I. Proposal's context

Violin made by Zacharias Fischer, 1767, Grassi Museum of the University of Leipzig

Cello made by the Voller brothers, late 19th c. , Musée de la musique, Paris

In France, 1.600.000 musical instruments are sold every year for a turnover of 500 million euros. 400.000 instruments are sold online and 200,000 between private individuals. Unfortunately, the traffic of instruments has increased considerably on the Internet without any traceability or possible identification of the instruments. Manufacturers and instrument makers have been trying for years to structure a standard instrument maintenance logbook without success. Innovative digital technologies and image recognition can be harnessed to create a comprehensive and transferable tool that allows for the history of an instrument to be archived and updated over time while protecting it against theft in a universal way.

This original project brings together an expert in music and musicology, Frederic BILLET, and a researcher in artificial intelligence, Victoria EYHARABIDE, to develop an innovative and interdisciplinary joint research programme. **Focusing on the string quartet instruments (violin, cello, viola), which are frequently stolen due to their high commercial value and easy portability, we propose creating an AI-based identity card for musical instruments. This project will enable musicians and instrument makers to identify their instruments and reduce instrument trafficking by allowing for the individual authentication of instruments**.

II. Objectives

a. Data: music, images, and instruments

• **Key objectives:** Researchers need data to develop innovative AI-based methods, i.e. annotated texts and images. Therefore, we will first curate datasets to enable machine learning work on the authentication of instruments from images and texts of musical instruments. Second, we will make those datasets available to the scientific community using IIIF manifests. The International Image Interoperability Framework (IIIF) has become the de-facto standard in image delivery for cultural heritage institutions on the Web. Thus, the image datasets are preserved and maintained over time and remain available for reuse to the scientific community at large

• **Scientific barriers to be lifted:** The organization, integration, and annotation of iconographical data is challenging since images, and textual descriptions describing the ***instruments are dispersed worldwide*** in public and private collections. Another drawback may be to ***homogenize dissimilar and scattered data*** (descriptions in different natural languages) and images (different formats, resolutions, and views) between several collections. Another challenge will be ***agreeing on a validation protocol and metrics to evaluate the manual annotations*** (intra and inter-annotator agreements). In this sense, CERES's technical and methodological support will be essential to curate, analyze, annotate, and publish the digitized corpora.

b. Hybrid AI: computer vision and knowledge engineering for cultural heritage

• **Key objectives:** The recognition of stringed musical instruments is usually carried out by expert luthiers who authenticate an instrument mainly based on shapes, that of the instrument as a whole, and the soundhole proportions. The extraction of conceptual entities in datasets (i.e. shape, model, color of the varnish, profile of the table, aspect of the head, etc.) through computer vision and knowledge engineering will allow for the analysis and cartography of semantic elements, which can then be included in a more extensive "***identity card***" of the instrument.

We propose a new approach that combines computer vision and knowledge graphs to create a novel musical instruments' identification method. Knowledge Graphs embeddings [10,11] form a projection of those knowledge graphs into a lower dimension, where entities and relations are represented into continuous vector spaces. Our method will incorporate these vector representations as a key ingredient to guide the identification process.

• **Scientific and technical barriers to be lifted:** Some difficulties may arise from the amount of data available and/or the difference in quality between images. Depending on the case, augmentation techniques will be considered, and eventually, transfer learning methods will be implemented. A difficulty related to the text can come from the different languages used by the instrument museums in the descriptive notes. For this, a collaboration is envisaged with the multilingual scientific team of the *Musée de la Musique - Philharmonie de Paris*.

III. Position of the project as it relates to the state of the art

•**Musical and musicological data-driven approach:** While some small datasets exist for musical instruments, they are generally only used for working with audio data. Iconographic datasets are rare and often incomplete. This project's innovative character and originality are that the created datasets will be openly available for reuse. F. Billet et V. Eyharabide recently published two papers using MIMO images and computer vision techniques in a more diachronic approach [4,5]. They also work together in the development of MUSICONIS, a music iconography database.

•**Computer vision-based cultural object authentication:** Conventional machine learning methods are mainly based on a single view and do not take into account other sources of information, such as the descriptors of images (color, shape, or texture) or textual data, which can dramatically improve accuracy and performance in the authentication of instruments. To our knowledge, this is the first implementation of biometric authentication techniques for the recognition of musical instruments.

• **Knowledge engineering for cultural heritage:** Knowledge graphs have flourished in the Humanities since they are suitable for dealing with cultural heritage data. However, examples concerning musical instruments fail to unify existing models or to propose efficient methodologies to transfer data from diverse sources to these data models. While knowledge engineering programs have already been set up with MIMO data [8,9], none have an application used in real life by musicians. V. Eyharabide has strong expertise in developing ontologies related to cultural heritage and music [7].

IV. Methodology

The first step of our method starts by ***creating a knowledge graph*** that describes our source dataset of musical instruments, such as the parts of the instruments, the material used to create them, or information about the owners. We will build on our work [7] to implement complete transcultural and diachronic knowledge graphs. Then, jointly with domain experts, we will ***link the images to concepts in the graph*** using the musicological descriptions of instruments. We will ***create an embedding of each musical instrument*** based on its connections and ***combine these knowledge graph embeddings with visual embeddings*** from the images with the other nodes in the graph. These embeddings will allow us to generate concept-level embeddings, which will later be used as the anchors for training neural networks. Later, we will ***train neural networks with the combined embeddings as anchors*** using an extension of Fisher's linear discriminant. The idea will be to minimize the overall distance between the center of visual concept embeddings and the normalized center of the knowledge graph concept embeddings. This method will enable the encoder part of the network to extract class-informative and structured latent space, allowing the classifier to generalize better to other domains. As a result, we will ***explore knowledge graph embedding, visual embedding, or combined embedding to calculate the similarity between instruments***. The results will be validated with musicologists, musicians, and instrument makers.

V. Role of each supervisor as well as the scientific skills provided.

In this project, the future Ph.D. student will be supervised by two scholars with complementary expertise. As a domain expert, Billiet specializes in instrumental practice and musical iconography [1-3]. Eyharabide provides expertise on knowledge graphs applied to musicology [7] and transfer learning for medieval musical instruments recognition [4-6].

VI. Candidate profile

With a master's degree in computer science, the candidate will have the necessary skills to work at the crossroads of knowledge representation and computer vision. An interest in cultural heritage and musicology would be an added asset. Motivation, self-discipline, and maturity are required for this highly interdisciplinary and collaborative project. The candidate should be comfortable with communicating in French or English (oral and written). No prior knowledge in the musicology domain is needed. However, the candidate will work with musicologists and history art experts to model/validate the algorithms' results. Knowledge of any of Protégé, PyTorch BigGraph, YOLO, and Tensorflow would be beneficial.

VII. References

Publications of Frederic BILLET related to the project

[1] F. Billiet, E. C. Block (edit), ***Lexique des stalles médiévales***. *Lexicon of Medieval Choir Stalls*, Brepols Publisher, 2019.

[2] F. Billiet, **L'interopérabilité des bases Musiconis, Tunicomus et Musicomed pour une nouvelle compréhension de la musique médiévale autour de la Méditerranée**, *in L'iconographie musicale dans les pays méditerranéens à l'heure de la numérisation intensive des images*(Dir. F. Billiet, H. Makhlouf), Tunis, Centre des Musiques Arabes et Méditerranéennes, 2020, p. 31-44.

[3] F. Billiet, **La modernité des scènes musicales dans les stalles de Gaillon au début du xvie siècle**, *Musiques-Images-Instruments, Mélanges en l'honneur de Florence Gétreau*, F. Guilloux, C. Massip, A. Framboisier, Y. Balmer (eds.), XIX+645 p., 2019 - ISBN: 978-2-503-58371-6

Publications of Victoria EYHARABIDE related to the project

[4] Imad Eddine Ibrahim BEKKOUCH, Nicolae Dragoș CONSTANTIN, Victoria EYHARABIDE, Frédéric BILLIET (2021). **Adversarial Domain Adaptation for Medieval Instrument Recognition**. Intelligent Systems Conference (IntelliSys) 2021. Amsterdam, The Netherlands. September 2021. https://saiconference.com/IntelliSys

[5] Imad Eddine Ibrahim BEKKOUCH, Victoria EYHARABIDE, Frédéric BILLET (2021). **Dual Training for Transfer Learning: Application on Medieval Studies**. International joint conference on neural networks (IJCNN) 2021. July, 2021**.** https://www.ijcnn.org/

[6] Victoria Eyharabide, Xavier Fresquet, Fouad Aouinti, Marcus Liwicki, Susan Boynton (2020). **Towards the identification of Medieval Musical Performance using Convolutional Neural Networks and IIIF**. International IIIF Annual Conference, Harvard University and the Massachusetts Institute of Technology (MIT), USA. June 2020.

[7] Victoria Eyharabide, Vincent Lully, Florentin Morel (2019). **MusicKG: Representations of sound and music in the Middle Ages as Linked Open Data**. Proc. of the leading European conference on Semantic Technologies and AI (SEMANTiCS 2019), Karlsruhe, Germany. September 2019 https://doi.org/10.1007/978-3-030-33220-4_5

[8] J. Kusnick, et al. (2020), A Timeline Metaphor for Analyzing the Relationships between Musical Instruments and Musical Pieces. VISIGRAPP: 240-251. https://doi.org/10.5220/0008990502400251

[9] S. Mukherjee, et al. (2020), Integration of Domain-Specific Metadata Schema for Cultural Heritage Resources to DSpace: A Prototype Design, Journal of Library Metadata. https://doi.org/10.1080/19386389.2020.1834093

[10] M. Cochez et al. (2017), *Global RDF vector space embeddings,* The Semantic Web – ISWC 2017:190-207. https://doi.org/10.1007/978-3-319-68288-4_12

[11] P. Ristoski, et al. (2016), *RDF2Vec: RDF Graph Embeddings for Data Mining,* The Semantic Web – ISWC 2016:498-514. https://doi.org/10.1007/978-3-319-46523-4_30