

DMP First draft April 2022

Project Name On the Art of Counterpoint - DMP First draft April 2022

Project Identifier 11A9922N

Grant Title 11A9922N

Principal Investigator / Researcher Vicente Parrilla Lopez

Project Data Contact +34 606678070, vicente.parrilla@kuleuven.be

Description In the Renaissance, musicians moved from reading to improvising music with total confidence and their compositional craft was based on their contrapuntal improvisations. Improvised counterpoint is understood here as the technique to add unwritten parts to a preexisting work or melody. However, the ancient art of improvised counterpoint is lost, and people can no longer experience it. This research project intends to revive it in all its complexity, as both academic knowledge and through sounds, with the key objective of presenting the results artistically in the form of concerts and recordings open to the public. The fundamental research hypothesis is that, if I achieve a clear understanding and deep assimilation of the Renaissance contrapuntal style and techniques, then I will be able to successfully present the results both to the academic world and to modern concert audiences. By becoming a fluent improviser in Renaissance style, I will provide a more accurate picture of how performers produced music at the time and, by extension, of the sixteenth-century musical soundscape.

Institution KU Leuven

1. General Information

Name applicant

Vicente Parrilla López

FWO Project Number & Title

FWO Project Number: 11A9922N

Title: On the Art of Counterpoint.

Affiliation

- KU Leuven

Doctorate in the Arts with two other affiliations: LUCA School of Arts and docARTES (Orpheus Instituut, Ghent)

2. Data description

Will you generate/collect new data and/or make use of existing data?

- Generate new data

Describe in detail the origin, type and format of the data (per dataset) and its (estimated) volume. This may be easiest in a table (see example) or as a data flow and per WP or objective of the project. If you reuse existing data, specify the source of these data. Distinguish data types (the kind of content) from data formats (the technical format).

My corpus of data comprises several 16th-century sources containing musical examples and explanations about contrapuntal techniques.

1. I will transcribe a selection of those examples into modern notation using FinaleMusic software (.musx and .musicxml files).
2. I will identify the particular interval combinations and techniques employed, which I will manually label into the music examples, producing .pdf files
3. Then I will record them using audio equipment and Logic Pro software.
4. Later on, I will listen to and play along with those examples for a prolonged period with the aim of absorbing the data.
5. Finally, I will produce audio and video recordings (in .flac .mp3 and .mp4 formats) of my artistic output which I will then analyse and compare to the Renaissance sources.

This methodology will allow me to produce new counterpoint lines that strictly follow the musical style codified by the aforementioned Renaissance sources, thus reviving the lost art of 16th

century improvised counterpoint.

Type of data	Format	Volume	How created
Manuscript treatise Paris, BnF Esp. 219	.pdf	27 MB	Open access file (public domain) on Gallica: https://gallica.bnf.fr/ark:/12148/btv1b10033631k/f3.image
Music examples	.musx	30 MB	Manually-entered data into FinaleMusic software
Music examples	.xml	300 MB	Exported files from the previous .musx files
Music examples	.pdf	300 MB	Exported files from the previous .musx files
Music examples	.png	100 MB	Exported files from the previous .pdf files
Music examples	.midi	50 MB	Exported files from the previous .musx files
Preexisting audio recordings	.mp3	200 MB	Downloaded from the open-access website: http://josquin.cesr.univ-tours.fr/lusitano/collections/show/20
Audio recordings - experimental, observational	.flac	5-10 GB	Home audio recordings using Logic Pro X audio software
Video recordings - experimental, observational	.mp4	1TB	Home video recordings using iPhone SE's video camera
Filenames table in spreadsheet	.numbers	1 MB	Manually-entered data into Numbers software
Filenames table in spreadsheet	.csv	1 MB	Exported files from the previous .numbers files
Text notes, Textual data	.txt	10 MB	Manually-entered data into iA Writer software
Text notes, Textual data	.txt	10 MB	Extracted as plain text files from annotated PDF files via PDF Expert software
Project website	.txt, .png, .mp3, .midi	600 MB	Content is stored in plain text files within folders (no database needed)

Of the file formats described above, only .musx (FinaleMusic) and .numbers are proprietary. To ensure that my data will remain accessible and reusable, I will export all data in .musx format to MusicXML (.musicxml) after the end of the project. MusicXML is the standard open format for exchanging digital sheet music, its format is open, fully documented, and can be freely used under the W3C Community Final Specification Agreement. Similarly, I will export all data in .numbers format to .csv after the end of the project.

In addition, I have created a project website for increasing the visibility and impact of my research work using Kirby CMS (<https://getkirby.com>). Kirby stores content in simple text files and folders, to which audiovisual files can be added and which work as site pages. It is a lightweight and highly performant CMS, easy to maintain and to backup and whose complete code base is available on GitHub. Kirby is designed with privacy in mind and is GDPR (DSGVO) compliant by default and websites are accessible on phones, tablets, and desktops.

3. Legal and ethical issues

Will you use personal data? If so, shortly describe the kind of personal data you will use. Add the reference to your file in KU Leuven's Register of Data Processing for Research and Public Service Purposes (PRET application). Be aware that registering the fact that you process personal data is a legal obligation.

- No

Privacy Registry Reference: -

Short description of the kind of personal data that will be used: -

I have used the research ethics flowchart provided by KU Leuven

(<https://www.kuleuven.be/english/research/ethics>) to verify that there are no potential ethical issues associated with my data.

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, add the reference to the formal approval by the relevant ethical review committee(s)

- No

Does your work possibly result in research data with potential for tech transfer and valorisation? Will IP restrictions be claimed for the data you created? If so, for what data and which restrictions will be asserted?

- No

Do existing 3rd party agreements restrict dissemination or exploitation of the data you (re)use? If so, to what data do they relate and what restrictions are in place?

- No

To avoid any uncertainty about who holds the copyright in the data used in the project, I have taken measures to ensure that data not created by me can always be distinguished from data created by me, and this throughout the research process.

4. Documentation and metadata

What documentation will be provided to enable reuse of the data collected/generated in this project?

All the information users would need to understand and reuse the data will be discussed and documented as part of the dissertation, which will contain detailed information on the methodology used to collect the data. In addition, I will add the following documentation to my data:

- Project-level documentation: a README.txt file contained in [location to be determined], based on the template README.txt provided by research support staff at KU Leuven and adapted to the needs of my project and my discipline.
- File-level documentation: additional README.txt files contained in [location to be determined]

I will describe how audiovisual raw data will be processed into other forms of data and, when possible, metadata will be embedded in the software used to produce the files.

Finally, providing I get the necessary training and support, I plan to use Git (a free and open source resource), via GitLab KU Leuven, as a distributed version control system hosted on github.com, both to document data processing and list changes and as an extra backup.

Will a metadata standard be used? If so, describe in detail which standard will be used. If no, state in detail which metadata will be created to make the data easy/easier to find and reuse.

I will add metadata to my data as follows:

- I will use Zotero (an open-source reference management software) for organizing bibliographic references and retrieving metadata in the required formats. Zotero allows to export reference data to open-source formats like bibtex, improving the accessibility of data.
- I will add metadata to the website files as I will be producing them by collecting standardized metadata on a spreadsheet in .numbers format. Since this format is proprietary, I will later export all data to Open Document Format for Office
- I will add metadata to image files that I will collect on a spreadsheet in .numbers format,

later exported to OpenDocument.

- I will embed metadata to sound files using Piezo software (<https://rogueamoeba.com/piezo/>) that I will also collect on a spreadsheet in .numbers format, later exported to OpenDocument.
- I will use a controlled vocabulary and file naming conventions for variables.
- I will use a uniform naming convention for files and folders.
- I plan to use unique identifiers for my data, to be collected on a spreadsheet for easy reference.

Newly generated images will include the following metadata:

- Media Type: image/png
- Size: 1234 KB
- Dimensions 1324x1234 Pixel
- In the case of uploaded, online images, Title, Caption and Alt text fields will be added as well.

When sharing my data, I will include rich metadata by using web standards (HTML and CSS) and the appropriate machine-readable metadata embedded in image, audio, and video files.

Finally, I will seek the advice of colleagues in my discipline and research support staff at my institution to decide which metadata standard (if any) is appropriate for me.

5. Data storage and backup during the FWO project

Where will the data be stored?

All my research data files (musx, pdf, png, midi, mp3, flac, mp4, and numbers) are thus stored through a combination of physical drives (two computers and several external hard drives) and the cloud (Dropbox and Backblaze). In my current configuration, text files (.txt), which I use for notes and for drafting all types of texts are stored similarly, except that I sync my main text editor (iA Writer) to Apple's iCloud instead of Dropbox. It is done automatically, and iCloud files are also stored in Backblaze and in the same physical drives. iCloud allows to view and restore past versions of documents. My text editor automatically saves versions of documents as I work on them. At any time, I can browse through document versions and go back to an older version. I can also explicitly save a version. A version is saved automatically every hour, or more frequently when I'm making many changes. A version is also saved when I open, save, duplicate, lock, rename or revert a document.

This storage solution is appropriate for my purposes because it allows me to keep multiple copies of my data (both locally and in the cloud) in an automatic way.

How is backup of the data provided?

During the research, my data will be stored using the following storage solutions, allowing me to keep up to five copies of my data:

1. Copy 1: on my two computers. Two local hard drives from a desktop and a portable computers
2. Copy 2: on a first cloud service (Dropbox). The two computers are synchronised via cloud storage (Dropbox, Plus account, with 2TB of storage capacity and version history allowing to revert files to any previous version saved to Dropbox in the past 30 days).
3. Copy 3: on a second cloud service (Backblaze). My files are backed up automatically and securely in the cloud via an additional backup service, Backblaze (<https://www.backblaze.com> - Personal Backup plan with extended 1-year version history). Backblaze has no file size or bandwidth speed limits and unlimited cloud backup. It constantly scans for new or changed files and automatically backs them up. No need to pick files or folders, we backup everything automatically. All versions of my files that are backed up —whether I have updated, changed, or fully deleted them from my computer— will remain in my Backblaze backup for 1 year after being modified or deleted from my device. This service backs up all user-created data and covers unlimited backup of data on the computer itself as well as any internal drives, and external drives physically connected to the computer. Extended 1-year version history allows me to retain deleted data or disconnected drive data for a year. Personal Backup is an active, evolving mirror of the current state of my computer and drives.
4. Copy 4: on a first external storage (disk 1, encrypted). An USB hard disk connected to my desktop computer. It allows me to combine local and cloud storage with automatic back-up procedures using "snapshot" technology via Time Machine, a Mac OS system tool. Time

Machine automatically makes hourly backups for the past 24 hours; daily backups for the past month; and weekly backups for all previous months. The oldest backups are deleted when the backup disk is full. Time Machine backs up only the files that changed since the previous backup.

5. Copy 5: on a second external storage (disk 2, encrypted). An USB hard disk connected to my portable (notebook) computer and following the same procedure as before via Time Machine.

Is there currently sufficient storage & backup capacity during the project? If yes, specify concisely. If no or insufficient storage or backup capacities are available then explain how this will be taken care of.

- Yes

What are the expected costs for data storage and back up during the project? How will these costs be covered?

The costs are as follows:

1. Backblaze Yearly Computer Unlimited Backup (\$70.00)
2. Backblaze Extended Version History (\$24.00)
3. Dropbox Plus account (2TB, €119.88 per year)
4. Toshiba external hard drive (USB 3.0, 2.5") connected to my desktop computer for automatic Time Machine encrypted backups (1 TB, €47,89)
5. Seagate Portable Drive external hard drive (USB 3.0) connected to my notebook computer for automatic Time Machine encrypted backups (5 TB, €113,99)

I had a Backblaze subscription on my own initiative before getting FWO funding, so I have payed myself this online backup service until now. I have recently renewed my subscription on 02/19/2022 for one year (until 02/18/2023).

The same goes for my Dropbox Plus account (2TB, €119.88 per year) which I have payed myself for the period 29/8/2021 to 29/8/2022.

When it comes to iCloud, all Apple device owners get 5GB of iCloud storage for free. Since the only research data I store on iCloud are my text files (txt), the free subscription is enough for now.

I acquired the afore-mentioned external hard drives on my own initiative. The Toshiba drive was bought on February 4 2014 while the Seagate drive was bought on August 3 2020.

I plan to acquire at least two extra external hard drives (4 or 5 TB each) for backing up the research data I will produce, as audiovisual files take up a considerable amount of hard drive space. I may consider getting a third new external drive to replace the 2014 Toshiba drive as well.

That would mean an extra investment of approximately 350 euros. To cover these costs, I may use part of the allocated FWO project budget.

Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

I use FileVault (<https://support.apple.com/en-gb/HT204837>) to encrypt the startup disk on my two Mac computers. FileVault full-disk encryption (FileVault 2) uses XTS-AES-128 encryption with a 256-bit key to help prevent unauthorized access to the information on my startup disks.

I use two different services to locate my computer in case of lost.

Backblaze can help recover my lost computer or stolen laptop by tracking it, if and when it connects to a network again. It provides the following data:

- Last Time Online: 03/01/2022 07:57 PM
- Last time location reported: 03/01/2022 05:02 PM
- IP Address (Network Address): 91.183.2.20
- A "Map This" function.
- A "Files backed up in the last 24 hours" function, which may help to see which files the thief has uploaded in recovering my computer.
- It displays the ISP my computer is using, which helps police determine location.
- ISP contacts: ISPs know the physical address of their IP addresses.

The process works as follows:

1. First, I prepare a restore of my data. Backblaze's "View / Restore" files feature allows me to download files or have them FedEx me a hard drive with the encrypted data. I even have the option to "roll back time" to the day before a thief may have corrupted or removed any of my files. It allows me to roll back time for up to 30 days.
2. Second, I will make use of Backblaze's recovery data: it would allow me to file a police report with useful recovery data if I were to contact the police in case my computer was stolen. I could include the IP address mentioned above in the police report. The police will use the IP address to ask the ISP for the physical address of the computer. This is a useful feature given that, typically, neither the ISP nor the police will provide me the address.
3. Third, I would periodically check back Backblaze's status page to see if my stolen computer has connected to the internet. With their "Locate My Computer" feature turned on, if the computer is connected to the Internet, I'll be able to get a location to give to the police.
4. Finally, in case I could not recover my computer and I would end up getting a new one, I could inherit my previous computer's backup state via Backblaze.

Besides Backblaze, I use Apple's "Find My" app (<https://www.apple.com/icloud/find-my/>), which makes it easy to keep track of my Apple devices even if a missing device would be offline. It allows me to see all my devices on a real-time map, to play a sound to find my devices and to display a message for someone who finds them. It also allows me to lock my device down automatically via the "Activation Lock" service, which is designed to prevent anyone else from using or selling my device. When I enable "Find My" on my device, Activation Lock is turned on automatically. My Apple ID and password will then be required in order to reactivate my device. As an extra security tool, I use 1Password (<https://1password.com>) to store and use strong passwords (up to 100 characters long) and two-factor authentication codes. With end-to-end encrypted data via AES 256-bit encryption, it is a safe way to manage my login info. The information stored in 1Password is encrypted, and only I hold the keys to decrypt it. 1Password can't see my 1Password data, so they can't use it, share it, or sell it. The Time Machine backups on both external hard drives are encrypted and automatic.

I do not plan to make use of Personal data or other sensitive data during my research.

6. Data preservation after the FWO project

Which data will be retained for the expected 5 year period after the end of the project? In case only a selection of the data can/will be preserved, clearly state the reasons for this (legal or contractual restrictions, physical preservation issues, ...).

In accordance with FWO regulations, I will retain all my research data for 5 years.

In addition, given that KU Leuven research data management policy expects that relevant research data generated are retained for a period of minimally 10 years after the end of the project in a safe, secure & sustainable way for purposes of reproducibility, verification, and potential reuse, I will ensure that all relevant data is preserved by selecting data on the basis of what can't be recreated and what is potentially useful to others. Besides, I will make sure to preserve them in non-proprietary formats, choosing those in common usage by the research community that adhere to an open, documented standard and in an unencrypted, uncompressed format.

Apart from the main source I plan to work on, a manuscript treatise scanned and preserved by Gallica (<https://gallica.bnf.fr/ark:/12148/btv1b10033631k/f3.image>, open access PDF file, public domain, 27 MB), I plan to preserve the following datasets (described in section 2) after the project:

- Music examples, musx, 30 MB, Manually-entered data into FinaleMusic software, Exported as musicxml files.
- Music examples, pdf, 300 MB, Exported files from the previous musx files.
- Music examples, png, 100 MB, Exported files from the previous pdf files.
- Music examples, midi, 100 MB, Exported files from the previous musx files.
- Preexisting audio recordings, mp3, 75 MB, Downloaded from the open-access website: <http://josquin.cesr.univ-tours.fr/lusitano/collections/show/20>
- Audio recordings, flac (experimental, observational), 5-10 GB, Home audio recordings using Logic Pro audio software.
- Video recordings, mp4 (experimental, observational), 1TB, Home video recordings using iPhone SE's video camera.
- Filenames table in spreadsheet, numbers, 1 MB, Manually-entered data into Numbers software.

- Text notes, .txt (Textual data), 10 MB, Manually-entered data into iA Writer software.
- Text notes, .txt (Textual data), 10 MB, Extracted as plain text files from annotated PDF files via PDF Expert software.
- Project website, .txt, .png, .mp3, midi, 600 MB, Content is stored in plain text files within folders (no database needed).

I will safely dispose of other data by deleting or erasing it when the right time comes.

Where will the data be archived (= stored for the longer term)?

After the end of the project, I will store my research data (including metadata and all documentation necessary to reuse the data) on a new data repository option at KU Leuven available since January 2022: RDR (<https://www.kuleuven.be/rdm/en/rdr>). RDR is a free service that allows to archive and/or share data with great data control and description capabilities that may serve as a platform for long-term preservation of data. In addition, RDR allows to store many kinds of data and to choose which specific files within a dataset are accessible to others, while preventing access to the remaining files.

In the event I use personal data, I will ensure that personal information is properly protected by clearly indicating which files contain personal information in the course of the project and anonymizing all relevant textual or audio-visual data files by using pseudonyms or generic descriptors and editing identifying information, rather than blanking-out information. Besides, I will follow the suggestions offered by the UK Data Service on <https://ukdataservice.ac.uk/learning-hub/research-data-management/#anonymisation>

I will seek the advice of legal support staff at my institution to properly anonymize relevant files and choose the correct storage solution.

What are the expected costs for data preservation during the retention period of 5 years? How will the costs be covered?

The audiovisual files that will be produced and compiled during the research project will be hosted on the servers of KU Leuven. In view of the expected total size of the files (< 5 TB), I will store them on KU Leuven OneDrive for Business at no cost: although the standard every user gets is 2 TB, this capacity can be extended up to 5 TB without costs, according to KU Leuven ICTS Storage guide (<https://icts.kuleuven.be/sc/english/storage/storageguide>)

In the event of extra costs at a later stage, the storage costs will be paid by using part of the allocated FWO project budget.

7. Data sharing and reuse

Are there any factors restricting or preventing the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)?

- No

The main source I plan to work on which is not produced by me is a public domain, open access PDF file, preserved by Gallica (<https://gallica.bnf.fr/ark:/12148/btv1b10033631k/f3.image>, 27 MB). Given it is shared as public domain, there are no factors restricting or preventing its use and sharing.

Which data will be made available after the end of the project?

I will not share the following dataset because the files were obtained from a copyrighted website (<http://josquin.cesr.univ-tours.fr/lusitano/collections/show/20>), unless I get written permission from the creator:

- Preexisting audio recordings, mp3, 75 MB, downloaded from the open-access website: <http://josquin.cesr.univ-tours.fr/lusitano/collections/show/20>

In addition, I will not share the following datasets because of its lack of relevance, unless advised by research support staff at my institution that this data can be shared if appropriate conditions for access are met:

- Filenames table in spreadsheet, numbers, 1 MB, Manually-entered data into Numbers software.
- Text notes, txt (Textual data), 10 MB, Manually-entered data into iA Writer software.
- Text notes, txt (Textual data), 10 MB, Extracted as plain text files from annotated PDF files via PDF Expert software.

However, I will store them in a data repository with the appropriate data descriptions. By allowing free access to metadata (instead of files) within the data repository, I can inform about the existence of these data and make it available upon request in case someone finds it useful. By storing it on RDR instead of on Zenodo, I can create a dataset with a mix of open and closed files.

The remaining datasets will be uploaded in a cvs format in Zenodo under a CC-BY license:

- Music examples, musx, 30 MB, Manually-entered data into FinaleMusic software, Exported as musicxml files.
- Music examples, pdf, 300 MB, Exported files from the previous musx files.
- Music examples, png, 100 MB, Exported files from the previous pdf files.
- Music examples, midi, 100 MB, Exported files from the previous musx files.
- Audio recordings, flac (experimental, observational), 5-10 GB, Home audio recordings using Logic Pro audio software.
- Video recordings, mp4 (experimental, observational), 1TB, Home video recordings using iPhone SE's video camera.
- Project website, txt, png, mp3, midi, 600 MB, Content is stored in plain text files within folders (no database needed).

Where/how will the data be made available for reuse?

- In an Open Access repository

Selected datasets with documentation will be uploaded in open file, standards formats in Zenodo (<https://zenodo.org/>).

I will apply the FAIR principles as thoroughly as possible when sharing my data:

- I will ensure the data is findable by sharing it on Zenodo, a solution that gives the data a DOI.
- I will ensure the data is accessible by clearly indicating how others can get access to the data.
- I will ensure the data is interoperable by using open file formats and standards whenever possible.
- I will ensure the data is reusable by connecting it with an appropriate license, documentation, and metadata.

When will the data be made available?

- Immediately after the end of the project

My intention is to make the data available immediately after the end of the project, unless advised by research support staff at my institution in case I need to publish research findings.

Who will be able to access the data and under what conditions?

The selected datasets with documentation will be uploaded in open file, standard formats to Zenodo (<https://zenodo.org/>) as an open access dataset under a CC-BY license. Therefore, it will be available to anyone for any purpose, provided that they give appropriate credit to the creators.

What are the expected costs for data sharing? How will the costs be covered?

Since Zenodo is a free service, no costs are expected for the datasets shared through this data repository.

On the other hand, since the datasets hosted on the servers of KU Leuven are expected to be smaller than 5 TB in total, I will store them on KU Leuven OneDrive for Business at no extra cost (although the standard every user gets is 2 TB, this capacity can be extended up to 5 TB without costs, according to KU Leuven ICTS Storage guide (<https://icts.kuleuven.be/sc/english/storage/storageguide>))

In the event of extra costs at a later stage, the storage costs will be paid by using part of the allocated FWO project budget.

8. Responsibilities

Who will be responsible for data documentation & metadata?

In the course of the research, the principal investigator (Vicente Parrilla,

vicente.parrilla@kuleuven.be) is responsible for implementing research data management. The supervisor, (David Burn, david.burn@kuleuven.be), is responsible for ensuring that the principal investigator implements research data management.

Who will be responsible for data storage & back up during the project?

The person carrying out the research day to day who will be implementing the DMP and keeping it up to date throughout the project is Vicente Parrilla (vicente.parrilla@kuleuven.be).

Who will be responsible for ensuring data preservation and reuse ?

After the research project ends, preservation of data is the responsibility of the supervisor. The supervisor needs the permission of the principal investigator for any reuse of the data that exceeds the scope of the license attached to the data.

Who bears the end responsibility for updating & implementing this DMP?

The PI (Vicente Parrilla, vicente.parrilla@kuleuven.be) bears the end responsibility of updating & implementing this DMP. However, his supervisor will likely be responsible for overall data management, especially in the long term [to be discussed].