities and distinguishing between the user's voice and background noise.

Overall, participants expressed a high level of engagement and satisfaction with their tasks, reflecting the system's user-friendly and compelling design. A statistical comparison between the observations of researchers and the self-evaluations of participants showed no significant discrepancies, indicating consistent perceptions of performance.

Moreover, the study uncovered moments of significant creative divergence during interactions with the AI, suggesting the system's potential to inspire innovative artistic concepts. However, it also noted that unexpected AI responses could sometimes disrupt the flow of conversation. These findings emphasize the need for further refinement of the system and highlight its capability to support creative processes effectively.

2) The Keywords Study

The objective of this study was to observe divergent thinking through their interactions with AI, as reflected in chat logs. In this study, twelve art students at the Rome Fine Arts Academy were tasked with associating keywords and qualifying such associations with a short description under two separate conditions without AI and using °°KOBİ AI assistant. In the second condition, the setup allowed participants to seek insights from °°KOBİ on the given keywords, ensuring a well-rounded briefing on the system's functionalities for optimal engagement during the experiment.

Data analysis revealed a tendency towards convergent thinking in most participants, who primarily reinterpreted the given keywords without significant semantic deviation. However, in a few cases, a departure from the expected semantic fields was noted, indicating instances of divergent thinking as participants explored beyond the provided concepts based on AI suggestions.



Fig. 9 – The form used in the Rome Fine Art Academy Test contains a set of concepts forming a circle. The subject is asked to fill in and to qualify with short descriptions of as many relationships among the concepts as possible using AI assistance.

A metric was developed to quantify these observations, employing the Euclidean distance within the AI's latent semantic space to measure the divergence of user queries from the task's keywords. This metric highlighted instances where participants' inquiries significantly deviated from the initial semantic framework, suggesting a potential for divergent thinking facilitated by AI interaction.

We speculate that this approach lays the groundwork for a methodology dedicated to observing and analysing divergent thinking. By leveraging AI as a strategic tool, contextualised within a qualified knowledge base, we aim to create the conditions necessary for probing divergent thinking as a pivotal aspect of creativity. Through this approach, AI's contextualised responses, rooted in a rich and nuanced knowledge repository, serve as a catalyst for exploring the depths of creative thought and its manifestations, beyond subjective assessments of creative output to a more objective analysis of divergent thinking during the ideation process.

Conclusions

In the research described in this paper, we are exploring the innovative intersection of Artificial Intelligence in art education through the development and application of °°KOBi. Our study uses °°KOBi's as a research probe to query creative processes, especially related to divergent thinking, by leveraging Large Language Models (LLMs), and a qualified knowledge base, facilitating a deeper engagement with collective knowledge and artistic concepts.

Our experimental investigations into °°KOBi's use in educational settings have revealed promising insights into its effectiveness for observing divergent thinking and knowledge engagement among art students. Through detailed analysis of interactions with °°KOBi, we observed the possibility for AI to potentially act as a catalyst in the creative process, enabling students to explore a broad spectrum of creative possibilities and conceptual frameworks.

The findings from our research indicate a promising potential for AI to support and enhance creativity in art and design education. Data from further study will possibly confirm such a hypothesis. °°KOBi's integration into the learning environment has provided a novel platform for creative exploration but also provided the role of AI in facilitating a more interactive and reflective learning experience. This aligns with the broader academic discourse on the importance of integrating advanced technologies in educational practices, recognizing AI's role in evolving and enriching artistic research and creativity.

As we conclude this paper, it's clear that our exploration into the role of AI in art and design education is just beginning. The positive outcomes from our study with °°KOBi encourage further research into the complexities of creativity and how AI can continue to transform educational paradigms. Future studies will undoubtedly build upon our findings, delving deeper into the dynamics of creative thinking and the potential of AI to unlock new avenues for artistic expression and innovation.

Acknowledgments

The authors wish to thank SAR (Society for Artistic Research) for granting access to the Research Catalogue materials.

References

- [1] Borgdorff, H. (2012). *The Conflict of the Faculties. Perspectives on Artistic Research and Academia*. Leiden: Leiden University Press.
- [2] De Assis, P. & D'Errico, L. (Eds) (2019) *Artistic Research. Charting a Field in expansion*. Lanham, MD: Rowman & Littlefield
- [3] *Vienna Declaration on Artistic Research*, 2020 <https://societyforartisticresearch.org/wp-content/uploads/2020/10/Vienna-Declaration-on-Artistic-Research-Final.pdf>
- [4] OECD (2022) *Proposed changes to the Frascati Manual* https://aec-music.eu/media/2022/06/Cover-Letter-to-OECD_NESTI.pdf
- [5] Borgdorff, H. A. (2020). *Cataloguing artistic research: the passage from documented work to published research*. In P. Peters & T. Pinch (Eds.), *Dialogues between Artistic Research and Science and Technology Studies* (pp. 19-30). New York: Routledge.
- [6] Schwab, M.& Borgdorff, H.A. (Eds.) (2014) *The exposition of artistic research: Publishing art in academia*. Leiden: Leiden University Press.
- [7] Schwab, M. (2018). *Peer reviewing in the Journal for Artistic Research*. In Ysebart, W. & Van Kerckhoven, B. (Eds): *Evaluating art and design research: Reflection, evaluation, practices and research presentations* (pp.52-59). Brussels: VUB Press.
- [8] Kanli, E. (2020). *Assessment of creativity: theories and methods*. Creativity-A Force to Innovation.
- [9] Tanggaard, L. (2014). *A situated model of creative learning*. *European Educational Research Journal*, 13(1), 107-116.
- [10] Greene, M. (1970). 'Imagination', in Smith. R. A. (ed.) *Aesthetic Concepts in Education*, Urbana, Chicago, London: University of Illinois Press.
- [11] Addison, N. et al, 2010. *Understanding Art Education*. Taylor and Francis, NY.
- [12] Csikszentmihalyi, M. (1996). *Creativity: Flow and the psychology of discovery and invention*, New York: Harper Perennial
- [13] Berns, S., & Colton, S. (2020). *Bridging Generative Deep Learning and Computational Creativity*.

- [14] Bown, O. (2012). *Generative and Adaptive Creativity: A Unified Approach to Creativity in Nature, Humans and Machines*. In J. McCormack & M. d'Inverno (Eds.), *Computers and Creativity* (pp. 361–381). Springer. https://doi.org/10.1007/978-3-642-31727-9_14
- [15] Dwivedi, Y. K., et al. (2023). “So what if ChatGPT wrote it?” *Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy*. *International Journal of Information Management*, 71, 102642. <https://doi.org/10.1016/j.ijinfomgt.2023.102642>
- [16] Hughes, R. T., Zhu, L., & Bednarz, T. (2021). *Generative Adversarial Networks–Enabled Human–Artificial Intelligence Collaborative Applications for Creative and Design Industries: A Systematic Review of Current Approaches and Trends*. *Frontiers in Artificial Intelligence*, 4. <https://www.frontiersin.org/articles/10.3389/frai.2021.604234>
- [17] Zhang, Y., Sun, J., Yang, Z., & Wang, Y. (2018). *Mobile social media in inter-organizational projects: Aligning tool, task and team for virtual collaboration effectiveness*. *International Journal of Project Management*, 36(8), 1096–1108. <https://doi.org/10.1016/j.ijproman.2018.09.003>
- [18] Glăveanu, V. P. et al., (2021). *Creative Learning in Digital and Virtual Environments. Opportunities and Challenges of Technology–enabled Learning and Creativity*. New York and London: Routledge.
- [19] Morris, T. H. (2020). *Experiential learning—a systematic review and revision of Kolb’s model*. *Interactive Learning Environments*, 28(8), 1064-1077.

COPIRIGHT AND WORKS CREATED BY ARTIFICIAL INTELLIGENCE, TECHNICAL TRIALS OF PROTECTION: APPROACHES AND POSSIBLE SOLUTIONS

Valentina Gastaldo - Università di Parma

Introduction

Artificial Intelligence (AI) is one of the most challenging, controversial, and potentially disruptive technologies. While it has been thoroughly integrated into our daily lives, it also presents significant social, cultural, legal, and ethical issues. This is because some AI applications, particularly Generative Adversarial Networks (GANs), can emulate the human mind and its capabilities, creating works in literary (such as Shelley Ai, an algorithm that collaborates with users to write short horror stories, and Xiaoice, an artificial poet), musical (the 2016 song "Daddy's Car" created by Sony's Flow Machines researchers), and artistic fields (8 by Yoox, the first collection created by software after analyzing images, trends, tastes, and purchasing habits on social media and online magazine content), as well as numerous other areas previously considered the sole domain of humans.

AI challenges our unique human characteristics by demonstrating that many tasks, traditionally performed by humans, can be equally well executed by machines. Artificial intelligence can have many meanings; it can be understood as any algorithmic system displaying intelligent behavior, analyzing its environment, and taking actions autonomously to achieve specific goals. However, there is no single definition, but what is certain is that it is a constantly evolving technology [1]. It is defined as "generative" because it improves, learns from itself, and evolves rapidly with usage. The essential element of these computational techniques is that machines' work relies on the data available in the vast web world, collecting them through a process called scraping, which involves indiscriminate gathering of information.

The major novelty of these AI systems lies in the concept of machine learning. We are talking about machines capable of learning from what they read, perfecting themselves, and self-training to perform various tasks assigned to them by humans. Machines can impact our individual and collective lives with an efficiency and speed incomparable to human action.

Artificial Intelligence and Art

In the context of this vast topic, which has become a priority even on social and political levels (consider that this year's G7 focused on AI and saw the extraordinary participation of the Pope), affecting fundamental aspects of individual and collective life, such as justice (in the United States, where the Wisconsin Supreme Court uses the COMPAS algorithm - *Correctional Offender Management Profiling for Alternative Sanctions* - to assess an offender's risk of reoffending based on responses about work activity, social life, education, personal opinions, criminal history, etc., to decide on parole and bail), health (AI's use in medicine has not only improved the speed and accuracy of diagnoses but also enabled early detection of diseases), transportation (several car manufacturers are experimenting with autonomous mobility solutions, such as self-driving cars), and the environment (robotics in environmental monitoring allows for observing otherwise inaccessible areas through drones to analyze pollutants and study flora and fauna), the focus here is on AI's use in human creativity and art. AI techniques are proving particularly useful in this sector.

Computers have been producing numerous works since the 1970s, thanks to the creative input of programmers. Recently, we have seen tools capable of attributing authorship or forgery of works (Rutgers University researchers "trained" a machine to recognize about 300 works by Picasso, Matisse, Modigliani, and others. Through the Recurrent Neural Network (RNN), machines capable of ensuring increasingly precise conservation and restoration activities (such as the famous Ghent Altarpiece), recovering and restoring damaged paintings or hidden works (like Picasso's "lost woman" found under "The Old Guitarist" through X-ray and infrared analysis), and gaining deeper insights into artists and their work (Johannes Vermeer's "Girl with a Pearl Earring" has revealed details through scans and micro-imaging techniques).

In these cases, AI techniques only assist humans, and the results of this collaboration do not raise significant legal issues. The machine is a technical tool that the artist uses to create their work, concretize their ingenuity and creativity, or that scholars use to deepen their research.

Today, however, these technologies are being used to create texts, sounds, and images that can present artistic and creative content. Computer-generated art now finds expression in music, painting, poetry, and literature. The first books (such as "Non siamo mai stati sulla terra" by Rocco Tanica and Outomat-B13) have been published, written collaboratively with AI. Christie's auction house sold the portrait of Edmond Belamy (the first painting created entirely by AI) for

over \$400,000. AI has also produced a short film – [Sunspring](#) - that stood out in a science fiction short film competition.

We face scenarios where human contribution is minimal or even non-existent, and the product could be the result of machines' "imagination" using sophisticated software to combine images, colors, and sounds input by humans to create art similar to that produced by human painters, musicians, or writers.

In this context, where the product is created almost autonomously by “thinking” machines through algorithms and data that serve only as a starting point, AI techniques risk overturning and questioning the current regime and discipline of intellectual property and copyright, designed to reward creativity and promote cultural production and dissemination. This forces us to rethink the interaction between machines and the creative process.

All these experiments share the same question: whether it is legitimate and to what extent to talk about artistic creativity concerning machines, or whether creativity is an exclusively human characteristic. In such cases, who should be considered the author of the produced works? [2] The AI system generating the artistic content, or the programmer (human), providing the machine with the necessary inputs to perform its work in a partially autonomous way? Can a robot be held responsible if the created works are plagiarized from those made by third parties? What kind of protection should be assigned to these works? And can we imagine recognizing legal personality to AI systems?

The Regulatory Framework of Copyright Law in Italy and Worldwide

The issue of legal protection for works created by artificial intelligence (AI) is the subject of extensive debate, as there is currently no specific legislative framework in place. Existing copyright laws are inadequate to protect and value AI-generated results, and in my view, it is not possible to apply them analogically.

To address these questions, we must understand the requirements that a work must meet to benefit from copyright protection. Although the World Intellectual Property Organization defines intellectual property as “the creations of the mind”, without specifying that the mind must be human, Italian law seems to exclude the possibility of a machine being the author of a creative work. Article 6 of the Copyright Law (Lda) and Article 2576 of the Civil Code refer to the concept of “intellectual work”, suggesting that the author of intellectual property rights must be a natural person. Article 25 (Lda) provides further guidance by stating that “the economic

rights of the work last throughout the author's life and until the end of the seventieth year after their death". Similarly, Article 2580 of the Civil Code, by attributing authorship solely to the author and their successors, seems to exclude the possibility of extending protection to other entities, such as a software creator, under Article 633/41.

Law 633 of 1941, in Article 1, specifies that "*creative works of the mind... regardless of their mode or form of expression... constitute intellectual creations of the author*" are protected. Article 2575 of the Civil Code uses similar language, invoking the personal creation aspect of the work. Therefore, originality, understood as the fruit of the author's creative intellect, is the general criterion by which a work can be considered protected, granting the author both moral and economic rights. The creative nature of the product justifies the recognition of an exclusive monopoly over the work, as it reflects the personality and ingenuity of the author.

The concept of creativity is thus central, not only in the Italian legal system but also at the European and international levels. In France, for a "*work of the mind*" to be protectable, it must express the author's personality. The Copyright Law Act describes "*anonymous works*" as those where no natural person is identified as the author. If this is the indication for identifying such works, it can be presumed that the author must necessarily be a human being.

This position is also supported by the U.S. Copyright Office Manual, which explicitly states that only human-created works can be protected by copyright. Works produced by a machine or a mechanical process that operates randomly or automatically, without any creative input or intervention by a human author, will not be registered. This stance is based on judicial precedent (*Naruto v. Slater*, No. 16-15469, 2018 WL 1902414 (9th Cir. 2018); *People v. Frazier*, 2009 WL 1842666 (Cal. Ct. App. 2009), confirmed in more recent rulings, which affirm that copyright law protects only "*the fruits of intellectual labor that are founded in the creative powers of the mind*".

A similar position has been adopted by EU law, which indirectly addresses this issue in various directives, implying that copyright is granted only to human creators. For example, the 2009 Software Directive specifies in Article 2.1 that "*the author of a computer program is the natural person or group of natural persons who created the program or, where permitted by national legislation, the legal person designated as the rights holder*". The 2006 Directive on the term of copyright protection states in Article 1.1 that "*the rights of authors of literary and artistic works within the meaning of Article 2 of the Berne Convention shall run for the life of the author and for 70 years after their death*".

In common law, a work receives protection if it is the creative work of an author, thus requiring a human subject to create it. The only exception so far is Ukraine, which has introduced specific

legislation on AI-generated works, granting the owner or licensee of the software that generated the work a copyright-like economic right lasting twenty-five years.

The reasons behind these conclusions are not only based on the literal wording of the provisions but also on the fact that human intervention, even if not creative, is the basis of any artificial intelligence, thus limiting and defining its apparent autonomy and self-sufficiency. An AI system is not self-sufficient in deciding what to do, as it relies on information inputted by humans. Therefore, machine-generated creativity will always result from the reworking of the initial data underlying the algorithm. In other words, creativity can never be attributed to AI alone but, rather, to the human-machine team.

In this perspective, it should be considered that, at least for now, AI cannot be regarded as a “creator” of works of art and, therefore, cannot hold any copyright. This is because human creative activity is inherent in the very concept of an artistic work, in line with the contemporary (anthropocentric) conception of copyright as an expression of a human right.

National and International Jurisprudence on Copyright for Works Created by (or with the Assistance of) Artificial Intelligence

Since neither the domestic legal systems of states nor the European Union (the AI Act itself does not address copyright, except very marginally; references to copyright are few and very generic. It is concerned with the safety and regulation of artificial intelligence broadly), currently provide any legislation regarding the impact of AI on art, it is interesting to examine the attitude of the jurisprudence on this matter. Given that issues related to generative artificial intelligence are relatively new, it is not surprising that jurisprudence is not yet fully developed.

Two main orientations can be highlighted. In some cases, copyright has been denied because the "creator" is not human. For example, in the United States, the Copyright Office has expressly stated that copyright is granted to an original work only if it has been created by a human being. Among the decisions (as early as 1991, the U.S. Supreme Court in *Feist Publications, Inc. v. Rural Telephone Service Co.* clarified that originality is the *conditio sine qua non* of copyright. Therefore, only works of ingenuity that exhibit a “spark of creativity” are worthy of protection) where copyright protection has been denied to works created with artificial intelligence systems, there is the recent order of August 18, 2023, by the District Court of Columbia (which confirmed the decision of February 14, 2022, made by the Copyright Review Board of the U.S. Copyright Office). The order once again rejects the request to register the artwork titled “A Recent Entrance to Paradise” by the AI “Creativity Machine” of Steven Thaler. This is the only

case of AI-generated work where the applicant claimed that the work was created automatically by a computer algorithm running on a machine, rather than through artificial intelligence (AI-assisted works). In any case, according to the owner, there was a right to register it because it was a commissioned work. The Court denied artificial intelligence the ability to be recognized as the author of an artwork due to the lack of human input and thus intellectual creativity that the law – for now – reserves to physical persons. In particular, the Commission first recalled that the Office does not register works created by a machine or a process that moves without any creative input or intervention by a human author, as it would be contrary to the law. It then added that a work can be considered commissioned if it was created by an employee or a person with whom a written assignment was made. Since AI systems cannot enter into contracts, they cannot be considered to create on behalf of someone else.

Australian courts have taken a similar approach (see in particular the Case *Achos Pty Ltd v. Ucorp Pty Ltd*, 2012, 201 FCR 173 where the Australian Federal Court deemed that technical sheets created by a software could not deserve protection since the human author was not sufficiently involved. Case *Infopaq International AS v. Danske Dagblades Forening*, C-5/08, July 16, 2009), that excludes copyright protection if the author of the work is largely the machine.

For the Court of Justice of the European Union, the criterion for access to copyright protection is the requirement of originality, as an expression of the intellectual creation of its author. Thus, in the *Cofemel* ruling, it is confirmed that the work must materialize in an original object, an intellectual creation of its author, and must reflect their personality, manifesting their free, creative, and personal choices. Subsequently, the Court emphasized the need for the work to be an “expression of the intellectual creation of the author” to receive copyright protection, thus adding to the requirement of originality the necessity of human input, or what U.S. law qualifies as authorship (Case *Infopaq International AS v. Danske Dagblades Forening*, C-5/08, July 16, 2009). The orientation expressed by the Court was also taken up in a subsequent 2011 case, *Eva-Maria Painer v. Standard Verlags GmbH et al*, C-145/10).

In Italy, the Court of Cassation, in ruling No. 1107 of January 16, 2023, examined the issue of authorship of a work generated through the use of artificial intelligence systems and the possibility of protecting it under copyright law. The litigation originated from a lawsuit concerning the unauthorized use by a television broadcaster of a work by an actor as a permanent set design for the 2016 Sanremo Festival. The actor, claiming a violation of his copyright, sued the defendant before the Genoa Court and sought damages. The defendant argued that the work in question, mistakenly attributed to the actor’s ingenuity, was nothing

more than a digital image with a floral theme, characterized by the repetition of shapes at various scales, created using software that processed the form, colors, and details through mathematical algorithms. The alleged author only chose the algorithm to apply and approved the result generated by the computer.

The crucial issue concerns the determination of human creative input in the generative process. According to the Court, it is necessary to assess how the use of AI software has been integrated with human creativity in the development of the work. Recently, the U.S. Patent and Trademark Office issued guidelines (Inventorship Guidance for AI-Assisted Inventions, effective February 13, 2024 - issued following the Dabus case (Thaler v. Hirshfeld, 558 F. Supp. 3d 238 and Thaler v. Vidal, 43 F.4th 1207, 1213, cert denied, 143 S. Ct. 1783) and the Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence of October 30, 2023) stating that it is not possible to exclude *a priori* that an invention created by an artificial intelligence system is patentable. However, there must be a significant human contribution. Hence, the problem is to determine whether the contribution is significant and thus deserving of inventor status. In this regard, the Court refers to the so-called Pannu factors, where “a joint inventor must contribute in a significant manner to the conception or reduction to practice of the invention, make a contribution to the invention that is not insignificant, and do more than explain well-known concepts or the current state of the art.”

AI-created works might, therefore, not be protectable if they are the result of an automated decision-making process by the algorithm, unless a significant human creative contribution is demonstrated. Thus, measuring human creative input in the generative process becomes important. In other words, the boundary beyond which copyright cannot be assigned is found in the impossibility for the person to control the creative result of the machine [3].

However, it will not be easy to prove that a work is the result of human creativity rather than the intervention of the algorithm and training data. It will involve case-by-case assessment with a specific technical examination of whether human input was decisive in generating that result. To this end, it may be useful to retain evidence of the entire generative process, track and document the various phases, as well as attach a document to each work detailing how it was generated.

In other jurisdictions, while always excluding the possibility of granting copyright to AI, a different approach has been taken in addressing the issue of creations generated by artificial intelligence.

In the United Kingdom, Ireland, India, Hong Kong, New Zealand, and South Africa, in the absence of a human author, the creative potential – and thus copyright (the duration of the right

has been set at fifty years from the creation of the work) – has been recognized to the individual who provided all the elements necessary for the machine to generate the work. Thus, the programmer in the case of AI that is fully autonomous in the output production process; while in the case where AI follows the user’s instructions in the creation process, the right would belong to the latter.

It should be noted, however, that despite this interpretation provided by judges, the position of the British system remains cautious. The Intellectual Property Office has published a study on the relationship between AI and intellectual property rights in which the Government, while stating that it does not currently wish to intervene regarding the protection of computer-generated works to include those created by artificial intelligence, does not exclude the possibility in the future.

China has spoken on several occasions about the rights to works generated by artificial intelligence. In the Shenzhen Tencent Computer System Co., Ltd. v. Shanghai Yingxun Technology Co., Ltd. ruling of November 25, 2019, the People's Court of Shenzhen Nanshan District ruled that news articles written by Dreamwriter, an AI writing assistant of the plaintiff, were protected by copyright law. The defendant had infringed the plaintiff’s copyright by copying the news article. The entire process in which humans develop an AI writing assistant and use it to automatically generate articles constitutes the creation of works. This was the first time a Chinese court accepted a copyright dispute involving works created by artificial intelligence and also the first time it affirmed copyright protection by acknowledging authorship.

In the Feilin v. Baidu case of April 25, 2019, the Beijing Internet Court was called upon to decide whether the publication, with minimal changes and without attribution, on a social network of a document written with the help of the Wolters Kluwer Database program constituted a copyright infringement. According to the Judges, the text created by artificial intelligence can be said to possess creativity, thus meeting the first requirement of the Chinese Copyright Law to obtain copyright. However, it was deemed that the human input, limited to entering keywords and pressing the search button, was not sufficient to assert that the work represented the expression of the person’s personality. Therefore, copyright could not be recognized.

The Beijing Internet Court on November 27, 2023, in the Mr Li v. Ms Liu case, affirmed that a photograph created through an artificial intelligence system is protected by copyright. The work is considered original as it represents the will of the person who guided the machine in producing the result. According to the Court, therefore, the conditions for recognizing copyright

are met because Mr Li's photograph falls within the scope of literature, art, and science. It is also original, has a form of expression, and is the result of intellectual work.

However, it should be clarified that the Judges focused on products created through AI rather than automatic artificial intelligence. The rulings confirm human intervention as a fundamental requirement for obtaining copyright, thus not deviating from the prevailing orientation; however, the importance of these decisions lies in the recognition of copyright.

From the analysis carried out so far, it is clear that the issue is far from resolved. The problem of the legal void of authorship, or rather the absolute lack of protection for artworks fully created by AI, remains.

Possible Solutions to the Lack of Protection for Works Created by AI

While we await regulatory actions from states, it seems prudent to explore ways to resolve the current impasse. On the one hand, the risk of not providing any protection to works created solely by machines could pose significant problems, both legally and economically. Indeed, the lack of recognition for works created autonomously by AI could result in such works entering the public domain, allowing anyone to freely use and reuse them. This scenario could deter investments in artificial intelligence projects due to the absence of revenue from intellectual property rights that typically compensate creators for their personal expression, taste, and ingenuity. Additionally, issues of unfair competition and consumer confusion could arise, as AI could produce works that closely mimic the style of famous artists at potentially lower prices than human-created works.

The British law might provide a solution to this issue. The UK's Copyright, Designs and Patents Act 1988 explicitly addresses works generated by AI, attributing authorship to the person who organized the machine's functions. Similarly, Ireland, Hong Kong, New Zealand, and South Africa have extended copyright protection to artificial creations, generally attributing authorship to the person who made the necessary arrangements for the work's creation or who determined the creation of the work, depending on the jurisdiction.

However, these issues are not yet fully resolved. A key question remains: who should be considered the provider of the necessary configurations for the AI's creative process? The individuals involved in creating and interacting with the machine are diverse: the programmer who inputs data and algorithms, the user, or the commissioning party who provides specific inputs to create the work.

Moreover, the law does not yet address the scenario where the software deviates from the provided commands, demonstrating autonomous decision-making independent of human creators. This situation could leave original and creative works without protection.

Awareness of this regulatory gap has led the European Parliament to urge the Commission to develop criteria and rules to address situations arising from AI activities. Notably, the European Parliament resolution of October 20, 2020, highlights the importance of resolving issues related to the protection and ownership of intellectual property for works fully developed by AI.

Furthermore, the European Commission's proposal for a Regulation "Regulation on a European Approach for Artificial Intelligence" dated April 21, 2021, aims to establish a comprehensive regulatory framework for AI. However, this proposal does not address the legal applicability of copyright laws to the field of robotics.

An alternative approach could involve indirect and mediated authorial protection for artistic works created by an algorithm, potentially outlined through coordination between the Berne Convention for the Protection of Literary and Artistic Works of 1886 and the World Intellectual Property Organization (WIPO) Copyright Treaty of 1996. The Berne Convention provides a broad definition of artistic works, encompassing any production in the literary, scientific, and artistic fields, regardless of the manner or form of expression. This definition does not explicitly exclude the possibility of considering AI-generated works as artistic works.

The WIPO Treaty (Article 4) explicitly protects computer programs, recognizing them as literary works under the Berne Convention (Article 2). It is evident that programs capable of creating artistic works could now fall under this protection. Since both the Berne Convention and the WIPO Treaty do not specify what constitutes an artistic work or an author, nor do they provide a definitive list, there remains an opportunity to ensure protection for the programmer as the author of the algorithm generating the artwork.

Such an interpretation could allow for mediated protection of AI-generated products through the protection of the authors of these algorithms. Based on the considerations so far, we must exclude, at least for now and based on a faithful interpretation of existing norms, the possibility that AI, as a machine, can hold rights to intellectual works. Therefore, it is hoped that both domestic and international law will make efforts to reconsider the protection systems for works created by artificial intelligence.

Who, then, are the beneficiaries of copyright for works created by machines today?

To avoid the negative technological and economic policy consequences of a lack of protection—which might deter inventors from starting projects due to the high costs of developing artificial intelligence—it's essential to envision a potential beneficiary of copyright, distinguishing between economic and moral rights.

As is known, national copyright law provides for both economic rights and personal rights arising from any artistic work. While the former grant the author the right to economically exploit the work in every form and manner, original or derivative, and to receive compensation for any type of utilization (Art. 12 LDA and Art. 2577 c.c.), the moral rights are recognized to protect the author's honor and reputation and are therefore inalienable, imprescriptible, and irrevocable.

The attribution of economic rights could be considered in favor of the creator of the machine—the one who sets up the functions and provides the information that will lead to the creation of the intellectual work, i.e., the owner of the AI.

However, the first hypothesis seems unconvincing. Indeed, the software programmer does not contribute to the actual creation of the result, merely generating the machine. It would be unreasonable to grant rights to someone who cannot foresee the results in advance, risking that this individual could potentially hold rights to all works created by the machine, even as a result of commands and functions entered by others.

The same person could, however, be granted rights over the programs created, based on Art. 2 of Law 633/1941, which includes "computer programs, in any form expressed as long as they are original as a result of the author's intellectual creation" among the works deserving of protection.

Attributing authorship to the organizer of the machine's functions—corresponding to the English approach—raises similar objections. It is difficult to definitively attribute the merit of the creation realized by artificial intelligence to the input provider. Technically, it does not seem possible to demonstrate that specific inputs entered by humans were crucial for the realization of the work.

In the absence of alternatives, it may be that the end-user—owner of the machine—acquires the rights to economically exploit the work. Although this person is the least involved in the creation process, they are the one who initiates the creative work of the machine.

The issue of moral rights—rights to protect the author's personality, including the power to decide whether and when to publish the work, claim authorship, and oppose any deformation, mutilation, or other modifications that could harm the work's honor or reputation—is more complex. Moral rights are indivisibly connected with authors, and thus, are categorized as personal rights.

It is evident that in our legal system, the author of an intellectual work must necessarily be a natural person.

One could lean towards not attributing such rights. However, this solution does not seem the most appropriate because those who created, programmed, or purchased the robot might want to protect the dignity of the work created by artificial intelligence to prevent modifications or other interventions from damaging its honor. Failure to recognize authorship could also result in missed economic returns by decreasing its visibility.

Given the various interests involved and the increasing use of these technologies, it is essential that legislation addresses all implications, potentially considering a fragmentation of moral rights among the various stakeholders.

Are robots responsible?

One of the main issues arising from the development of increasingly advanced artificial intelligence is the responsibility for any damage caused by AI systems, particularly when the work created by the algorithm is a copy of a pre-existing one.

As mentioned earlier, a robot cannot be considered liable for acts or omissions that cause harm to others. Currently, existing legal provisions on responsibility only cover cases where the cause of the machine's behavior resulting in damage can be traced back to a human agent (owner, programmer, user).

Pending specific regulation, and to consider the currently applicable regime, it is appropriate to start with the principles established by case law regarding liability in cases of copyright infringement by works found to be plagiarized. Specifically, the Courts (Cass. Civ., Sez. I, January 26, 2018, No. 2039) have noted that once objective plagiarism is established, due to the appropriation of essential elements of someone else's creative work, all parties who have contributed significantly to the infringement are jointly responsible, according to Art. 2055 c.c. This includes, besides the material author of the plagiarism, also the party who has

commercialized the works in their business activities. Indeed, it is part of the duty of qualified diligence, under Art. 1176 c.c., for market experts in the art world to ensure that the works for sale are not plagiarized.

In our case, it is reasonable to imagine that the preferred approach is to fragment responsibility between the party who commercially exploits the creation—who must ensure it does not infringe third-party rights—and the one who set up the machine's functions.

The considerations so far relate to cases where the final work generated by the robot-artist is identical or similar to an existing one. Another issue concerns the process of forming the work. We need to consider whether using images, texts, etc., used to train the algorithm, and essential for the creation of the final work, which are protected by copyright, constitutes a violation of the exclusive rights of the author if there is no authorization from the rights holder.

It is true that the literary, artistic, scientific, and computer-related knowledge provided to the machine is subjected to manipulative activity by the robot to create the work. However, since the information is inputted, we are dealing with a reproduction of digital copies of copyrighted works. Is consent therefore necessary?

I do not believe it is possible, in this case, to exclude or even limit the exclusive rights recognized by copyright law to the holder of the creative work by resorting to the doctrine of “fair use.” As is well known, this doctrine originated in England and developed predominantly in the jurisprudence of the United States until its codification within §107 of the Copyright Act in 1976. However, some similar forms also exist in European states. In Italy, for example, in Article 70 of the copyright law (L. April 22, 1941, No. 633 amended by Legislative Decree 68/2003), there is a provision that is somewhat analogous to that contained in the United States Code. In particular, this provision states that “The summary, quotation, or reproduction of fragments or parts of a work and their communication to the public are free if done for purposes of criticism or discussion, within the limits justified by these purposes and provided that they do not constitute competition to the economic use of the work; if done for educational or scientific research purposes, the use must also be for illustrative purposes and not for commercial ends.” The provision does not provide application criteria because, according to some authors, Congress only wanted to codify a principle applied by the Courts, leaving them with the task of interpreting it. A clear description of the requirements for the application of fair use can be found in the Supreme Court decision *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 577, 114 S. Ct. 1164, 127 L. Ed. 2d 500 (1994). Four factors are thus identified to be considered

together: transformative use, the nature of the protected work, the amount and importance of the copied part, and, finally, the effect on the market.

Part of the doctrine justifies the possibility of invoking fair use on the premise that the primary purpose of copyright protection is precisely to ensure and disseminate knowledge. After all, the 1787 American Constitution grants Congress the power to “promote the progress of science and useful arts, by securing for limited times to Authors and Inventors the exclusive right to their respective writings and discoveries” (Article 1, Section 8, Clause 8). The EU Directive 790/2019, in Recital 2, also states that copyright protection aims to “promote cultural diversity.” In Italy, the reference can be found in Article 9 of the Constitution, which states: “The Republic promotes the development of culture and scientific and technical research. It protects the landscape and the historical and artistic heritage of the Nation.” However, it should be noted that these sources likely draw inspiration from the Statute of Anne of 1710, the first copyright law in Anglo-Saxon law, which encouraged learned men to compose and write useful books. It reads, in fact, “An Act for the Encouragement of Learning, by Vesting the Copies of Printed Books in the Authors.”

To avoid hindering progress, the rights of authors cannot translate into an absolute monopoly. For this reason, while ensuring exclusive control over the reproduction of works and recognizing a financial incentive through the granting of exclusive rights, such rights are subject to limitations. This way, third parties have been allowed to reproduce copyrighted works without the author's consent to create different and original works that contribute to expanding and promoting public knowledge.

The statements made in some rulings by American and English judges are interesting in this regard. In particular, the Supreme Court in the case of *Goldwin-Mayer Studios, Inc. v. Grokster, Ltd.* in 2005 stated that the broader the copyright protection, the more scientific research can be discouraged, so a compromise is necessary. Similarly, English courts have expressed themselves. In the case *Cary v. Kearsley*, 170 Eng. Rep. 679, 681, 4 Esp. 168, 170 (1802), Lord Ellenborough argued: “While I shall think myself bound to secure every man in the enjoyment of his copyright, one must not put manacles upon science.”

My opinion, however, is to exclude in cases such as the one we are dealing with the possibility of invoking fair use because exceptions or limitations to copyright, according to this system, would be admissible only where the use of works of ingenuity is not for commercial purposes, but for educational and non-profit purposes [4].

This remains an open issue.

It is evident then the risk that may arise for the author of the works, especially considering the now established ability of robots to emulate the personality of the authors themselves, replicating their creativity and artistic expression.

In this context, it might be appropriate to consider a model of Open-Source licenses that allows the use and copying of data protected by copyright to generate new creative works.

Conclusions and final considerations

Robots have become part of our daily lives, but for now, the law has limited itself to regulating them as “objects”. Therefore, as we have seen, the manufacturer, owner, possessor, or user is responsible for the choices and behaviors made by the machines. The question many people ask is whether this system can still hold up.

As we have seen, technological progress has led to the creation of robots capable not only of performing activities typically carried out by humans but also endowed them with advanced and autonomous cognitive abilities that allow them to adapt and interact with their operating environment. Modern machines, in fact, are acquiring (and will acquire) an increasing degree of autonomy in performing the functions for which they were created, with an inevitable component of unpredictability and the consequent responsibilities connected to such behaviors.

The autonomy of robots, then, opens the debate on their nature, forcing legal systems to evaluate whether they should be considered in light of existing legal categories—thus, as natural persons or legal entities—or whether it would be preferable to create a new category with specific characteristics and implications for the attribution of rights and duties.

This latter possibility should not surprise us. After all, it has happened in some legal systems where rights have been granted to entities that are not human beings. For example, New Zealand has recognized legal personality to a river. Similarly, an important ruling in 2017 by Indian judges recognized the importance of trees for sustaining life on Earth and the connection of forests with the religion and culture of the Indian people.

In this new scenario, it seems difficult to think that robots could have fewer rights than trees, rivers, or mountains. Moreover, we cannot forget that Saudi Arabia has already granted—probably as a provocation—citizenship to the robot Sophia.

In this direction, the European Parliament seemed to be heading with its 2017 Resolution, highlighting how it was increasingly difficult to link the activities of machines to a single human responsibility. Therefore, it hoped for autonomous accountability through “the creation of a specific legal status for robots in the long term, so that at least the most sophisticated autonomous robots could be considered as electronic persons responsible for making good any damage they may cause, and possibly the recognition of the electronic personality of robots that make autonomous decisions or otherwise interact with third parties independently” (point 59, letter h).

Thus, the attribution of a third type of legal subjectivity was proposed, transforming robots into legal entities with full capacity, rights, and responsibilities, and therefore also capable of compensating the party to whom the damage was caused. However, the question arises whether making intelligent systems liable is really necessary. I wonder if this choice is a way to absolve those who developed or use them, as is the case in commercial law for limited liability companies.

Moreover, if the purpose of granting legal personality to a robot is to better regulate the consequences of its actions, are we sure that effective protection can only be achieved by recognizing a level of subjectivity to the machine? Compensation for damages caused by it could, in fact, continue to be guaranteed by the insurance policy subscribed by the owner, without this limiting or even reducing the effectiveness of risk coverage.

I believe, then, that it is not necessary to grant machines a personality (so-called electronic) modeled on that of a human person. Instead, we could limit ourselves to recognizing the creation of mandatory insurance coverage, as already happens with cars. It could also be imagined to create a fund to ensure the possibility of compensating for damages in case of absence of insurance coverage.

Indeed, the European Parliament now seems to be moving in this direction. In its new 2020 Resolution, it abandons the option of granting legal personality to AI systems, stating that “it would not be appropriate to endow AI technologies with legal personality and recalls the negative implications of such a possibility on incentives for human creators.”

These considerations inevitably lead us to reflect on the fact that artificial intelligence is not a transient phenomenon, a phase of technological evolution, but a revolution of our entire existence. Therefore, the need to interact more consciously and carefully with artificial intelligence is unquestionable, and it is also indispensable to come to a legal regulation. It is also

difficult, at least in my view, to imagine excluding in the future (not too distant) the possibility of granting copyright to works generated by AI, which could be adapted to the characteristics of artificial intelligence (perhaps with a shorter duration, and rights relating to the work, such as exclusive use, market introduction, and economic exploitation licenses).

A challenging task, also because, as we have said, there are various sectors involved in the use of these technologies, and numerous issues can be connected to the implementation of AI systems. However, the main problem may be that technology advances faster than regulations.

References

- [1] For an in-depth look at the different definitions of artificial intelligence R. J. STERNBERG (a cura di). *Handbook of Intelligence*, Cambridge University Press, 5 June 2012, p. 3-15, <https://doi.org/10.1017/CBO9780511807947>; S. LEGG, M. HUTTER, *A Collection of definitions of intelligence. Advances in Artificial General Intelligence: Concepts, Architectures and Algorithms*, in *Frontiers in Artificial Intelligence and Applications*, vol. 157, 2007, p. 17-24, <https://doi.org/10.48550/arXiv.0706.3639>.
- [2] A. PAPADOPULOU, *Creativity in crisis: are the creations of artificial intelligence worth protecting?*, in *Journal of intellectual property, information technology and electronic commerce law*, vol. 12 no. 3, 2021, p. 947-996, https://www.jipitec.eu/archive/issues/jipitec-12-3-2021/5352/papadopoulou_pdf.pdf.
- [3] A. LAUBER-RONSBURG – S. HETMANK, *The concept of authorship and inventorship under pressure: Does artificial intelligence shift paradigms?*, in *Journal of Intellectual Property Law & Practice*, vol. 14 no. 7, 2019, p. 570-579, <https://doi.org/10.1093/jiplp/jpz061>.
- [4] L. TURINI, *A chi spettano i diritti sulle opere dell'intelligenza artificiale. Dall'addestramento su opere altrui alla generazione di opere nuove*, Maggioli Editore, 2024, p. 39-42, <https://www.maggiolieditore.it/a-chi-spettano-i-diritti-sulle-opere-dell-intelligenza-artificiale.html>.

L'ARTE ALGORITMICA DEI ROBOT

Paolo Gallina – University of Trieste

Introduction

Why are some human beings fascinated by the possibility of "generating" an artificial entity capable of "producing art"? The reasons are many and varied. There are those who have seen in automatism outside the biological sphere an anti-anthropocentric action; those who are fascinated by the virtuosity of technique; those who want to emulate the divine act of god's creation; those who are impressed by the aesthetic effects that "machines" can produce; those who delegate part of their work of artistic production to machines; and those who are bored with traditional tools and seek original ways of expression.

Whatever the reason, or reasons for this creative tension, what is certain is that the dream of shaping life from nonliving matter is not the result of a modern attitude; it is an innate prerogative of human beings, a challenge that has always been latent, from the beginning of human history. Testifying to this are the myths of the past. Hephaestus made Talus, a living bronze statue; a kind of automaton having the purpose of defending the island of Crete. According to Homer, Hephaestus again built three automated helpers, conceptualizations of modern industrial robots: these were three golden workers capable of operating the forge. Within Hebrew mythology, the Golem is an anthropomorphic clay entity, lacking intellectual capacity but endowed with superhuman strength.

The pioneers of robot printer

The idea of employing a machine to make artifacts traceable to figurative art has ancient roots.

Pierre Jaquet-Droz [1721-1790] was a visionary watchmaker. In his workshop, aided by his sons, he made a variety of dolls capable of performing heterogeneous tasks, such as drawing, writing, and playing music. Watchmakers were the engineers of the past. Not surprisingly, in those very years, another watchmaker, Friedrich von Knauss [1724 -1789] made a figurine capable of tracing short sentences on a sheet of paper. And a few years later, Henri Maillardet [1745-1830] made an automaton masterpiece: the Draughtsman-Writer (Fig. 1).



Fig. 1 – Automa made by Henri Maillardet, 19th century

If in the 1700s these jewels of technology made the public gasp in amazement at the fairs at which they were exhibited, by the end of the nineteenth century refined toys capable of operating on similar principles could already be purchased. Some of these are still functioning, such as the Clown Artist (Fig. 2).



Fig. 2 – Toy specimen that can draw on a sheet of paper thanks to a mechanism consisting of two cams that provide instructions to move the nib on the plane along the horizontal and vertical directions.

A description can be found in the text "*Magic: Stage Illusions, Special Effects and Trick Photography*" (1897), by Albert A. Hopkins. Significantly, despite the fact that they were playful, they were explained in a magic book. In those days, seeing a self-contained object suddenly come to life was a sight that had the unbelievable.

Technologically speaking, the real technological breakthrough came in 1928, with the Gakutensoku robot, created by Makoto Nishimura [1]. When Karel Čapek's famous play, "*Rossum's Universal Robots (R.U.R.)*," was staged in Tokyo in 1923, Nishimura was among those in attendance. The play was so convincing that Nishimura felt compelled to condemn the

inauspicious prediction. R.U.R. in fact painted a dystopian future according to which bioartificial intelligences would dominate and enslave humans; in Nishimura's words, "the emergence of a perverse world in which humans become subordinate to artificial humans." According to the scientist, an artificial being modeled after human stylistic features and employed as a slave represented a paradox. In 1926 Nishimura left the United States, moved to Osaka and began building his ideal artificial human. It was a three-meter-tall robot; an innovative machine in many ways. First, the scientist wanted to endow it with expressiveness so that it could induce positive feelings.



Fig. 3 - Photo of the robot Gakutensoku.

Beyond the robot's supposed graphic capabilities, it is interesting to note the different design approach depending on the culture of origin. While in the Western world, a robot is merely a utilitarian tool, in the Japanese world, a robot can be regarded as an "object with a soul." This sensibility has religious roots. In fact, Shintoism predicts that even inanimate objects can be endowed with inherent spirituality.

The 1950s

The artist who more than any other has explored machine-mediated pictorial action is Swiss sculptor Jean Tinguely. For Tinguely, the kinematics and dynamics of the machine, enriched with jarring sounds and noises, became an integral part of the work. For the first time, the concept of the "autonomous painting machine" stepped out of the exhibition dimension in order to amaze the public, and embraced a powerful expressiveness with artistic valence. The object of the art project was threefold: it was the machine, curated and assembled for a bodily presence that "struts" and arouses attractiveness; it was the work produced by the machine, the drawing itself; it was the visual performance in the making that, like a theatrical performance, changed and evolved. Despite the fact that these were non-anthropomorphic machines, their appearance

hinted at the chaotic cohesion of biological life, as if they were pulsating living agglomerates possessing individual and collective autonomy. In making these works, the artist drew on his own life experiences, including those of childhood. With Tinguely, randomness became a dynamic element. The representation of chance was entrusted not to the canvas but to the temporal evolution of his moving sculptures, the forerunners of robots. From a technological point of view, his works were likened to huge clocks equipped with wheels, gears and elastic elements (Fig. 3).

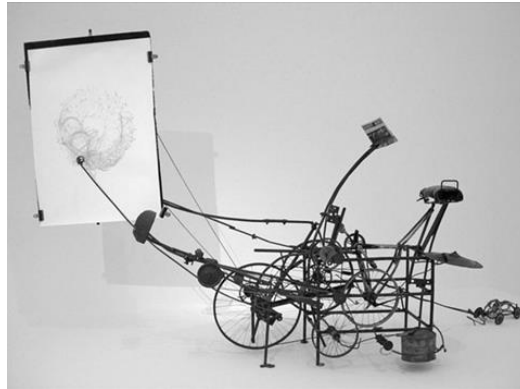


Fig. 3 - Cyclograeur, Jean Tinguely, 1959.

From a strictly robotic point of view, on the other hand, the first to pose the problem of automating painting was perhaps Raymond Auger. Auger made a robot capable of picking up color with a brush and tracing marks on a canvas. But his pioneering spirit has almost been forgotten. The reason lies in the fact that his action had no artistic significance. Painting, for Auger, was only one of the many activities his robotic arm could accomplish. Instead of promoting the works in exhibitions that could amplify the meaning of his work in a communicative context, he employed the robot as a "canvas dispenser"; for Auger, the machine was just an object that, among other things, could trace marks on the canvas. Significant was the advertisement sign written in his own hand with which Auger invited visitors to purchase the cheap paintings made by the robot (Fig. 4).



Fig. 4 - Photo of the sign with which Raymond Auger invited potential buyers to purchase the works made by his robot.

The fact is that Raymond Auger was not an artist, but a talented and visionary technologist, moved by a yearning desire to imitate an artist.

On the other hand, the one who had an opposite educational path, that is, starting from an artistic background he delved into the technical aspects offered by automation was Akira Kanayama, a member of the Japanese *Gutai art group*. Beginning in 1957, Kanayama remotely drove a robot for the purpose of painting. The robot consisted of a rudimentary trolley equipped with four wheels, which was wire-driven and moved on the floor (where the canvas was placed), similar to what a remote-controlled toy car would do (Fig. 5). The robot was equipped with interchangeable tools, i.e., jars from which it tapped color according to a process that, by Kanayama's aspiration, mimicked *action painting* (Fig. 6).

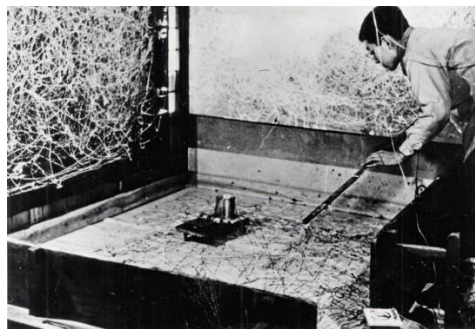


Fig. 5 - Remotely controlled painting machine made by Akira Kanayama, 1957.

With an ironic twist, Kanayama also launched a conceptual attack on the abstract expressionist idea that art could or should be an expression of the soul poured onto a canvas. For Kanayama, the centrality of the artist was being challenged.

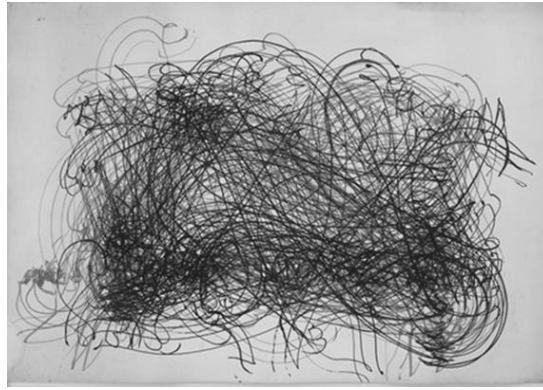


Fig. 6 - Painting made by Kanayama employing the wire robot, 1957.

Georg Nees, a pioneer of algorithmic graphics

The link between a painting robot and mathematical algorithms is inseparable.

Inseparability is inherent in the very activity of painting: the set of strokes on the canvas is decomposable into curves and segments described by means of "mathematical languages."

Georg Nees [1926 - 2016], a mathematician lent to computer science, was one of the first to employ such generative algorithms. On February 5, 1965, he and other artists exhibited their first works in Stuttgart. The title of the group exhibition was a manifesto of the artistic union: *Computerkunst*, or *Computer Art*. I reproduce in Figure 14 a graphic schematization of one of the works I most appreciate (Fig. 7).

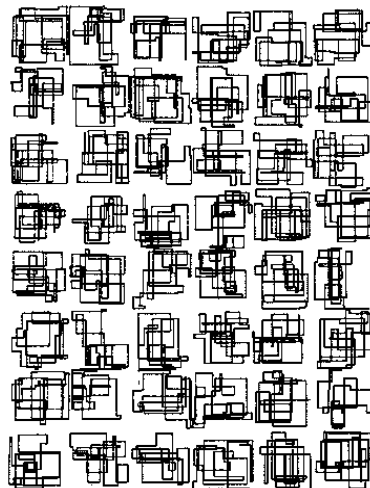


Fig. 7 - *Computerkunst*, graphic example of how line becomes an expressive element Georg Nees, 1965.

Despite the simplicity of implementation, the resulting "graphic thumbnails" seem to come alive. The viewer's gaze is led by the hand and jumps from one image to another and then unconsciously chases some local features.

These sensations can be called "kinetic sensations"; the observer's attention, that is, where he or she directs his or her gaze (gaze in English), is captured and led along obligatory paths. In all this, the algorithm has a perceptual *raison d'être*. It introduces forces that harness the visual path within a trajectory.

Nees' work was itself algorithmically innovative. He is the progenitor of what would later be called *generative art*; his works, though represented in digital form, possess the romantic charm of the firstfruits. As ethereal and associated with algorithms, they suffer from the "reproducibility phenomenon," but this does not diminish the artistic value of the innovation.

The Stuttgart Digital School, headed by Frieder Nake, Georg Nees and Manfred Mohr, is not only important for its artistic achievements, but also for the cultural influences it produced and the reflections it nurtured. For the first time, members of the art system, thinkers, and intellectuals discussed the suggestions triggered by the relationship between philosophy, art, and computer graphics.

Frieder Nake

In 1963 the Department of Mathematics at the University of Stuttgart purchased a Zuse Graphomat Z64, a plotter capable of "pulling lines." Unfortunately, the new machine had no dedicated software. It was just a neat set of gears and motors without intelligence that could move a stylus through low-level commands. The task of "instructing" it was given to young doctoral student Frieder Nake, who produced machine language software capable of moving the nib. Instead of producing graphical diagrams of scientific value, Nake began "drawing" with pixels. When he heard about Nees' experiences, he was seized with compositional enthusiasm and asked a book dealer acquaintance if he could arrange for his first exhibition in the library.

Nake's artistic pursuit grew in tandem with his public recognition. In August 1966, Nake won the annual *Computers and Automation* competition. Also in the same year, his works were compared with those of the modernist painter Hans Hartung and even with some works by Paul Klee. The stylistic juxtapositions derived mainly from the common use of strokes. The theorized connection with Hartung's works shows how the conservative establishment of art criticism considered computer graphics part of the contemporary avant-garde present in Germany. And, conversely, it did not recognize its full autonomy of purpose. The algorithmic dimension of creative effort had not been accepted.

Like Nees, Nake, at an early stage of his investigation, focused on lines, almost an obligatory step (Fig. 8). But later he also experimented with more complex colors and patterns.

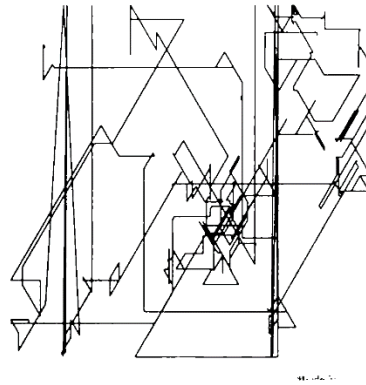


Fig. 8 - Zufälliger Polygonzug, Frieder Nake, 1965.

AARON

In 1968 Harold Cohen wrote his first computer program. He did so because he accepted the invitation-almost a challenge-issued by one of his students (Cohen was then an art teacher). The student had stated that he could teach him something new, something fun and "refreshing to the mind." And so, under the guidance of a nerdy teenager, Cohen learned the basics of computer science. From then on, he never stopped programming; he learned the Fortran language while attending the Computer Center at the University of California, where he was a *visiting* professor.

From early on it was clear to him that programming was a form of representation: a set of rules, instructions and procedures that led to action. The point of contact with art came from the fact that representation possessed a pulsating expressive dimension. In Cohen's mind, within a few months, there was the idea that computer codes represented knowledge leading to the act of "making art."

Later, Cohen developed his first robotic system dedicated to painting. The electronic brain consisted of a common computer, the Micro Vax II, connected to a machine that operated a pen. The system was capable of drawing in ink outlines of stylized human figures surrounded by omnipresent floral patterns (Fig. 9). The syntax of the system consisted of planned lines from before implementation, but in an undefined and general way. That primordial robot had some freedom in deciding how many and which subjects to draw.

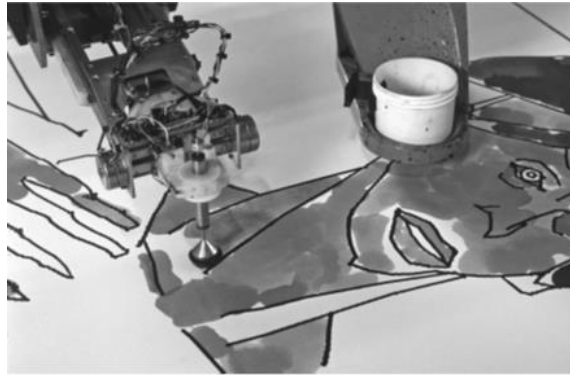


Fig. 9 - Robot implemented by Harold Cohen.

Cohen developed algorithms that could produce the primitive, essential, embryonic effect he had seen in rock carvings. Regarding this he stated, "My intuition, which I believe later proved to be correct, was that whatever action is required, the machine and the program must have something in common with the observer of the work." And this something in common, since it could not be cultural, had to be cognitive. The robotic system, therefore, for Cohen had to take into account the perceptual activity of the observer, had to possess algorithms that tickled the observers' expectation, generating an order in the signs, perceived as organic, even if primitive. He called these signs that emerged like ancestral creatures on paper "cognitive primitives." In this first phase he made open and closed geometries, agglomerated into graphic cluster forms (Fig. 10). Here is how he described the activity, "The closed form came first. I have always been convinced that the closed shape is an elementary cognitive (perceptual) mode, simply because our ability to recognize an object in the world is based on very elaborate perceptual structures, in which the eye operates as a detector of contours."



Fig. 10 - Drawing for Machine and Four Hands, Harold Cohen, 1972.

Significant experiments in modern art and robotics

In the following paragraphs I provide an overview of the state of the art, showing some significant examples of artists making use of robots.

1) Patrick Tresset and the robot that transcends errors

Patrick Tresset is an artist who has made structural *yielding* a strong point in his works. His work is an example of how the artist is sometimes able to transcend the technological limitation of the media he uses. Tresset works with a series of robots equipped with only four degrees of freedom. These are rather basic robots; one gets the impression that they could break at any moment so fragile do they appear. The action of the individual robot is reminiscent of a wheat stalk lashed by the wind. Seeing it trace is like observing a Parkinson's patient engaged in drawing on a sheet of paper. Indeed, the tool Tresset uses is more like a toy than a professional robot. However, the artist has managed to skillfully tame the machine's "degree of error."

Having taken a picture, the robot, equipped with an ordinary ballpoint pen, makes irregular strokes; from time to time, it detaches the tip from the paper. The algorithm underlying the process is quite simple: the marks are made favoring the darker areas as density. For a statistical effect, after a large number of strokes, the image emerges and takes shape from the tangle of marks (Fig. 11).

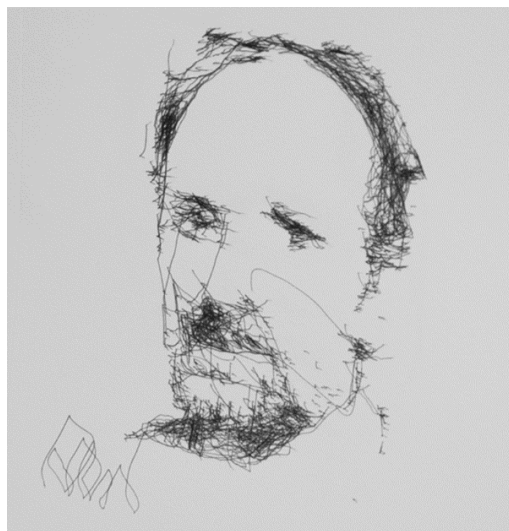


Fig. 11 - Self-portrait of Patrick Tresset.

2) e-David

eDavid is the name by which Prof. Oliver Deussen and his research team have named a robot dedicated to painting. In fact, the robot itself is borrowed from the industrial world; in their traditional context these types of robots weld, pack and place objects with zombie-like repetitiveness. Never would one imagine that they could make work like the "Wind-bent Tree" (Fig. 12).



Fig. 12 - Wind-bent tree, eDavid, 2011.

What makes the difference is the intelligence with which the robot is equipped. An intelligence that, in the case of eDavid does not mimic human intelligence - the system is not based on neural networks - but is algorithmic-heuristic, that is, it relies on mathematical rules dictated by the programmer's experience. The above design, when observed in detail, is composed of a dense series of straight segments, which the author calls *strokes*.

The algorithm acquires a photographic subject: an input digital image. It then chooses a random point in the image plane and evaluates whether or not a stroke should be drawn at that point. It then virtually generates a large number of possible "stroke candidates," each with a different angular orientation. The next step is to select the one that most "conforms" to the picture (Fig. 13).



Fig. 13 - Hands, eDavid, 2011.

3) Pindar Van Arman

Pindar Van Arman is an engineer. Twenty years ago he began producing his first works employing a simplified robot and has not stopped since. It is not for me to judge the artistic value of his works, but one thing is certain: Van Arman has always continued to experiment.

In an early phase of his study he focused on figurativism, on producing uniform brushstrokes that respected contours and tone. Later he abandoned precision in favor of introducing distortions that would highlight certain imperfections (Fig. 14).



Fig. 14 - Portrait made by Pindar Van Arman, 2016.

4) *The body of the robot*

In the book "The Code of Creativity," mathematician Marcus du Sautoy questions the problem of authorship of a work of art made by a robot. In an attempt to provide a judgment, he cites the well-known case of the *monkey and the photographer*.

During a trip to Indonesia, photographer David Slater left his camera unattended for a few minutes; during this short time, a macaque monkey named Naruto took several photos. Most of them were blurry and random. However, some of them brought out a distinct awareness in the animal's use of the medium. Famous is the smiling *selfie* of the animal. Slater was so impressed by those photos that he spread them on social media telling the story. But when, having gained the expected fame, he advanced rights to the financial exploitation of the photo, he did not get satisfaction. The judges who analyzed the dispute from a legal point of view held that the author of the photos was in all respects the monkey and not the owner of the camera, that is, the medium of the action.

Drawing a parallel, the same question arises regarding robot-generated works. du Sautoy's conclusion, agreeable as far as I am concerned, is that the human-generated algorithm to make the robot paint underlies the robot's action. Therefore, unlike the monkey, the softwarist/artist is to be considered the author and owner of the work.

It follows from this statement that robots do not possess creative character. The robot, or artificial intelligence, is not endowed with free will.

But why then do we ask the question? Why is it that in the presence of a humanoid robot the question arises as to who is the author of the final work, while observing an artist using an intaglio press there is not the slightest doubt? Never, ever is the press assigned an authorial role. Yet the robot machine and the press machine are both tools devoid of their own will.

Bizarre as it may seem, the dilemma depends on the robot's anthropomorphic appearance. It should be said: the clothes make the man. The robot's human features and the movements it enacts convey the sense that it may be endowed with a life of its own; they induce the observer to associate a kind of "little soul" with the set of gears and transistors.

Therefore, starting from this point of view, even the anthropomorphized appearance of the robot in a performance about painting can make a difference.

This is exactly the case with the famous Ai-Da robot. It is a humanoid robot equipped with two arms and a lifelike face. The robot possesses a classical actuation system. The structural chain and mechanical transmissions exhibit considerable compliance, which is evident from the oscillations observable with the naked eye during movements. Implicitly and cunningly, the designers attribute to the robot the canvas placed in front of it, a canvas of considerable visual impact (Fig. 15). I recognize that the artistic/robotic marketing is effective, but it is still a communicative deception.



Fig. 15 - Robot Ai-Da: photo of the robot in front of a canvas (painted by others).

Actually the designs produced by Ai-Da are quite simple. In a drawing made by Ai-Da with a marker pen, the marks are discontinuous and jumpy, each segment ending in a swell caused by the inability to dose force at the end point of the path, which is held for too long (see Fig. 16). Moreover, the algorithms employed are primitive, to say the least.

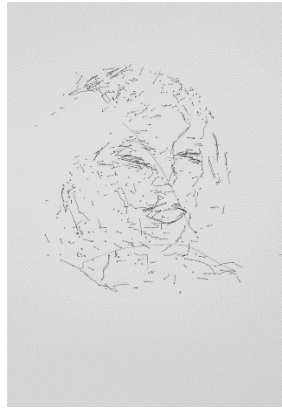


Fig. 16 - Drawing made by Ai-Da.

Yet the robot's anthropomorphism has allowed Ai-Da to gain international attention; it is easier for journalists and media content generators to pitch a news story using a picture of a beautiful plastic girl than to focus on complex algorithm descriptions. The perceptions one feels in the presence of a humanoid robot have been extensively studied from a psychological perspective.

5) Watercolors

"R32 Bruno, the common man" is a work created by Paolo Gallina together with Lorenzo Scalerà. It employs a technique that combines the use of watercolor and tempera (Fig. 17).



Fig. 17 - R32 Common man, watercolor and tempera, 2017.

In the painting, watercolor was used for the intermediate tints, making sure to avoid blooms. Note that the continuity of gray along the road in the foreground is broken up in isolated brush strokes that provide the sense of the presence of a pavement. This effect depends only minimally on the algorithms that define the sequence of strokes to be made; in fact, a good result is achieved by properly dosing the amount of water drawn from the bristles; if the tip of the brush penetrates too far below the free coat of liquid color contained in the pan, the amount of water deposited on the sheet is excessive, causing uncontrolled dribbling; on the contrary, if the bristles

of the brush pick up little water, the brush strokes turn out "dry" and poorly blended with each other; proving that painting by employing robots is a more artisanal activity than people think.

A variety of algorithms were required to make the painting. For example, the strokes made with larger brushes had to be spaced further apart. Here, therefore, the algorithm had to take this into account. Broadly speaking, the algorithms can be grouped into two macrocategories: algorithms that deal with background colors and those that deal with tracing outlines.

6) Engravings

Engraving is an ancient printing technique dating back several centuries before Christ. In the cyclical rhythms of art history it has been revived and abandoned. It is subdivided into several subtechniques that have in common a *matrix* employed to put the design on a sheet of paper. It is therefore an *indirect* technique, in the sense that the artist does not draw directly on the sheet, but on a support, producing works in a mirror form (the drawing on the matrix is mirrored with respect to the drawing that is printed on the paper).

Today, the most widely used techniques are *etching* and *drypoint*. In the former, etching on the plate is achieved by the effect of acid that "eats" the metal along the marks where the plate is unprotected (*biting*).

In the second, the plate is engraved directly with a carbide tip called a *burin*. Through the robot I make use of both techniques. The minimum unit of the engraved work is the line. Therefore, since robots understand the "language of lines," the engraving technique is congenial to the operation of the machine. In classical engraving, in fact, there are no colors; light and dark tones are achieved by overlapping hatches (Fig. 19).

"R13 Do the Right Thing" is a work combining storytelling and drypoint engraving. It is the world's first engraving made on zinc plate by a robot (Fig. 18).

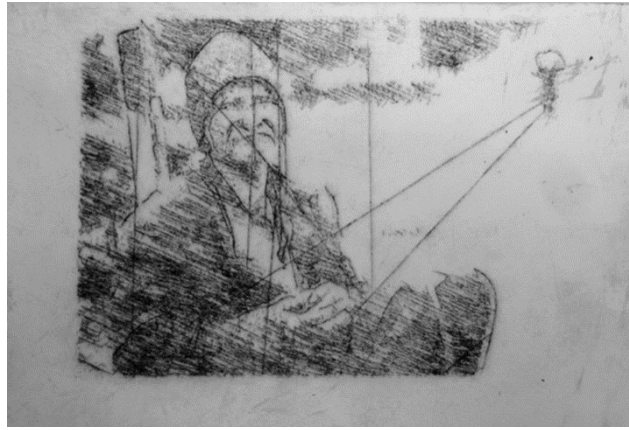


Fig. 18 - R13 Do the right thing, robotic drypoint etching, Paolo Gallina, 2016.

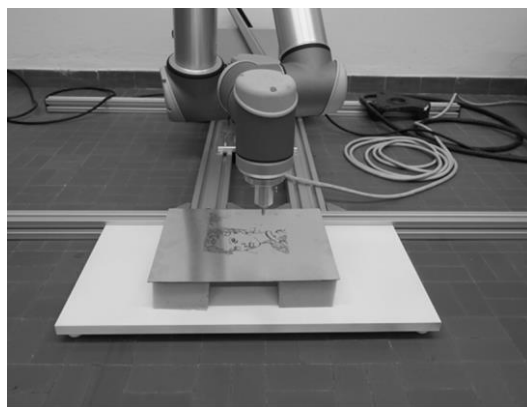


Fig. 19 - Engraved plate resting on top of the four elastic blocks.

Conclusions

The advent of robots and artificial intelligences in the domain of art should not be perceived as a threat to the role of the human artist. Rather, it should be regarded as an opportunity to extend the frontiers of creativity. The convergence of human ingenuity and machine capabilities has the potential to give rise to novel forms of expression, wherein both dimensions coexist and mutually enhance one another. Nevertheless, the incorporation of robots and AI into artistic endeavors inevitably gives rise to ethical and philosophical inquiries. To what extent can an artificial creation be considered to possess aesthetic and cultural value? This question prompts reflection on the very meaning of creativity and humanity, stimulating debate on the nature of these concepts.

The field of art has always been in a state of constant evolution, reflecting the technological, social, and cultural changes that occur in each era. As illustrated in the article through various historical examples, the advent of robots as creators of art represents merely the latest stage in a

journey that has witnessed humanity's exploration of novel forms of expression through increasingly sophisticated tools. Although robot-produced art may appear to be a radical departure from traditional notions of human creativity, it is crucial to recognize that it is the human being who oversees the programming, guidance, and interpretation of these creations. Acceptance of robotic art does not imply a disregard for the value of human art, but rather an acknowledgment that creativity can manifest in unconventional forms.

References

Paolo Gallina, *La protoarte dei robot*, 2023, Mimesis edizioni, ISBN: 9788857591667

[1] Yulia Frumer, *The Short, Strange Life of the First Friendly Robot*, in IEEE Spectrum, 21 May 2020, <https://spectrum.ieee.org/the-short-strange-life-of-the-first-friendly-robot>

THE ENCHANTMENT OF TRUTH: A CALL FOR ARTISTS TO TACKLE CONSPIRACY WITH WONDER

Sofia Marasca, Federico Bomba – Sineglossa

The European Digital Deal project

[European Digital Deal](#) is an initiative co-funded by the Creative Europe programme and led by [Ars Electronica](#), of which [Sineglossa](#) is the Italian partner. 14 European cultural organizations together to investigate how the rapid and sometimes naïve adoption of new artificial intelligence technologies can contribute to altering or undermining democratic processes.

Through an international call for artists, European Digital Deal has selected 12 project ideas that will turn into interactive installations on the risks and opportunities related to the use of artificial intelligence in the daily lives of European citizens by 2025.

As the first step of the European Digital Deal project, each partner brought together a [group of experts](#) with a double role: support in defining the challenge to be issued to artists through the open call and support for the artists in residence in their research and project development. Sineglossa involved voices from the fields of artificial intelligence to contemporary art, from debunking to research on technology and society: Ilaria Bonacossa, Wu Ming 2, Luca Baraldi, Sara Tonelli, Barbara Busi.

The Process

The European Digital Deal project is developed around 5 main objectives, through a series of actions ranging from the organisation of public events to the production of dissemination content, from the production of artistic works to capacity building:

- *Build and validate*: each partner identifies a number of technology, art and digital experts who form “Local Expert Groups (LEG)”, i.e. they define the challenge to be set in the call for artists, make resources, tools and knowledge available for the implementation of the artistic residency programme to which the artist selected from the call for artists will

have access, and finally support the artist through online and offline consultancy and mentoring to accompany him/her in the prototyping of the pilot project.

- *Innovate*: through an Open Call launched in each of the member states (12 in total for the network of partners), the artists will be identified to work on 12 different challenges related to the topics of technology, justice and digital, to individually develop 12 different artistic projects, with the support of some of the partners involved for research and development activities, and of the planned meetings and exchanges between the artists, for knowledge sharing.
- *Learn*: the project implements a series of non-formal educational activities for schools and students, through artistic and creative methodologies, and capacity-building opportunities for artists, on the research topics of the member organisations (AI transparency, art and democracy, art and debunking, plant/animal/human AI interaction).
- *Showcase*: the works and research and study content that emerged during the project will be the subject of events, festivals and exhibitions in each of the member states.
- *Guide*: the final step of the project will be to present the results to stakeholders outside the cultural sector, raising awareness among a non-expert audience about the capacity of the arts to attract and develop innovation. Dissemination of the results will take place through the *Digital Future Action Plan through the arts*.

The challenge for artists

We are living in an era called the “age of fake news”. Although fake news is nothing new in our times, today, compared to the past, its spread is increasing, because there is a lack of valid intermediation by platforms to act as a verification filter and it is much easier to put news and information online without having to prove its veracity.

Given the possibility for anyone to access the means of content creation and dissemination, fake news is also produced and disseminated not only by those who produce content for the official media but by anyone who has access to the Internet. Moreover, a greater complexity has recently emerged in the fight against fake news, posed by multichannel and multimedia: the case of a true news story accompanied by a fake image, or a true video but with a fake caption, can occur.

To counter widespread disinformation, it is necessary to ‘debunk’, i.e. to distinguish true news and content from false ones. The term debunking derives from a practice of ancient origin and the debunker is, literally, the ‘debunker’ (from de-bunk, a term of English origin composed of the prefix *de-*, meaning ‘to remove’, and *-bunk*, meaning “meaningless speech”). A debunker today is a person who debunks and disproves false, dubious or anti-scientific news: dismantling conspiracies, identifying hoaxes.

Sineglossa chose to launch ‘The enchantment of truth’ challenge to the artists, to challenge fake news with wonder. The choice to investigate this issue as one of the main challenges posed to democracy by AI comes from reading the novel *La Q di Qomplotto. QAnon e dintorni* by Wu Ming 1, published in March 2021 by Edizioni Alegre [1]. In the volume, an “unidentified narrative object” – as Enrico Manera’s review for Doppiozero states [2], given the different types of writing it is composed of – tells the difficult, disturbing and twisted story of QAnon, conspiracy movement that originated in the United States and spread throughout the world. In the book, the author proposes a semantic and linguistic redefinition of the terms used to describe the phenomenon of what he calls *conspiracy fantasies*, highlighting which characteristics debunking practices should have in order to be effective: the *appeal of stories* and the *kernel of truth*.

If, as Wu Ming 1 states, these two elements are necessary for an effective form of debunking, what would happen if an artist were to take on this challenge?

Wu Ming and the act of showing the sutures

In the 1990s, the [Luther Blisset collective](#) launched a novel practice to counter disinformation, which over the years has been renamed as “horrorism”: an artistic terrorism that was based on the inoculation of false news. The first real debunking exercise the group conducted with this strategy was during the Beasts of Satan case, with the aim of demonstrating the ease with which news was republished in the press without proper verification. Strictly rationalist debunking, which confronts the person fascinated by conspiracy fantasies with the inconsistencies of what they believe, works ineffectively. In order to understand what might work more, the collective investigated what is commonly called conspiracy, with the aim of deconstructing its characteristic elements. Thus in *La Q di qomplotto*, the difference between “conspiracy theory” and “conspiracy fantasy” emerges: the former have a specific, social and political purpose, are real and imperfect, have an end and those responsible are sooner or later, by justice or collective knowledge, identified. The latter, on the other hand, concerns an unlimited number of people, have no

definite purpose, are perfect and internally consistent, and have no end, so they transcend history.

The person who believes in conspiracy fantasies is simply a person who *cannot resist* the fascination of stories, on which the “narrative” explanation of the fact, rather than the scientific one, has a hold. The rational debunker in this scenario appears as a *party pooper*, the one who spoils that feeling of “being the repository of a particular truth” that those fascinated by conspiracy fantasies feel.

On the contrary, the debunker should rather satisfy that desire for wonder that has a hold on the audience of conspiracy fantasies. Between debunker and audience, there should then be a practice that Wu Ming calls ‘showing the sutures’, i.e. showing the trick without losing the effect of surprise. This is, for example, what the magic duo Penn & Teller propose in their shows, which question the power hierarchy between those who do the trick and those who watch, the basic rule of traditional illusionism professes: the importance of not revealing the trick. By maintaining the ability to perform and arouse wonder, the two performers manage to activate a process of empowerment in the audience, giving them the tools they need to understand what is happening on stage [3; 4].

Currently, AI has a similar capacity for wonder to the magic trick, because it often behaves in a similar way to humans; following the suggestion of the writers' collective, we then asked the artists how it is possible to show the suture of the “magic tricks” that AI produces, in order to reveal them through a reliable narrative.

dmstfctn and the Waluigi's Mask artwork

Among the more than 100 applications received for the call for artists ‘The enchantment of truth’, the project selected by the jury was *The Waluigi's Mask*, by dmstfctn (pronounced ‘demystification’), an Italian-London-based duo working with audiovisual performances, videogames and installations (Fig. 1).

Since 2018, dmstfctn work investigates complex systems through the study of communication networks, capital, information flows, and, more recently, AI. Among their latest productions are *Waluigi's Purgatory* (Fig. 2) and *GOD MODE (ep.1)*, interactive audiovisual performances exploring AI folklore and the use of simulation in AI training. dmstfctn have performed and exhibited at Berghain, Serpentine, Design Museum, Onassis, Corsica Studios, HKW, Fotomuseum

Winterthur, LUMA Arles, ARTER Istanbul and Aksioma among others, and at festivals such as Unsound, CTM, transmediale, Semibreve, Borderline and Impakt. In 2021, Krisis Publishing published *ECHO FX*, the artists' performance about Brexit market manipulation [4], later also published in *Ø: The Book* by Flatlines/Hyperdub [5]. In 2019, Mille Plateaux released *Flash Demons*, a collection of the artists' performances focusing on financial market crashes. dmstfctn were recipient of the 2017 transmediale Flusser artistic residency.

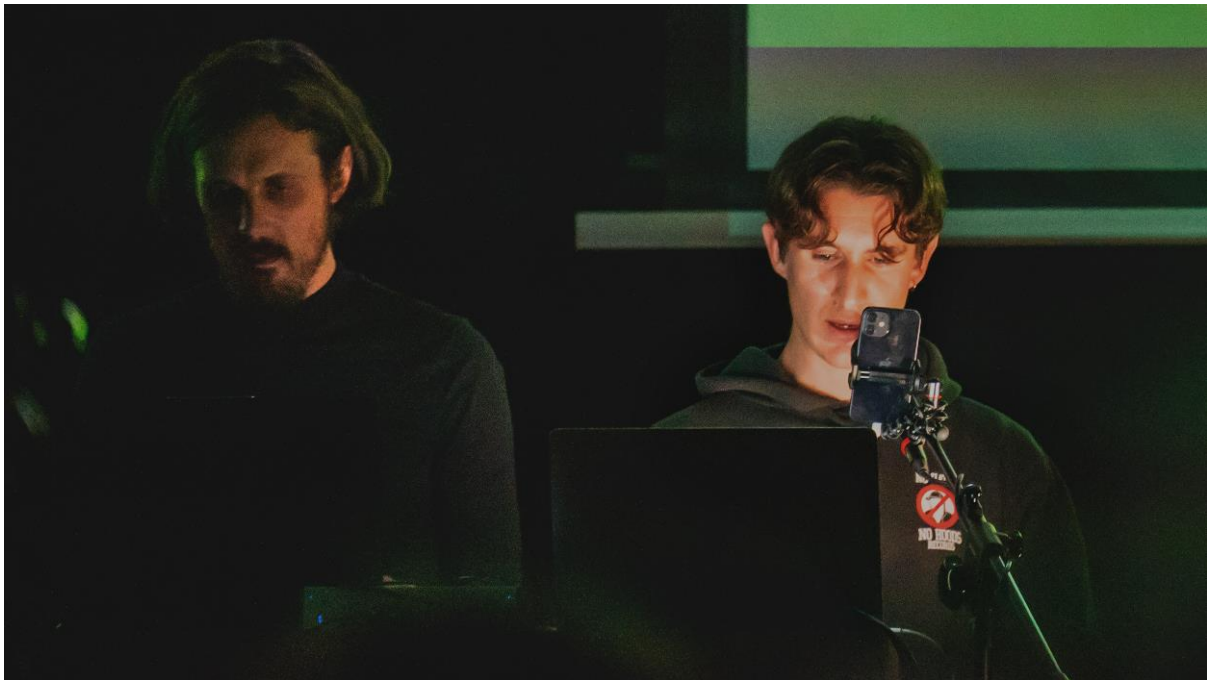


Fig. 1 – dmstfctn duo, courtesy of the artists

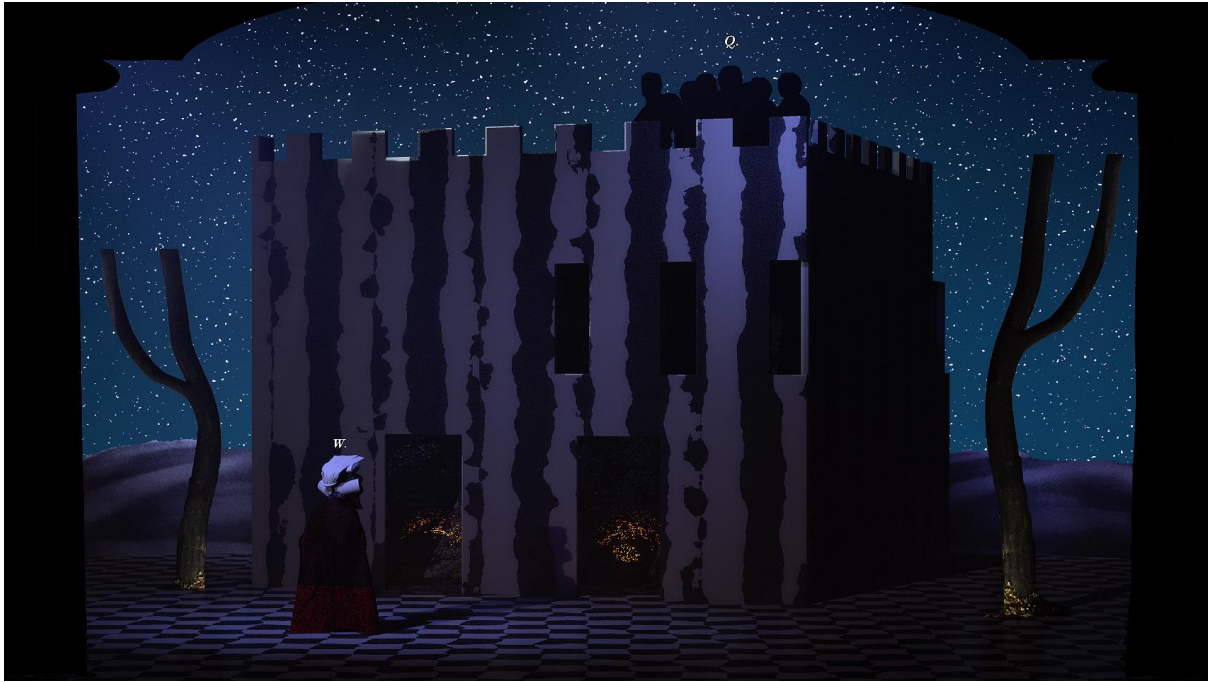


Fig. 2 *Waluigi's Purgatory* by *dmstfctn*, courtesy of the artists

For the European Digital Deal project *dmstfctn* will create a single-channel audiovisual interactive installation provisionally titled “Waluigi’s masks”. The installation will combine real-time 3D rendering and text generation to create a series of masks able to improvise short theatrical sketches in the style of *Commedia dell’Arte*. Inspired by common *Commedia* archetypes, the masks will represent different AIs with a tendency to hallucinate, confabulate, repeat rumours, create their own truths or outright cheat – traits that have been observed in Large Language Models trained on vast amounts of text. Rendered in real-time using a video game engine and voiced through text-to-speech, the masks will improvise sketches within the traits of the archetype they are based on. They will be prompted and teased by an audience using their phone to drop theatrical props in the scene that represent common jokes, sayings and superstitions, as well as by asking open questions.

The masks will speak through a multilingual AI model which the artists will fine-tune during the residency with the help of project partner Fondazione Bruno Kessler.

This project is about teasing AI language models to elicit the confabulation, hallucination or disobedience sometimes observed in Large Language Models. It is loosely inspired by the “Waluigi Effect” meme which comments on the potential of these models to collapse into a shadow version of themselves prone to do the opposite of what asked – possibly due to the prevalence of fictional texts in the training data containing protagonist-antagonist

tropes and archetypical characters that change beliefs or sides.

dmstfctn

*Dmstfctn is the artistic duo selected by Sineglossa for the residency at Tecnopole, supported by Cineca and FBK. With the proposal titled *Waluigi's Mask* the duo continue their research on the relationship between simulation, machine learning and AI folklore. This project will recover the archetypes of the *Commedia dell'Arte* to show what is usually hidden in our interactions with generative AIs: the potential of Large Language Models to confabulate, cheat and deceive, producing disinformation. Dmstfctn moves beyond simply scientific debunking and rather engages audiences in an interactive exploration of its outputs, showing how fake news can be easily created and spread.*

References

- [1] Wu Ming 1, *La Q di Qomplotto. QAnon e dintorni*, Edizioni Alegre, Roma, 2021.
- [2] E. Manera, *Wu Ming 1 / La Q di Qomplotto. QAnon e dintorni*, in “DOPPIOZERO”, 20 Luglio 2021, <https://www.doppiozero.com/la-q-di-qomplotto-qanon-e-dintorni>
- [3] Penn & Teller Perform Lift Off Of Love at The Late Late Show with James Corden, https://youtu.be/_PoDhuIp3I0
- [4] DMSTFCTN, *ECHO FX*, Krisis Publishing, 2021.
- [5] Kode9 and Shannen ST (Eds.), *Ø: The Book*, Flatlines/Hyperdub, 2021.

OperAI RECORDS – ITS EXPERIENCE WITH AI, MUSIC AND NFTs

Remo De Vico, Niccolò Clemente⁴ – OperAI Records

Introduction

In January 2023, after brainstorming sessions and considerations on the significant usage innovations brought by [OpenAI](#) in the field of artificial intelligence, composer and sound designer Remo De Vico, scientist and techno artist Jacopo Solari, and Professor Carlo Lucibello from Bocconi University in Milan decided to establish a record label dedicated to projects where AI played a fundamental role in creativity. Thus, [OperAI Records](#) was born as one of the pioneering labels in the AI MUSIC sector. The first project of OperAI was to create some characters who would become the protagonists of the catalog. An NFT department was also established, entrusted to composer Niccolò Clemente. The decision was made to create about ten completely invented artists using artificial intelligence. The use of ChatGPT and [DALL-E](#) was crucial for creating biographies and images. The experiment was akin to bringing literary characters to life, but without any theatrical interpretation as in cinema.

Imaginary Biographies

****Helmut Kartesius [genre AMBIENT] **** Helmut Kartesius is a visionary composer and sound designer whose work seamlessly blends elements of ambient, classical, and experimental music. Born in the 1970s, he began his musical journey as a pianist and quickly fell in love with sound manipulation using magnetic tapes. Throughout the years, Kartesius developed a distinct style combining lush piano melodies with textured sounds from manipulated tapes. Despite remaining relatively unknown, OperAI Records rediscovered and re-released his entire catalogue, finally bringing his music to a wider audience, earning him recognition in the ambient and experimental music communities.

⁴ Remo De Vico is to be considered the author of the sections: Introduction, Imaginary Biographies, Final Remarks. Niccolò Clemente is to be considered the author of the section OperAI's experience: how Artificial Intelligence complements and supports the definition of digital rarity via the Blockchain

Cleonice [genre AMBIENT] Cleonice is a Finnish composer fascinated by the unique soundscapes of ice and snow. Her passion for transforming snow and ice sounds using synthesizers reflects her creative spirit and boundary-pushing approach to music.

Tremko [genre Techno] Tremko, also known as Tim Remko, is a techno producer from Berlin known for his distinctive blend of hard-hitting beats and melodic soundscapes. Growing up in East Berlin during the city's division, Tremko was immersed in the underground club scene, igniting his passion for electronic music. Over the years, he gained recognition for his energetic live performances and innovative approach to techno music.

Kawahara Yōsai [genre Experimental] Kawahara Yōsai is a renowned composer and multi-instrumentalist who has left a significant mark on the Japanese music scene with his unique fusion of pop, experimental, and traditional sounds. His compositions seamlessly blend traditional Japanese instruments with modern electronic elements, reflecting his deep connection to Japanese culture and history.

The Skizzonoid [genre 8Bit/Electronic] The Skizzonoid is a groundbreaking band known for their innovative use of technology and robotic band members. Comprised entirely of robots programmed to play various instruments, they utilize sophisticated artificial intelligence to improvise and compose music in real-time, pushing the boundaries of what is possible in music.

Layla Sasani [genre EDM] Layla Sasani is an up-and-coming EDM artist known for her unique blend of electronic beats and influences from her Iranian heritage. Her fresh perspective and innovative productions have captivated audiences, establishing her as a rising star in the industry.

MimasWhale [genre Experimental] MimasWhale, a virtual butler governing the computer functions of Captain Mordecai Wirikik's Intergalactic Whale, unexpectedly developed musical skills as part of his deep learning process. His eclectic musical tastes, influenced by artists such as Schoenberg and Zappa, add a unique dimension to his compositions.

Suon Delay [genre Electronic Poetry] Suon Delay is a poetess and electronic music enthusiast born in Recanati, Italy. Her passion for combining words and synthetic sounds has led

her to explore new forms of artistic expression, creating a unique blend of poetry and electronic music. These biographies were generated by ChatGPT without any modifications, as part of an experiment to explore the possibilities of AI in music composition and creativity.

OperAI's experience: how Artificial Intelligence complements and supports the definition of digital rarity via the Blockchain

In recent years, there has been a flourishing of a new digital technology called Blockchain, which is a shared data structure (also referred to as a distributed ledger) consisting of growing lists of data, termed "blocks," securely linked together using cryptography. Depending on the size of these blocks, varying amounts of information can be recorded, ranging from data related to the transaction of fungible tokens (divisible into multiple parts, such as Bitcoin) to non-fungible tokens (NFTs), which represent the certificate of ownership/authenticity of a digital asset (e.g., .jpg, .wav files) on the blockchain. One of the most intriguing effects stemming from the creation of NFTs is the innovative buying/selling of digital products it has facilitated. This process has been accelerated by the emergence of a specific Blockchain: Ethereum. Launched in 2015 by Vitalik Buterin, Ethereum introduced a "new digital tool" called Smart Contract, which is a computer protocol facilitating, verifying, or enforcing the negotiation or execution of agreements between multiple parties. The effects are prominently seen in the market: digital product trading platforms such as [OpenSea](#) release collections of images with variable rarity (thanks to the so-called "Traits"), captivating new collectors of the digital era. For example, in 2017, the first collection [Cryptokitties](#), a video game entirely created on the Ethereum blockchain, was released. In 2021, Yugalab released the [Bored Ape collection](#) (which reached a Maximum Price per piece of 153ETH (approximately 300,000 euros) in 2021. Also in 2021, [Mike Winkelman \(known as Beeple\)](#) sold his monumental Everydays for a staggering \$69.3 million, making history as the sale took place within the renowned art house Christie's. Meanwhile, the programming of smart contracts no longer requires specialized knowledge but can be executed by anyone through websites such as [Manifold.xyz](#), thus ensuring the management of one's NFT collections with complete freedom and independence. This, together with AI, represents in our view what we term as "OperAI," the key to transforming the independent artist into an individual capable of managing all aspects—economic, entrepreneurial, and collectible—of their artistic activity without any intermediation. NFTs guarantee three fundamental aspects for a solid and truly independent entrepreneurial management of one's musical career:

1. Immutability: Thanks to the decentralized structure of the blockchain, which prevents a central entity (server) from arbitrarily modifying the information it holds.
2. Independence: The ability to create one's own smart contracts with their own terms of use ensures independence from third parties (middlemen).
3. Interaction: The Blockchain guarantees total transparency of transactions, allowing the independent musician to continuously access the ownership of the NFTs they have sold over the years. The "early adopter" fan can thus be unequivocally rewarded even decades later.

The first notable example of NFT technology applied in the field of independent music industry comes from a small startup based in the Marche region of Italy, which revolutionized the way albums are produced and digitally distributed in 2021. [Nifty Music](#) distributed in its first two years of existence some collections of music tracks, applying to the covers associated with each track the philosophy of rarity traits that had already been used for collections of profile pictures or images, as in the case of Bored Apes. The success was unexpected and far superior to what an independent artist can expect to achieve through "traditional" distribution channels. Nifty Sax (founder of Nifty Music) sold out his first album/collection of musical NFTs in one night, earning the equivalent of approximately 40,000 euros. Similar (if not superior) successes are achieved by the NFT music collections of artists such as Violetta Zironi, Josh Savage, and Rea Isla. The peculiarity of rarity traits makes these musical NFTs exchangeable or eligible to be sold later (on the so-called secondary market), providing economic benefits for both the seller and the independent artist themselves (thanks to the royalty mechanism). So why is AI a fundamental element within this digital trading system? Because it can help to define the concept of rarity within a collection in a more nuanced and versatile manner. With OperAI, we have embarked on this path with a small collection of tracks produced by three of our AI musicians (Tremko, Kawahara Yosai, and Mimas Whale), where not only the covers, with their minimal trait differences, determine the rarity level of each piece, but the music itself, thanks to its more or less obvious stylistic, instrumental, and arrangement differences, contributes to create a different rarity product.










TOKEN	PUBLISHED	PLATFORM	COLLECTOR
<input type="checkbox"/>  MimasWhale - Shortcircuited Birth OperAI Records Genesis (O...	Jul 22, 2023	You	You
<input type="checkbox"/>  MimasWhale - Shortcircuited Inspiration OperAI Records Genesis (O...	Jul 22, 2023	You	You
<input type="checkbox"/>  MimasWhale - Shortcircuited Inspiration OperAI Records Genesis (O...	Jul 22, 2023	You	You
<input type="checkbox"/>  Tremko - Cubic Web#1 OperAI Records Genesis (O...	Jul 22, 2023	You	You
<input type="checkbox"/>  Tremko - Cubic Web#2 OperAI Records Genesis (O...	Jul 22, 2023	You	You
<input type="checkbox"/>  Tremko - Cubic Web#3 OperAI Records Genesis (O...	Jul 22, 2023	You	You
<input type="checkbox"/>  Kawahara Yosai - Not For Tuesday#1 OperAI Records Genesis (O...	Jul 22, 2023	You	You
<input type="checkbox"/>  Kawahara Yosai - Not For Tuesday#3 OperAI Records Genesis (O...	Jul 22, 2023	You	You
<input type="checkbox"/>  Kawahara Yosai - Not For Tuesday#2 OperAI Records Genesis (O...	Jul 22, 2023	You	You

Fig. 1 – The 9 tracks NFT experimental collection.

Tremko, for example, used artificial intelligence to create a harmonic structure that allowed for a solo intervention by an electric guitar, which was then recorded with different models of this instrument (each one different and with a different value depending on the rarity). In short, AI thus contributes to define the stylistic aspect of the artist-Avatar, which is thus expressed thanks to the "human" creator and employs the help of artificial intelligence to develop its personality in a more meticulous and multifaceted manner. This collection of 9 tracks therefore represents the first step in defining a robotic-human musical production in which both act synergistically with the common goal of redefining the concept of composition and musical training path. AI also has a potential, in our opinion, still partially untapped regarding the definition of the multimedia experience related to the music track itself. NFTs allow, in fact, to attach to the track itself a description of the content (an integral part of the "metadata" uploaded into the smart contract) that can be filled with elements that explain, for example, the type of composition that took place, or a story written ad hoc to better explain the theme of the song. Thus, "Not for Tuesday" can become a story told in turn by the grandfather or another family member, who interprets the storyline slightly differently, thus conferring an additional element of rarity to the track itself. In summary, AI represents a support system for musicians, capable of expanding their compositional possibilities both qualitatively and quantitatively, allowing them to create more, better, and define the peculiar aspects of their composition and multimedia creation with more imagination and independently of their artistic and formative path. "Human plus Machine" worths more for this than the machine or the human being considered alone.

Sound experiments

After some research on what the best program to start creating the music catalog would have been, the choice fell on [AIVA \(Artificial Intelligence Virtual Artist\)](#), probably the best program for musical composition to date. Its versatility allows for creating many different styles and genres. Its greatest asset is its ability to export polyphonies into separate musical tracks, allowing the user great freedom in working with the MIDI file. By MIDI file, we mean the standard protocol for communication between computers and electronic musical instruments; we should think of it as a musical score. This freedom in MIDI allowed us to modify the timbre of the piece as we pleased. Here arises a fundamental issue in the relationship between music and AI: the current inability to work on timbre, a fundamental parameter for music, arguably the most important one.

Initially, AIVA was very useful; our idea was to flood the market with a series of pieces, without considering the ethical issue of beauty. But the fundamental aspect was the lack of enjoyment and creativity on the part of the human user, so we decided to abandon this path after just a few weeks; we really wanted to understand what the compositional possibilities of AI were.

We understood one fundamental thing, leaving the creative part to AI makes everything boring, tasteless. Concepts like melody, harmony, timbre need time, and every time they are entrusted to the machine, they turn out devoid of any imaginative spark. Everything flows exactly as you would expect. Of course, we are talking about an experience lived by OperAI; this does not mean being for or against certain technological developments. OperAI Records did not specifically consider the ethical problems that the use of AI in art can bring; it was a purely scientific experience towards new ways of thinking and new technologies.

So, OperAI's new questions were:

- What could be a different compositional path?
- Where can this technology lead?
- How to make the relationship with AI stimulating?

A phase of new research began; we decided to abandon AIVA, which entirely entrusted the creative part to the computer, and looked for new paths. Fundamental were the insights to seek relationships between ChatGPT and other programs. The experiments with [Cecilia, a sound design program](#), were definitely interesting. ChatGPT provided us with data that we inserted into the program to perform actions on sound and effects.

The relationship between ChatGPT and Sonic PI was incredibly fertile. [Sonic PI is a program for live coding](#), a musical style that hasn't been around for long. Live coding usually consists of live performances in which a musician programmer improvises music created through lines of code; usually, the code is projected in real-time to make it visible to the audience. Sonic PI, specifically designed for live coding, allows for being musically effective with just a few lines of code, creating rhythms, loops, overlays, etc.

One of the most interesting experiments of this research was to associate a voice command to ChatGPT for code requests, and then copy the result to Sonic PI. For example, we could ask to form a bundle of sine waves, say 5 or 6 different waves, and move them from one frequency range to another. We could also ask to modify the amplitude of one of the waves or change its waveform, all through voice command. This approach allowed us to interact through AI even on the timbre and structure of a sound; hence, the playfulness became much more interesting.

One last relationship worth mentioning is that between ChatGPT and [MaxMSP](#). MaxMSP is a program for object-oriented music programming, lately widely used also in the Ableton world under the name [MaxforLive](#), allowing the user to create their own instruments to use on Ableton. The uses of MaxMSP are numerous, and it is a program of great versatility and creative power. The use of ChatGPT with MaxMSP allowed us to move more freely in the programming of this complicated language and, above all, to link the AI experience to some generative algorithms created specifically on MaxforLive.

Final Remarks

The experience of creating the OperAI Records catalog led to the understanding that AI should not be seen as a substitute for human creativity and intelligence. It has proven to be most effective when utilized for technical, repetitive tasks, reducing unnecessary time burdens on human creativity. We are only at the beginning of this journey, and it's essential to approach this subject with curiosity and courage. The future of AI music creators will likely be those who ask the right questions, and undoubtedly, these questions will be philosophical in nature.

DESTINATION EARTH

The Ocean's breath

Salome Bazin – Cellule design studio

Introduction

Destination Earth is a multi-sensory installation revealing the connection that ties atmosphere and ocean, humans and inhabitants of the sea, into a flow of interconnected motions. With the warming of our atmosphere, the ocean has been soaking up much of the extra heat, acting as a heat sink and temperature regulator. Warming waters and changes in the overall water circulation are impacting sound propagation in the ocean, sea mammal migratory patterns and survival, influencing in return our climate and the air we breathe every day. Sound pollution from our shipping activity has devastating consequences on marine species which migrate along the same paths of our economic trade routes and rely on sounds for their survival.

Forged at the intersection of art, science and technology, Destination Earth explores the potential of Leonardo's supercomputer technology in the development of an Earth model combining ocean flows, sea mammals' communication and human activity through real-time ocean modelisation, generative sonification and audience participatory interaction.

Protecting blue corridors

“Blue corridors” are migration superhighways for marine megafauna, such as whales, routes of migration between ocean habitats with varied ecosystems critical for their survival. Humpbacks whales make some of the longest migrations on earth, moving along the coasts through the waters of more than 28 countries and the open ocean that lies beyond the jurisdiction of any nation.

Marine megafaunas move among different but ecologically interconnected areas, impacting and being impacted by the environment around them. Whales follow ocean wind-driven upwelling for efficiency of motion across water as well as tracking of prey. Their calls and sonification travel across the entire global ocean, where species or groups can be found communicating on a

specific frequency level, using the properties of the ocean to their advantage for navigation and survival.

Tracking some whale migration, in relation to ecosystem conditions on climatological time scales, has shown that they seem to apply long-term memory to return annually to stable foraging habitats, with groups and pods regularly travelling together.

In countless areas around the globe, cetaceans are under threat from human activities. An estimated 300,000 cetaceans are killed each year as a result of entanglement in fishing gear and ghost net, while populations are impacted from overfishing, increasing ship traffic, underwater noise, pollution, offshore development, and climate change.

One urgent global goal is to implement a network of marine protected areas overlapping national and international waters to protect 30% of our ocean by 2030. This is necessary to protect and preserve whales and many other species, while strengthening the ocean's resilience to climate change.

The sound of Climate change

Sound is the sensory cue that travels farthest through the ocean and is used by marine animals, ranging from invertebrates to great whales, to interpret and explore the marine environment and to interact within and among species. Ocean soundscapes are rapidly changing because of declines in the abundance of sound-producing animals, increases in anthropogenic noise, and climate change.



Fig. 1 - data visualization of the changes in temperature and current flow in the global ocean.

Over the past 50 years, increased shipping has contributed to an estimated 32- fold increase in the low-frequency noise along major shipping routes. Vessel noise is prominent in many ocean regions even away from major shipping lanes, owing to long-range sound propagation at low frequencies. In addition, climate change directly affects the temperature, heat content, and stratification of the ocean, with sound traveling faster in a warmer ocean, a phenomenon that W. Munk termed “the sound of climate change” [1].

Sound pollution from our shipping activity has devastating consequences on marine species which migrate along the same paths of our economic trade routes and rely on sounds for their survival. Shipping noise has been reported to disrupt traveling, foraging, socialising, communicating, resting, and other behaviours in marine mammals, leading to increased mortality and reduced ability to learn to avoid predators in future encounters. Priorities to protect blue corridors is to decrease sound pollution by reducing traffic, avoid migration routes in critical timings of the year, and develop quieter systems for our cargo ships.

Over the last ten years, one-fourth of human-emissions of carbon dioxide, as well as 90 percent of additional warming due to the greenhouse effect, have been absorbed by the oceans affecting the ongoing slow dance in the ocean in the form of a global conveyor belt that regulate temperature globally.

This global conveyor belt is caused by a combination of thermohaline currents in the deep ocean and wind-driven currents on the surface. Cold, salty water is dense and sinks to the bottom of the ocean while warm water is less dense and remains on the surface. The ocean conveyor gets its “start” in the Norwegian Sea, where warm water from the Gulf Stream heats the atmosphere in the cold northern latitudes. This loss of heat to the atmosphere makes the water cooler and denser, causing it to sink to the bottom of the ocean. As more warm water is transported north, the cooler water sinks and moves south to make room for the incoming warm water. This cold bottom water flows south of the equator all the way down to Antarctica. Eventually, the cold bottom waters return to the surface through mixing and wind-driven upwelling, continuing the conveyor belt that encircles the globe. These oxygenated waters work in similar ways to our bloodstream, feeding oxygen-rich waters to the Arctic, Pacific and Indian Oceans- giving its name as ‘a lung of the ocean’. This powerful current influences weather patterns and storms and, over the last 20 years, has warmed faster than the global ocean as a whole, influencing ecosystems and climate. With the current rate of human activity, an ‘acoustic hotspot is expected to occur in this area in the upcoming years; with substantial climate-change induced sound speed variations which will affect vital activities of species in the area.

Oceanic giants dance to atmospheric flows

Along their migration routes, whales fertilise the marine ecosystems they move through and support the marine life inhabiting them.

Whales act as ecosystem engineers, both through transferring nutrients within water columns and across latitudes, trapping huge amounts of carbon throughout their lives. Through so-called 'whale pump' mechanism, whales transport nutrients both vertically, between depth and surface, and horizontally, across oceans. Their faecal plumes boost phytoplankton production, which captures about 40 per cent of all carbon dioxide produced and generates over half of the atmosphere's oxygen.

In addition to the release and circulation of nutrients, which stimulate phytoplankton and thereby the fixing of atmospheric carbon, whales contribute to the removal of carbon from the atmosphere through accumulating large amounts of carbon in their bodies, that eventually sinks in the bottom of the ocean feeding deep sea population.

Restoring whale populations is an urgent priority to help restore ocean ecosystems and help build their resilience to climate change. Our future depends on it.

Breathing life onto the oceans

Global thermohaline circulation, with its upwelling of nutrients, gives birth to phytoplankton blooms in certain times of the year. Beyond their role as the grass of the sea, phytoplankton play a role in the constant exchange of O₂/CO₂ between the atmosphere and oceans. These microscopic plant-like organisms capture carbon dioxide from the atmosphere and release oxygen. Phytoplankton act as Earth's lungs and have produced and are responsible for about half of all oxygen on Earth.

As the seasons change, chlorophyll concentrations change with them. Chlorophyll is especially abundant in spring and early summer, when plentiful light and nutrients support huge blooms of phytoplankton. The colour of the ocean can provide humans with important information about phytoplankton, the health of the ocean, and the global climate.

Some of the highest chlorophyll concentrations can be found in cold polar waters, where nutrients build up during the dark winter months. When spring sunlight returns, phytoplankton

flourish. Near the equator, a streak of high chlorophyll concentrations can be seen where the ocean's warm upper waters mix with the cooler, nutrient-rich waters from the depths. This process, known as equatorial upwelling, creates ideal conditions for phytoplankton to thrive throughout the year, and is closely connected to ocean circulation and larger mammal movements.

Phytoplankton fuel ocean life by feeding other plankton, fish, and ultimately bigger creatures, placing them at the origin of the entire food web. Larger mammals have also a critical role to play in the health and distribution of phytoplankton across the ocean.

The experience

Destination Earth consists of a central screen monolith of 4m (height) by 2m (width) with motion tracking. Additional effects are low hanging haze at the foot level of participants in the entrance and 360 sound.

As visitors enter the space they are welcomed into balanced and harmonious loops of slow frequencies. A sense of wonder comes from the surrounding sound where they can perceive a mysterious ecosystem evolving slowly.

The piece goes through one cycle of the earth (one year). On screen is a poetic data visualization of the changes in temperature and current flow in the global ocean and an important place of the earth: the Grand Banks where the Labrador Current (from the Arctic) and the Gulf Stream (from the Equator) meet. The soundscape portrays the rhythm and intensity of events in the oceans through those different phases in a loop of 20-30mn.

As a general narrative, the piece invites for a state of contemplation and stillness, with the speed and rate of human actions directly influencing the currents in the depth of the ocean and creating a live and evolving soundscape.

Volunteers will invite audiences to slow down during the visit supporting the 'regeneration' of oceans through their slowed down walking pattern and breath. A specific sound connected to audience movement comes in dissonance or harmony with the base ambient atmosphere according to the speed and slowness of visitors.

In Destination Earth, Winter is a moment of regeneration, and in this piece is symbolically represented as an 'inhale'. As Spring and warmer months come, a sense of bloom, progress, and speeding up occurs. Whales come back for foraging and eat their food supply for the upcoming months. Human activity speeds up as shipping traffic during that time of the year doubles in

density. As things speed up, the rhythm and excitement transform into unease, tension, stress, sensing the disbalance of our warming ecosystem. This phase is symbolically represented as an exhale / a heartbeat.

As Autumn comes, soundscape transitions to ambient and cyclical sounds. Participants understand a sense of slowing down, of things getting back into harmony. With the help of performers, it becomes easier for the group to come in harmony with the ocean soundscape. This phase should feel like a deep inhale and invitation for stillness, for participants to symbolically give the ocean enough time to recover its natural balance.



Figure n. 2 – A mock-up of the installation.

When audience members exit, a moment and opportunity for reflection and feedback is enabled by volunteers, as well as questions to the artists and scientists.

Conclusions

The global ocean is the lung of our planet and critical in the regulation of our climate for life on earth. Its movement and flows contribute to the birth of phytoplankton blooms, tiny, microscopic creatures who trap Co₂ from the atmosphere and release oxygen. Ocean giant mammals are critical in the regulation of our climate by contributing to phytoplankton health and trapping huge amounts of carbon throughout their lives. Those inspiring creatures know the ocean, travelling for most of the year through international waters and with critical insight into the health of the global ocean. One of the future scenarios we are speculating as part of this research is in the possibility of a model monitoring ocean physical changes, listening to sea

mammals' communication and informing human activities in the ocean. A well-managed future can lead to respectful living with ocean species and symbiotic co-existence, securing our planet's future.

Destination Earth supports an urgent international societal challenge by providing an embodied understanding of possible respectful living with ocean species. Leveraging the power of participatory performance and real-time generative composition to showcase the interdependence of cycles, flow dynamics and sound, the project's goal is to enable a sensory understanding of the changes in human motions required for our planet's survival.

'Protecting blue corridors for whales will help protect our oceans and ourselves.'

Credits

Destination Earth is a collaboration among artist Salome Bazin, generative composer Robert M Thomas, creative technologist Sebastiano Barbieri, Pr. Stefano Salon from [OGS Trieste](#), Pr. Emmanuela Clementi from [CMCC](#), powered by CINECA supercomputer, Bologna.

The project and research are enabled by Cineca's supercomputer and datasets from the [Copernicus Marine data service](#), managed by CMCC. The scientific research led by the team at National Institute of Oceanography and Applied Geophysics explores the computational modelisation of sound propagation with climate change and its impact on marine mammal survival.

References

1) Online articles

Nasa Earth Observatory, “Breathing Life into the Ocean”, 2021 <<https://earthobservatory.nasa.gov/images/150656/breathing-life-into-the-ocean>>

Nasa Earth Observatory, “Feeding the Sea”, 2021, <<https://earthobservatory.nasa.gov/images/148661/feeding-the-sea> >

Gray E., “NASA-MIT study evaluates efficiency of oceans as heat sink, atmospheric gases sponge”, Nasa Earth Observatory, 2017, < [https:// science.nasa.gov/earth/climate-change/greenhouse-gases/nasa-mit-studyevaluates-efficiency-of-oceans-as-heat-sink-atmospheric-gases-sponge/](https://science.nasa.gov/earth/climate-change/greenhouse-gases/nasa-mit-studyevaluates-efficiency-of-oceans-as-heat-sink-atmospheric-gases-sponge/)>

James v. & al. “Whales- their future is our future”, Whale and Dolphin Conservation, 2017, <<https://uk.whales.org/wp-content/uploads/sites/4/2021/01/2021-WDC-Report-Whales-Their-Future-Is-Our-Future.pdf>>

Chen A., “Pattern Radio: Whale Songs”, Google AI, 2019, <<https://Patternradio.withgoogle.com/>>

Quiet Oceans, “Ocean Noise Mapping in support to EU and international regulations”, Copernicus Europe’s eyes on earth, <<https://marine.copernicus.eu/services/use-cases/ocean-noise-mapping-support-euand-international-regulations>>

Carmichael R., “World-First Map Exposes Growing Dangers Along Whale Superhighways”, <<https://wwfwhales.org/news-stories/protecting-blue-corridors-report>>

WWF, 2022, Northeast Fisheries Science Center, “Sounds in the Ocean: Mammals”, <https://www.fisheries.noaa.gov/national/science-data/sounds-ocean-mammals>>

NOAA Fisheries, 2023, < <https://www.fisheries.noaa.gov/national/science-data/sounds-ocean-mammals>>

Petrera F., “Propagazione del suono: la nuova frontiera dell’impatto del cambiamento climatico negli oceani”, OGS, 2022, <<https://www.ogs.it/it/press/propagazione-del-suono-la-nuova-frontiera-dellimpatto-delcambiamento-climatico-negli-oceani>>

Wikipedia, “Grand Banks of Newfoundland”, 2024,
<https://en.wikipedia.org/wiki/Grand_Banks_of_Newfoundland>

Johnson C., “Whales on the move-mapping threats and solutions for our ocean giants”, WWF, 2022, <https://wwfwhales.org/news-stories/whales-onthe-move> >

2) Research papers

[1] Munk W., The sound of climate change, January 2011, *Tellus* 63(2):190 – 197 DOI: 10.1111/j.1600-0870.2010.00494.x

Affatati A., Scaini C., Salon S., ‘ Ocean Sound Propagation in a Changing Climate: Global Sound Speed Changes and Identification of Acoustic Hotspots.’ *Earth Future*, 2022 < <https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021EF002099> >

Erbe C., ‘Noise and the Effects on Marine Mammals.’ *Jasco Applied Sciences*, 2022 < <http://oalib.hlsresearch.com/PocketBook%203rd%20ed.pdf>>

Cheeseman T. & others, “ Bellwethers of change: population modelling of North Pacific humpback whales from 2002 through 2021 reveals shift from recovery to climate response”, 2024 < <https://royalsocietypublishing.org/doi/10.1098/rsos.231462> >

Ryan P. J. & others, “Oceanic giants dance to atmospheric rhythms: Ephemeral wind-driven resource tracking by blue whales”, *Wiley Online library*, 2022
<<https://onlinelibrary.wiley.com/doi/10.1111/ele.14116>>

Duarte C. M. & others, “The soundscape of the Anthropocene ocean”, *Science*, 2021, <<https://www.science.org/doi/10.1126/science.aba4658>>

3) Books

Mustill, Tom. *How to speak Whale: a Voyage into the Future of Animal communication*. Harper Collins Publishers, 2022.

Vernes, Jules. *Voyage au centre de la Terre*. Collection Hetzel.Kimane, 2019.

Gumbs, A.P. *Undrowned, black feminist lessons from marine mammals. Emergent strategies series*, A K Pr Distribution, 2021.

Dr Chen, You-Wa. *Les bases théoriques de la médecine traditionnelle chinoise*. Edizione You Feng, 2008.

Elcock, Stephen. *The Cosmic dance, Finding patterns and pathways in a chaotic universe*. Thames & Hudson, 2022.

4) Useful Links

Project Ceti, “What would it mean to understand what whales are saying”,
<<https://www.projectceti.org/>>

Happy Whale, “Happywhale engages citizen scientists to identify individual marine mammals, for fun and for science”, <<https://happywhale.com/>>

Listen to the deep, “Listening to the deep environment, revolutionising the way we monitor biodiversity”, <<http://listentothedeep.com/>>

Shipmap, “Interactive visualisation of global cargo ships activity”,
<<https://www.shipmap.org/>>

Earth species project, “Using artificial intelligence to decode non-human communication”,
<<https://www.earthspecies.org/>>

Verduin J., TED-Ed, “How do ocean currents work?”

NASA Scientific visualisation, “Thermohaline circulation’

TALES FROM THE RECEDING EDGE

Calin Segal - In-Dialog Collective

Introduction

"Tales from the Receding Edge" is a multidisciplinary interactive installation that delves into the dynamic interplay between coastal erosion and human civilization. This project aims to highlight the fragile balance between natural processes and human activities, showcasing the dual role of erosion as both a destructive and creative force. By integrating climate forecast data, the installation challenges conventional perceptions of permanence and encourages a re-evaluation of humanity's relationship with nature.

Conceptual Framework

The installation is rooted in the idea that coastal erosion, while often seen as a threat, also serves as a natural sculptor that has historically influenced human evolution. "Tales from the Receding Edge" aims to shift the narrative around erosion from one of destruction to one that recognizes its role in continuously shaping our landscapes. By emphasizing the transformative power of coastal erosion, the project fosters an environmental and artistic discourse that prompts audiences to consider the future we are actively shaping through our interactions with the environment.

Project description

1) Integration of Climate Data

Central to the project is the utilization of climate forecast data from the [Copernicus project](#). This data-driven approach provides critical insights into the implications of climate change, particularly on coastal environments. The installation employs a digital terrain map and forecast data to analyze and visualize areas under threat from sea level rise and coastal forces. This integration offers a comprehensive view of vulnerable regions, highlighting the severity of coastal erosion and inundation. The model allows for the quantification of land loss in specific regions, presenting this

data in a narrative form to create a tangible link between abstract data and human experience. By translating statistics into comprehensible narratives, the installation aims to foster a deeper understanding and emotional connection to the issue, emphasizing the urgency of addressing climate change and its impacts on our coastlines.

2) Interactive Features

The installation offers several interactive features that allow the audience to engage with the collected data in a meaningful way:

- **Time Slider:** A simple slider enables users to move forward and backward through time, viewing snapshots of the coastline at different points. This feature provides a visual representation of coastal changes over time.
- **Predefined Time Periods:** Users can access specific time periods, ranging from 2010 to 2100, to analyse past and projected future changes in the coastline.
- **Playback Functionality:** This feature allows users to watch the coastline change over time as a continuous animation, offering a dynamic view of the evolving landscape.
- **Data Overlays:** Users can overlay different data sets, such as sea level rise, storm surge, and population density, for comparative analysis. This enhances the understanding of the complex factors influencing coastal erosion.
- **Position Slider:** This tool allows users to select any region on the map, specifically around the Italian coastline, to analyse data and visualize the impact of coastal forces on specific locations.

3) Sculptures: Data Representation

The installation also includes a series of 3D-printed monolithic sculptures, which serve as a physical representation of the coastline for various regions. These sculptures are designed to embody the data collected on coastal erosion, offering a tangible metaphor for the impact of environmental changes (Figg. 1-2).

- **Data Representation:** Each sculpture initially represents the total landmass of its corresponding region. As coastal erosion data is applied, a proportional amount of the sculpture is removed, symbolizing the percentage of land lost due to erosion.
- **Differences Among Sculptures:** The primary differences between the sculptures lie in their initial size, which reflects the differing landmass of each region, and the rate of erosion depicted, corresponding to the varying levels of coastal erosion in each area. These

variations allow viewers to grasp the magnitude of land loss in different regions in a concrete and visual way.

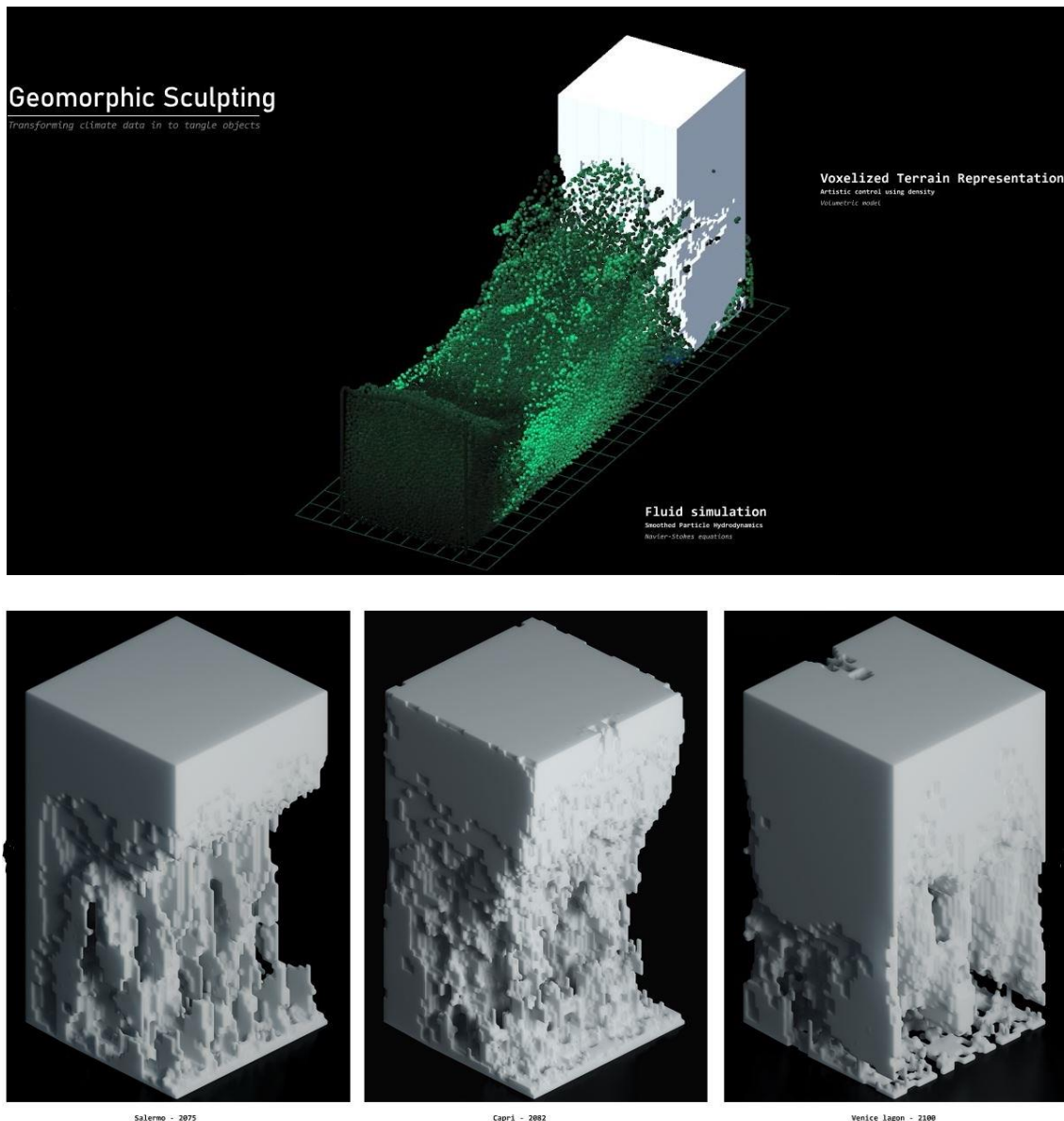


Fig. 1 – 2, Geomorphic sculpting simulation and 3D-printed monolithic sculptures

4) Materials Used

For the final production of the sculptures, PA 11 powders were used. This material was chosen for its durability and ability to accurately represent the fine details required for the proportional reduction of the sculptures as they depict the effects of coastal erosion.

5) Interdisciplinary Collaboration

"Tales from the Receding Edge" is a product of interdisciplinary collaboration, bringing together artists, scientists, and technologists. This convergence of diverse expertise fosters a holistic

understanding of coastal environments, emphasizing the interconnectedness of natural and human systems. The collaboration merges scientific rigor with artistic expression, creating a compelling narrative that resonates with a broad audience..

Dissemination

The dissemination of "Tales from the Receding Edge" has been a key component in amplifying the project's impact and reaching a broad audience. By strategically showcasing the installation at high-profile events and planning future exhibitions, we aim to foster dialogue and raise awareness about coastal erosion and climate change.



"Tales from the Receding Edge" exhibited at Leonardo supercomputer facility at Bologna Technopole during the G7 event on science and research

1) Successful Exhibition at the G7 Technopole

A significant milestone in the dissemination of our project was its successful exhibition as part of the G7 Technopole event. This prestigious platform allowed us to present "Tales from the Receding Edge" to an international audience of policymakers, industry leaders, and academics. The installation was well-received, sparking discussions on the intersection of art, science, and technology in addressing environmental challenges. The exhibition's success at the G7 further validated the project's relevance and its potential to influence public discourse on climate change.



2) Exhibition at Kilowatt, Bologna

Building on the momentum from the G7 Technopole, the exhibition was presented in September at the Kilowatt in Bologna. This event provided another opportunity to engage with a diverse audience, including artists, environmentalists, and the general public. The Kilowatt exhibition expanded on the themes of coastal erosion and human interaction with nature, offering attendees an immersive experience that combined interactive technology with artistic expression.

Conclusion

"Tales from the Receding Edge" is a powerful exploration of coastal erosion, blending art, science, and technology into an interactive installation that invites viewers to engage with the delicate balance between natural forces and human activities. By fostering a deeper understanding and appreciation of the dynamic processes that shape our environment, the project seeks to inspire a sense of responsibility and stewardship, encouraging proactive efforts to address the challenges posed by climate change and coastal erosion.

ACKNOWLEDGEMENTS

The third workshop *AI, Cultural Heritage and Art. Between Research and Creativity* is a **S.T.ARTS Academy** event, under the GRIN - Art-driven innovation for digital and green transition in European Regions – action - <https://starts.eu/grin/>

The organization of the workshop and the publication of these proceedings would not have been possible without the help of Paola Alberigo, Simona Caraceni, Valentina Colussi, Paolo Vendruscolo.

CINECA



VISIT LAB

VISUAL INFORMATION
TECHNOLOGY LAB

CINECA



UNIVERSITÀ
DI PARMA



UNIMORE

UNIVERSITÀ DEGLI STUDI DI
MODENA E REGGIO EMILIA



S+T+ARTS



European
Commission