Institut de Recherche et Coordination Acoustique / Musique (IRCAM)
Musical representations team

Overview of current research at IRCAM
with implications in orchestration research

Philippe Esling[1]
[1]Maître de conférences, représentations musicales
Université Pierre et Marie Curie (UPMC – Paris 6)
esling@ircam.fr

Overview of IRCAM

IRCAM – Research and creation

Founded in 1977 by Pierre Boulez
Associated with the Pompidou Center
Funded by the French Minstry of Culture
Unique institution for music tech. research.

Gathers **musicians, scientists and engineers** for

- renewing contemporary music expression through science & technology
- multi-disciplinary research applied to sound & music

Four main departments (~180 people)

- **STMS Sc. & Tech research Lab** : 100 persons including researchers,engineers, techs, PhD candidates : acoustics, signal processing, computer science, psychology, musicology
- **Creation.** 30+ works / year using the latest research technologies.
- **Higher education.** *scientific* and *artistic* courses, both hosted inside and with institutional partners
- **Research/Creation interface department**

IRCAM – R&D Department

R&D Department – By the numbers

Hosts the IRCAM-CNRS-UPMC (joint UMR 9919 STMS)

- 140 persons/ year : 100 researchers/engineers/PhD and 40 interns/guests
- Scientific topics : digital audio signal processing, computer science, acoustics, human perception/ cognition, musicology
- 20+ softwares environments distributed – forumnet.ircam.fr
- Very active in R&D collaborative projects, 20 ongoing, 33% as Coordinator
- Technology licenses : Several dozens ongoing

Research teams

- **EAC** - Acoustics and Cognitive Spaces
- **PDS** - Perception and Sound Design
- **AnaSyn** - Analysis/Synthesis
- **S3AM** - Physical modeling and synthesis
- **RepMus** - Musical representations and learning
- **APM** - Musicology

IRCAM – Inside ACTOR

S3AM

RepMus

AnaSyn

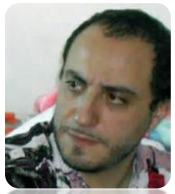
APM

PDS

EAC

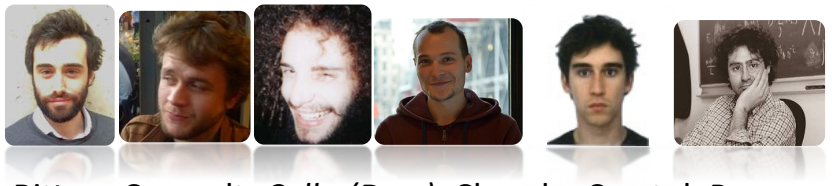
3 permanent researchers


Carlos Agon


Gerard Assayag


Philippe Esling

5 PhDs and 1 developer

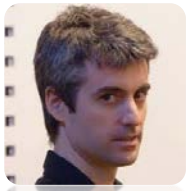

Bitton, Carsault, *Cella (Dev.)*, Chemla, Crestel, Prang

Large research group of composers


Y. Maresz


D. Ghisi


K. Haddad


C. Castellarnau

(and whole GdR Orchestration = ~30 composers)

IRCAM – Inside ACTOR

S3AM

RepMus

AnaSyn

APM

PDS

EAC

Patrick Susini

Olivier Warusfel
Markus Noisternig

Nicolas Donin

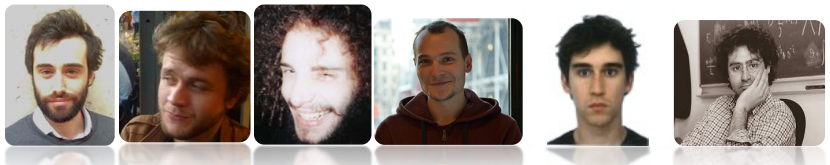
3 permanent researchers

Carlos Agon

Gerard Assayag

Philippe Esling

5 PhDs and 1 developer



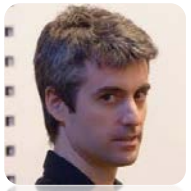
Bitton, Carsault, *Cella (Dev.)*, Chemla, Crestel, Prang

Large research group of composers

Y. Maresz

D. Ghisi

K. Haddad



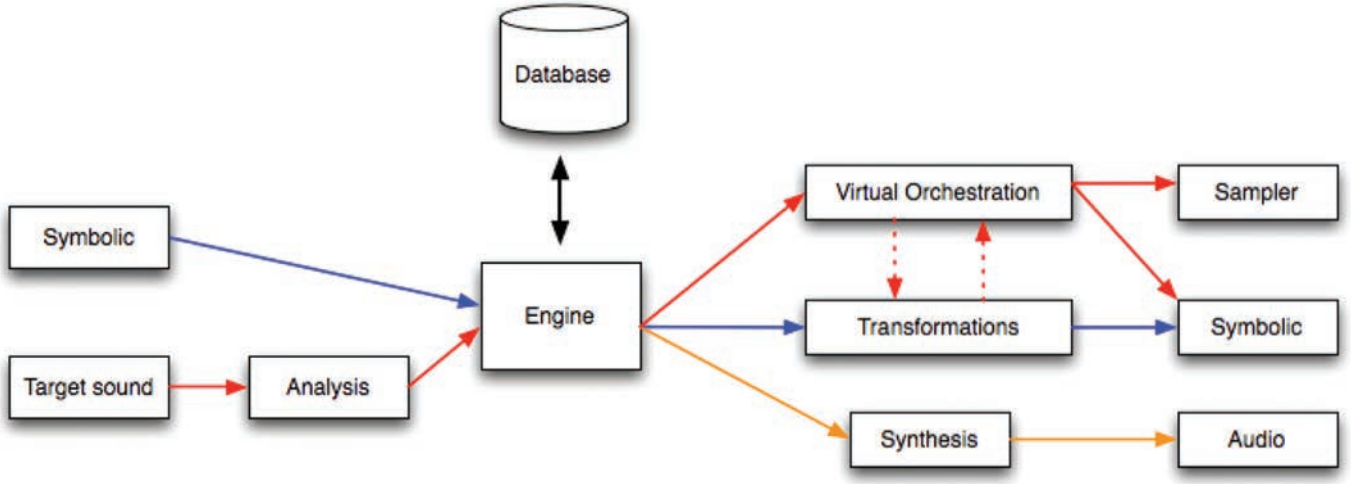
C. Castellarnau

(and whole GdR Orchestration = ~30 composers)

Computer orchestration at IRCAM

A historical project

- Started from composers discussion ~2003, first proposal by Yan Maresz



- Sample Orchestrator (ANR) project from 2008
- Lead to 3 PhD thesis (Gregoire Carpentier, Damien Tardieu, Philippe Esling)
- Several tools developed for computer assisted orchestration
- Now a full-time associate professor devoted to the topic
- Might lead to the creation of a dedicated team (ACIDITEAM)
- Orchestration **has all the questions !**

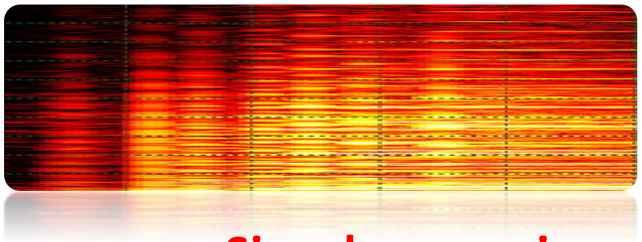
The spaces of orchestration

Acoustics

Pressure wave (acoustic raw signal)

$$\mathbf{s} = \{s_t, s_{t+\tau} \cdots, s_{t+N\tau}\}$$

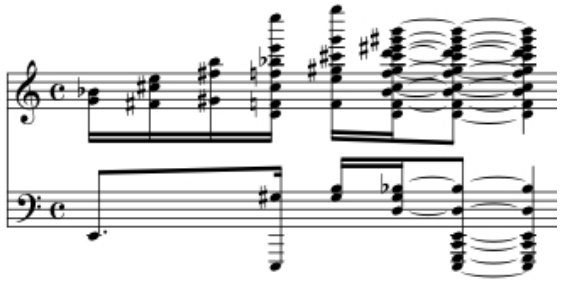
$$s(t) = \sum_{k=0}^{K-1} \alpha_k z_k^t + b(t)$$



Signal processing

Spectral transform (Fourier, wavelet)

$$f(\omega) = \int_{-\infty}^{+\infty} s(t) e^{-i2\pi\omega t} dt$$



Computer music

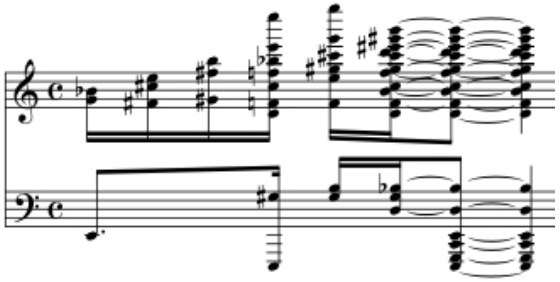
Symbolic score (computer music)

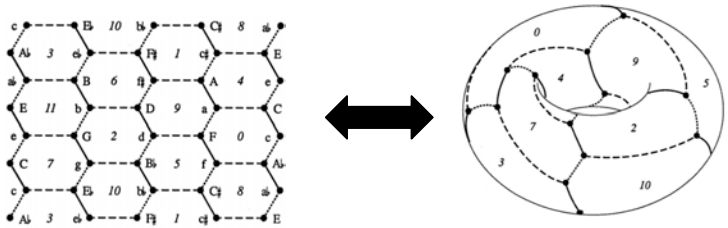
$$M = \{n_i \mid i \in [1, N], n_i \in G\}$$

$$G = \mathbb{Z}/n\mathbb{Z}$$

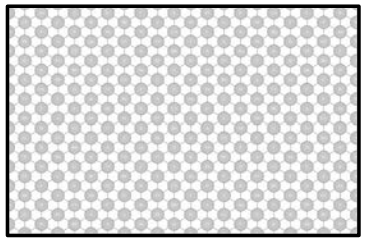
$$< T_k \mid (T_k)^{12} = T_0 = 0 >$$

Music as symbols

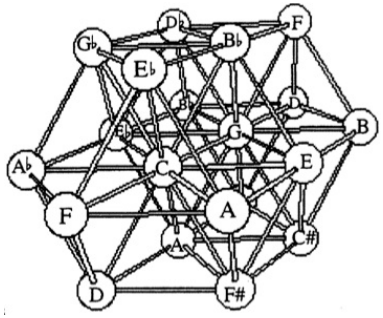




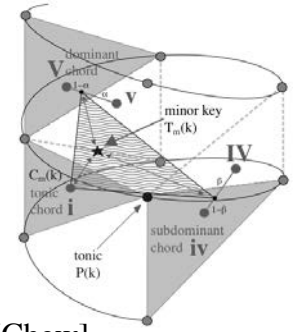
Chicken Wire Torus [Douthett & Steinbach]



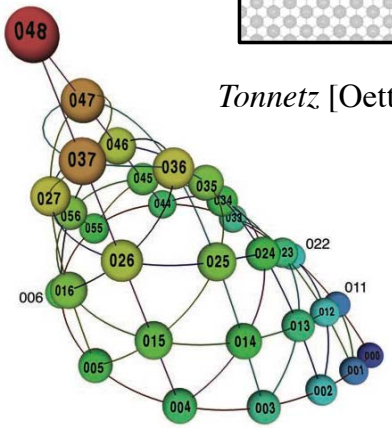
Tonnetz [Oettingen, Riemann]



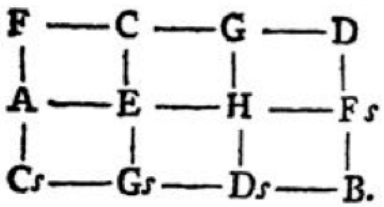
3D *Tonnetz* [Gollin]



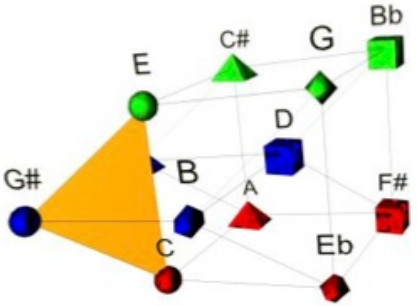
Spiral Array [Chew]



Orbifolds [Tymoczko]



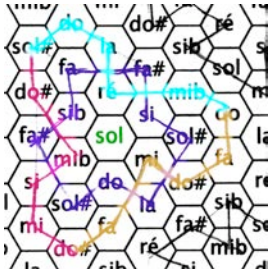
Speculum Musicum [Euler]

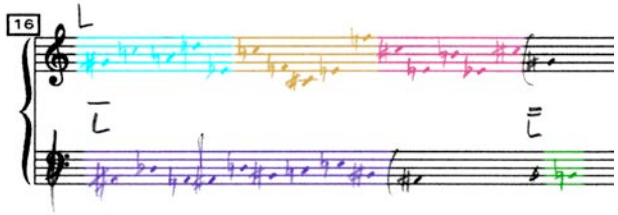


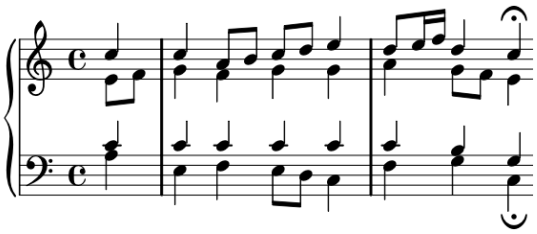
Model Planet [Barouin]

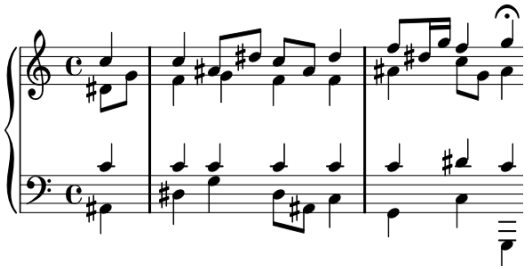
Representing symbolic music

- Multiple representations of music exists
- Work on spatial structures (Tonnetz) by Louis Bigo



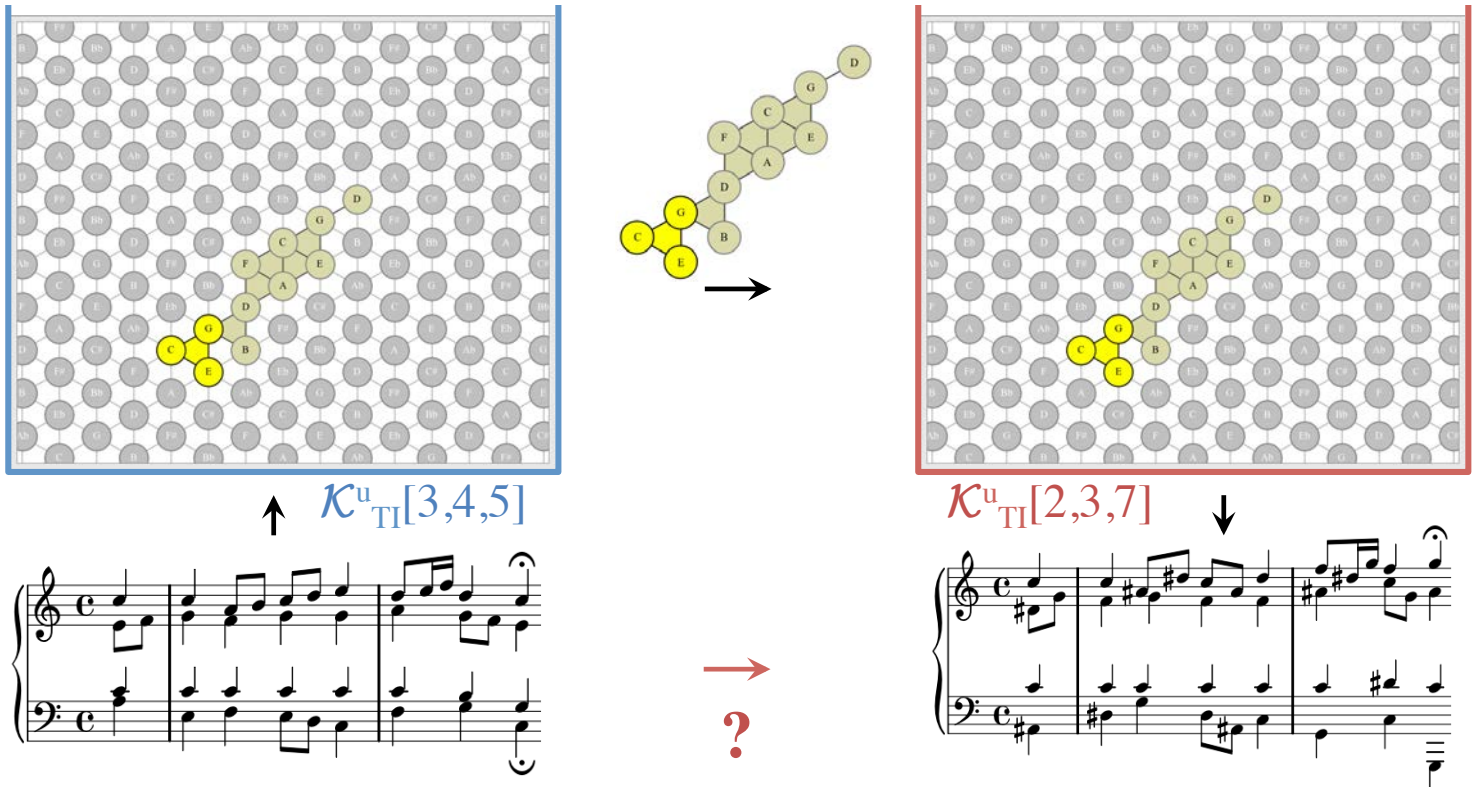






Representing symbolic music

- Multiple representations of music exists
- Work on spatial structures (Tonnetz) by Louis Bigo



Musical embedding space

Matthieu Prang, Léopold Crestel (prang@ircam.fr)

GloVe

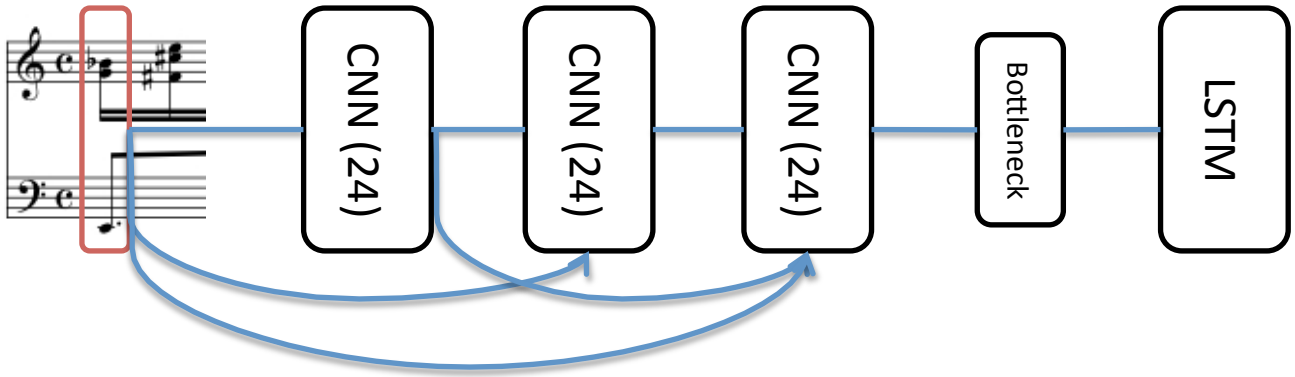
State-of-the-art word embeddings algorithm that encode the co-occurrence probabilities of two words.

$$J = \sum_{i,j=1}^{V} f(X_{ij})(w_i^\top \hat{w}_j + b_i + \hat{b}_j - logX_{ij})^2$$

CNN-LSTM

Specially tailored for musical symbolic datas : CNN for pitch-class invariant and LSTM for time series sequence

$$J = min(\frac{1}{N}\sum_{i=1}^{N}(v_{w_i} - v_{w_i'})^2)$$



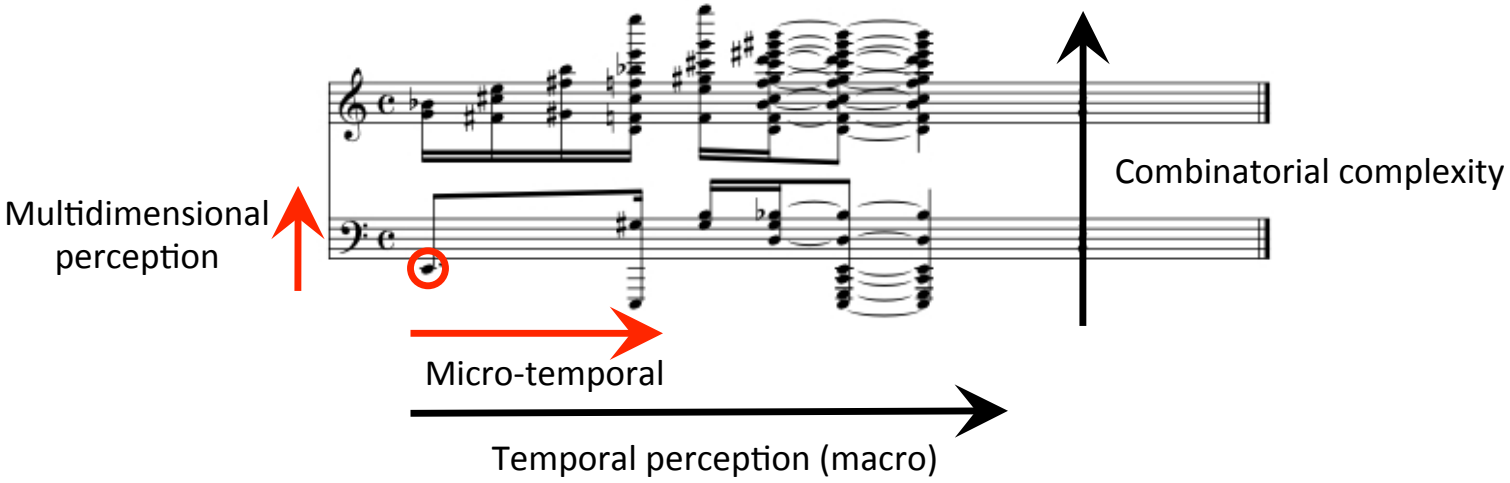
Musical embedding space

Matthieu Prang, Léopold Crestel (prang@ircam.fr)

- Machine learning requires loads of examples
- Defining a very large classical MIDI database
  - Collected from 10 independent sources
  - More than 15,000 composers
  - Above 90,000 MIDI files
  - Spanning wide eras
- Still needs a **checking / safety procedure**
- Already re-defined state-of-the-art in inference
- Defined first state-of-the-art in orchestral prediction
- Currently developing composition from these spaces
- Also working on interactive composition

Neural orchestration

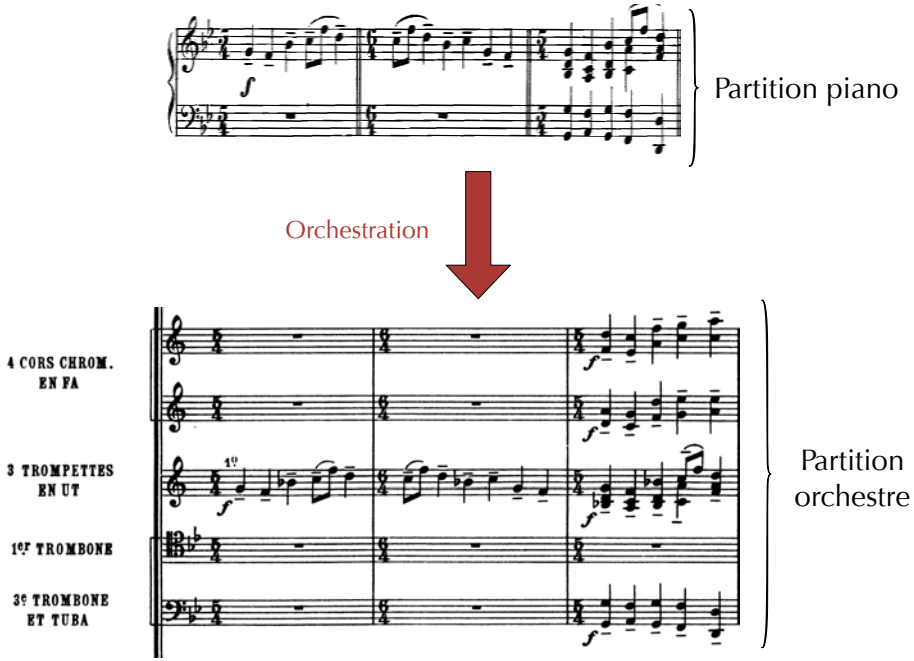
- Art of **writing** musical pieces for orchestras (symbolic view)
- Can be seen as the art of **mixing instrumental properties**. (signal view)
- Discovering how the orchestra is used to achieve a musical thought.
- At the crossroads between **signal and symbolism** (writing and timbre)
- How to combine **instrumental models** (spectral properties)

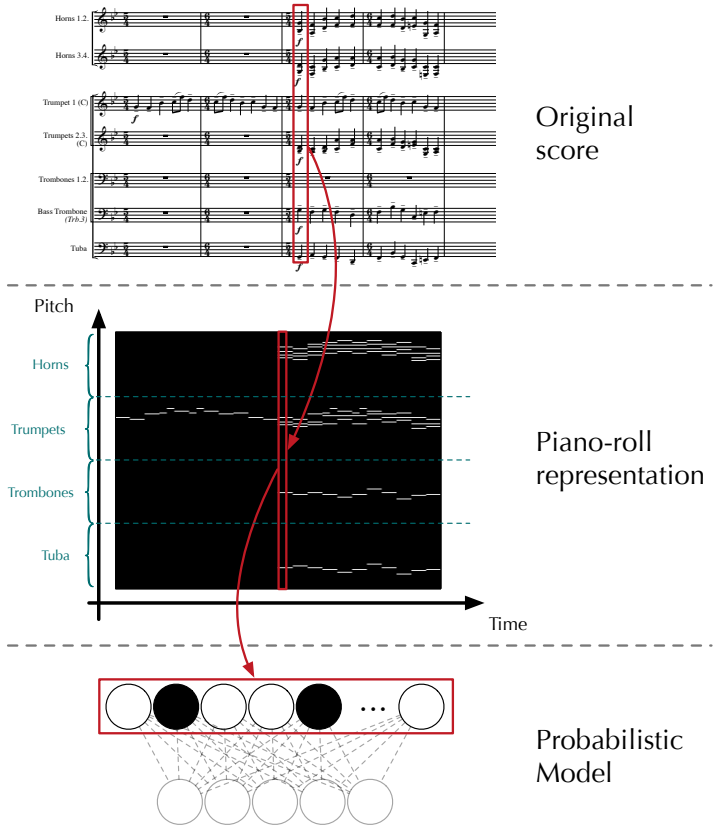


Neural orchestration

Léopold Crestel
(crestel@ircam.fr)

Oppositely to the « historical » computer-aided orchestration systems …





Neural orchestration

Léopold Crestel
(crestel@ircam.fr)

Dense-CNN

Partition piano

Orchestration

4 CORS CHROM. EN FA

3 TROMPETTES EN UT

1er TROMBONE

3e TROMBONE ET TUBA

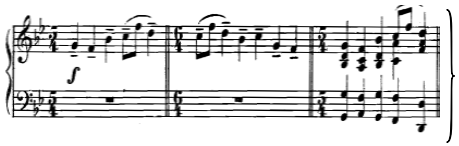
Partition orchestre

RNN-LSTM

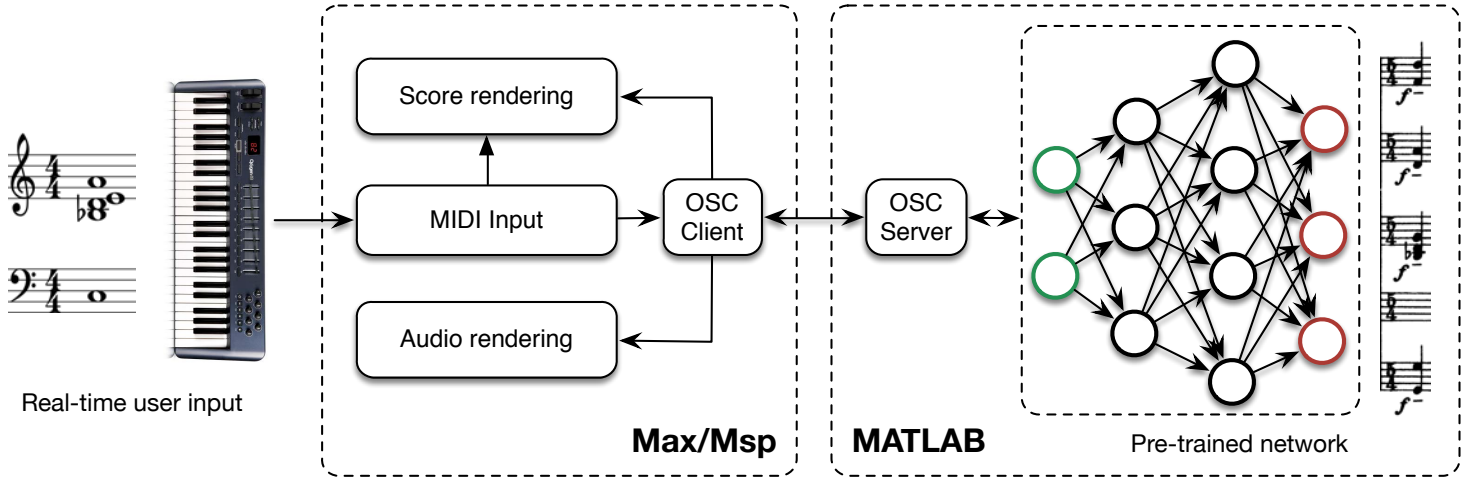
Binary cross-entropy prediction

Neural orchestration

Léopold Crestel
(crestel@ircam.fr)



Model	Orchestral Event-level (%)
Random	0.5
Repeat	12.6
cRBM	23.2
FGcRBM	6.2



[Esling Philippe, Leopold Crestel "*Live Orchestral Piano,* the first system for real-time orchestration", SMC, 2017]

Neural orchestration

Léopold Crestel
(crestel@ircam.fr)

Score

Neural Tchaïkovsky

Léopold Crestel and his RNN-LSTM string quatuor

Violin

Violoncello

Contrabass

Viola

fff

- Give the network the harmonic progression from Tchaïkovsky
- Let it go wild on an orchestration for a string quatuor

MIDI Render

The first orchestral MIDI DB

Léopold Crestel
(crestel@ircam.fr)

First collection of digital scores linking solo piano versions to their orchestration (or vice-versa)

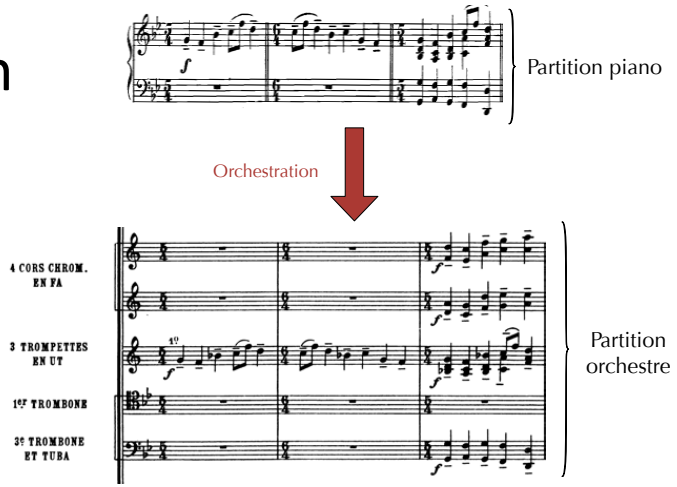
- Collected from various collections
- 223 MIDI files pairs
- 51 composers of different eras
- Hand-checked and corrected

Defined the first projective orchestration task and evaluation

Freely available for different formats

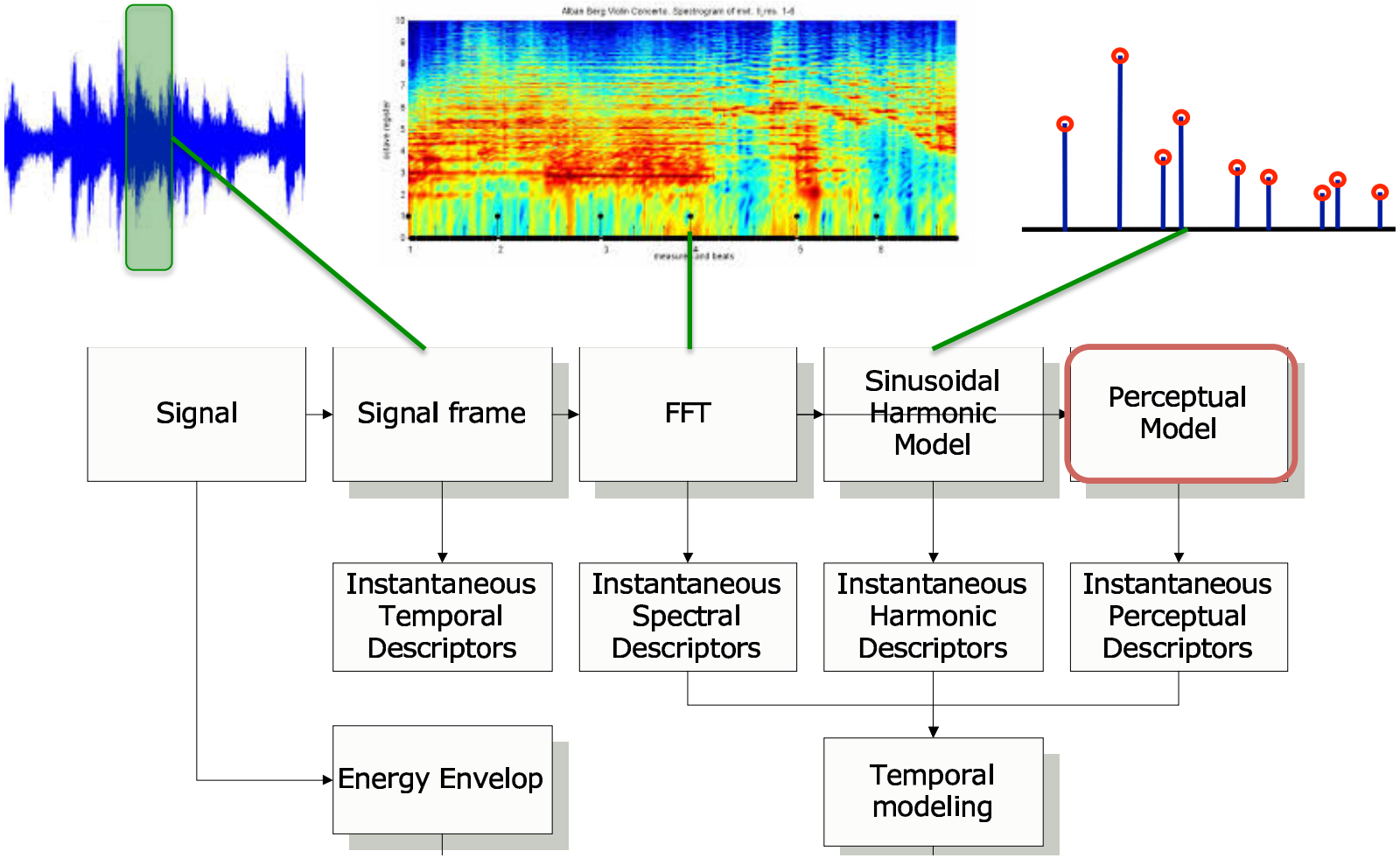
https://qsdfo.github.io/LOP/

We are still looking out for more



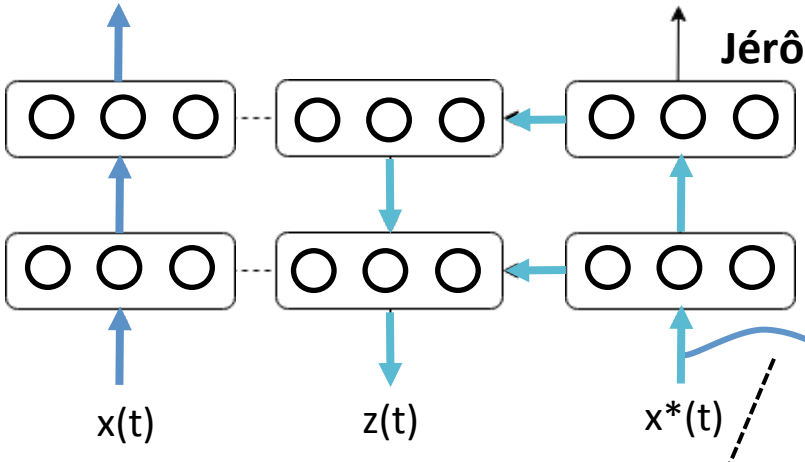
[Leopold Crestel, Philippe Esling, Lena Heng, Stephen McAdams "*A database linking piano and orchestral MIDI scores with application to automatic projective orchestration*", ISMIR, 2017]

Music as signals

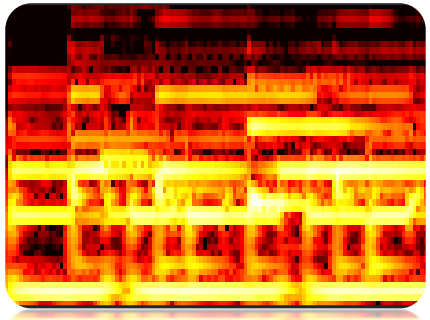


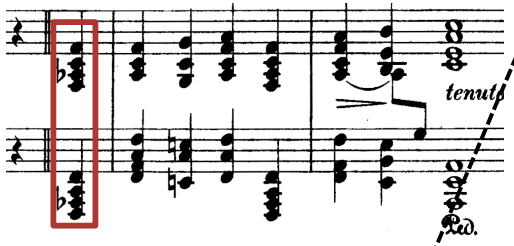
Musical co-improvisation

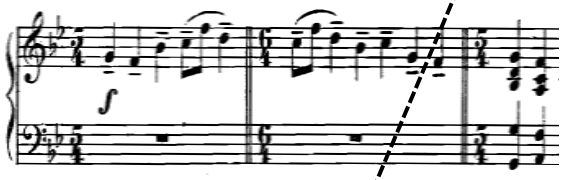
Jérôme Nika, Tristan Carsault



Scenario

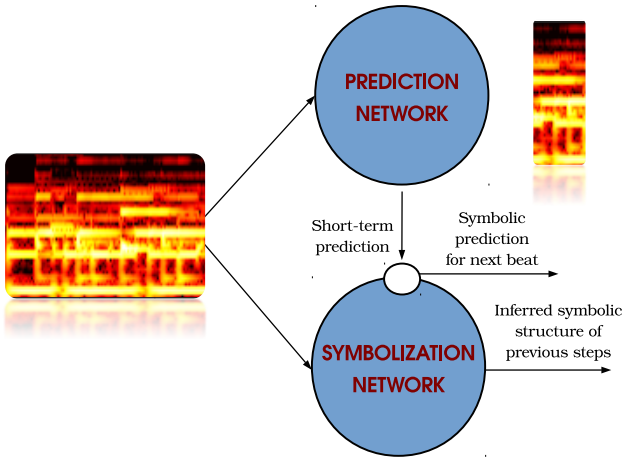






Future ?

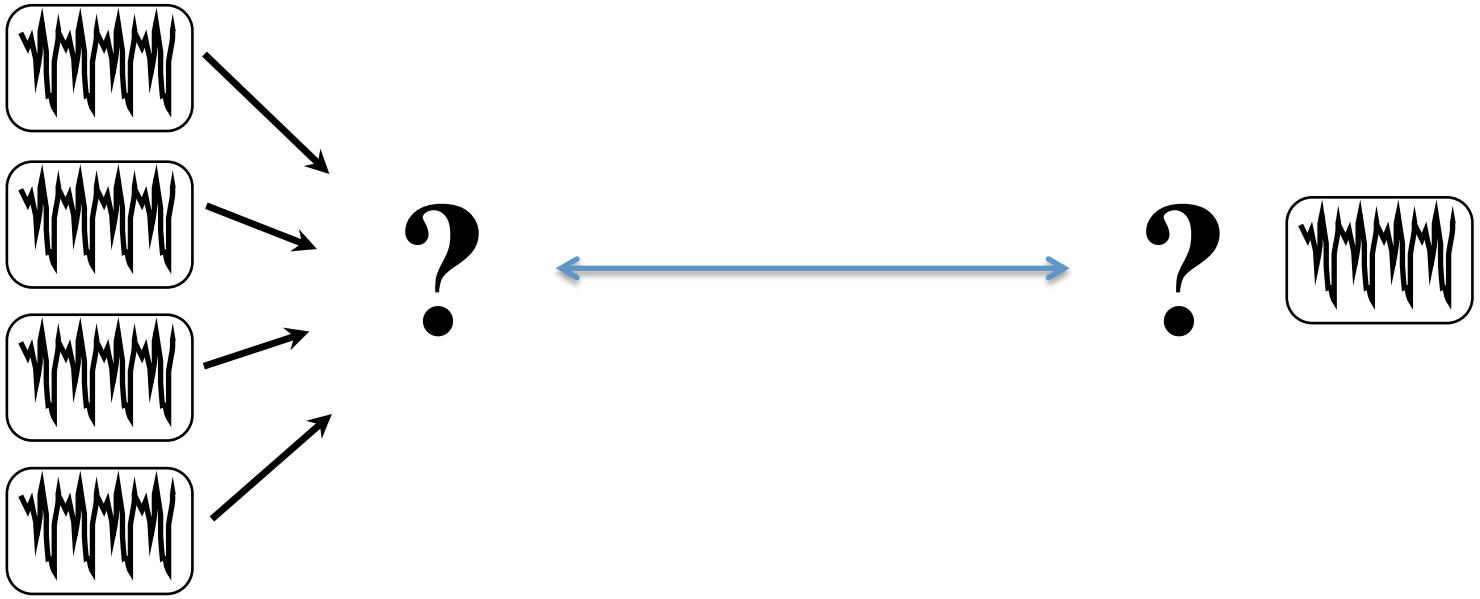
Memory



ircam Centre Pompidou

Spectral orchestration

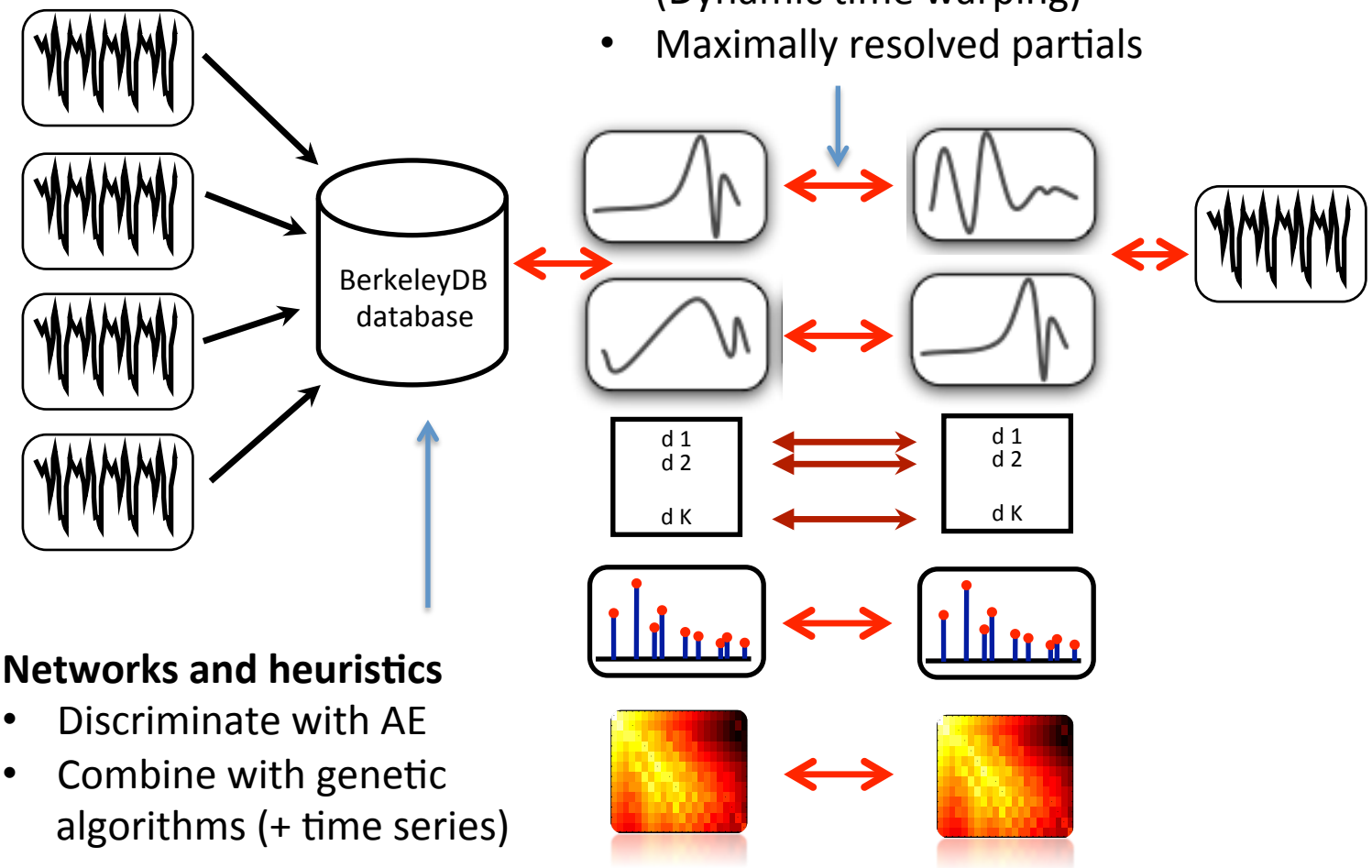
Spectral orchestration



Spectral orchestration

Perceptually motivated distances

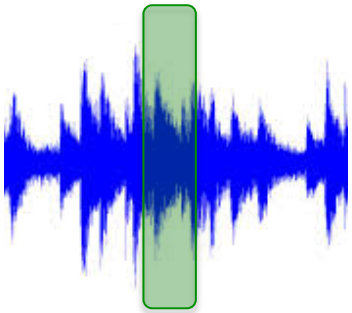
- (Dynamic time warping)
- Maximally resolved partials

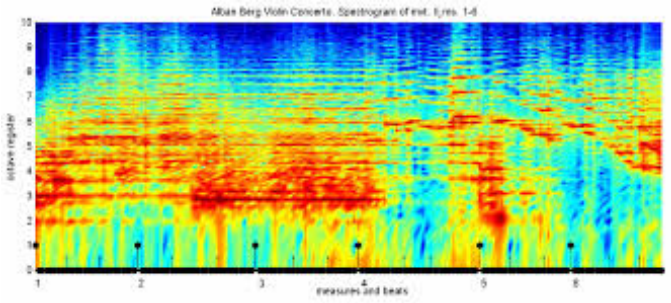


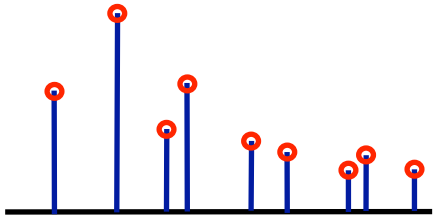
Networks and heuristics

- Discriminate with AE
- Combine with genetic algorithms (+ time series)

Spectral orchestration

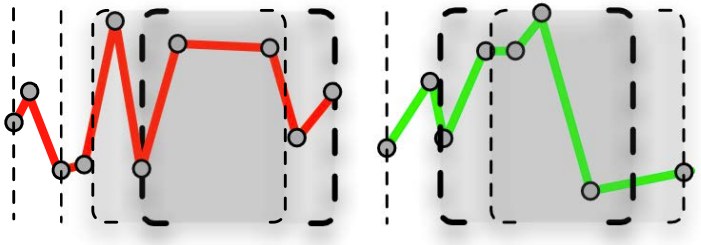


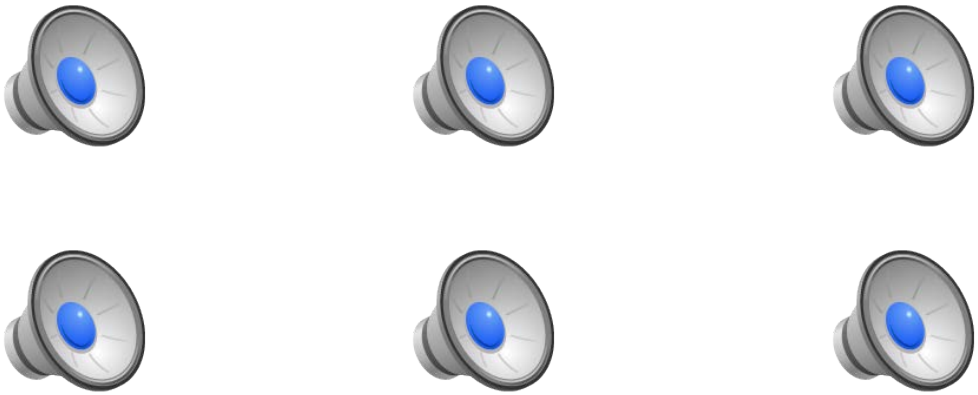




[(Target - Solution) features]

[Entropic segmentation]

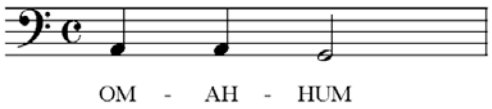




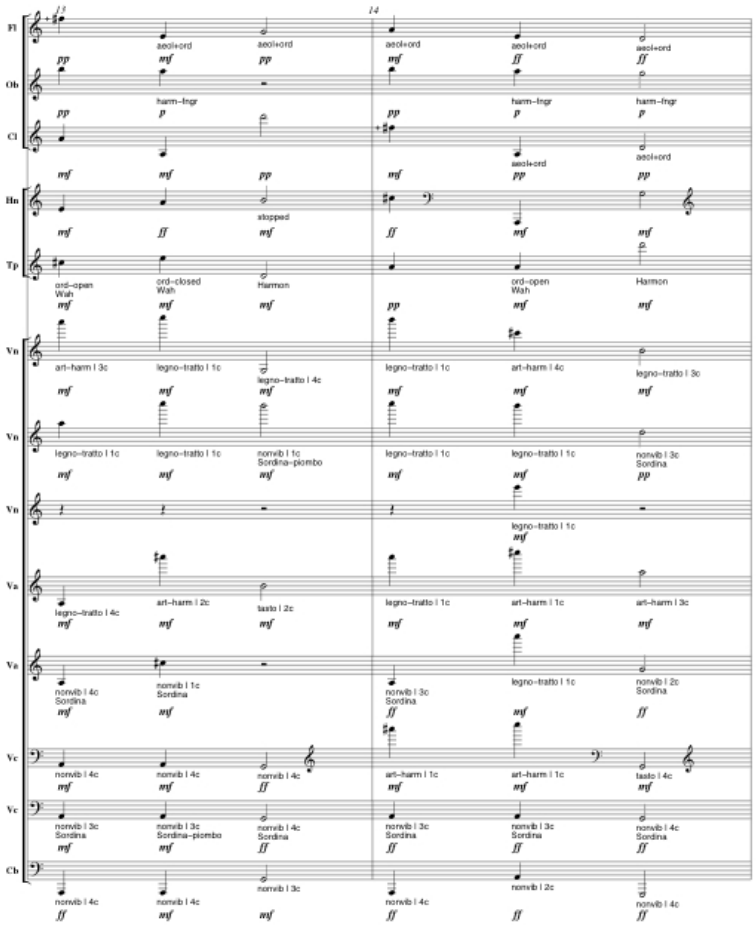
Spectral orchestration

Speakings (Jonathan Harvey)
[2008] For orchestra and electronics

Harmonic background line
(beginning of the 3rd part)

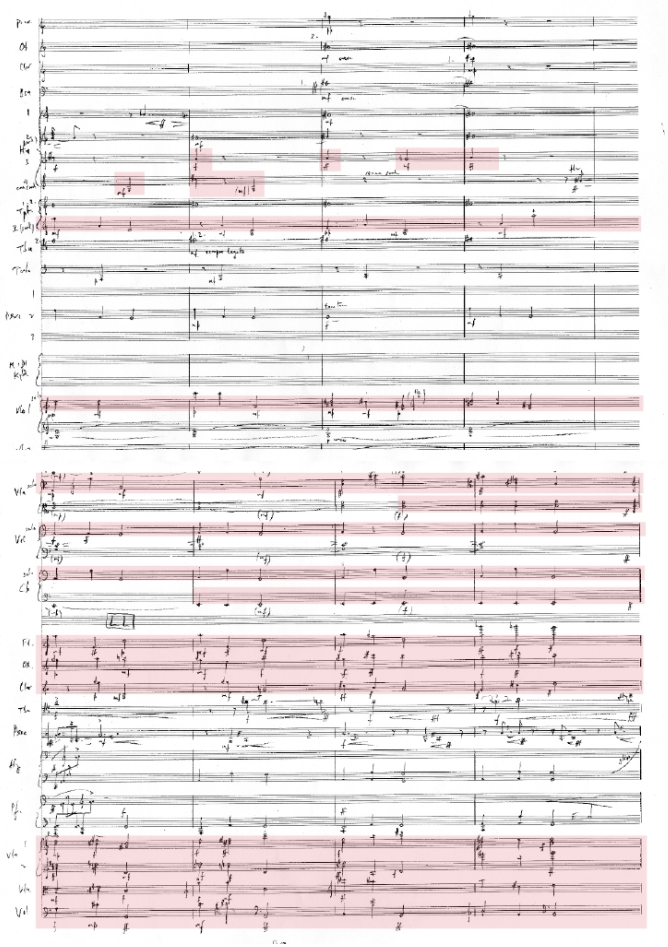


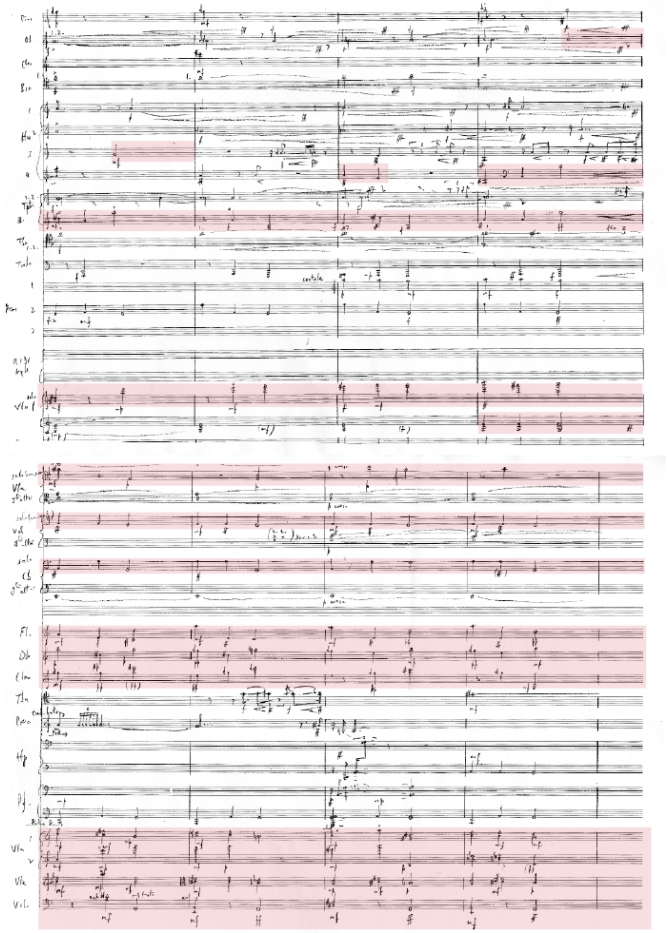
Orchestration :
- 22 ostinato repetitions
- Temporal evolution controlled by constraints



Spectral orchestration

Speakings (Jonathan Harvey) - Created August 19th, 2008 in Royal Albert Hall, London (BBC Scottish Orchestra, director Ilan Volkov)



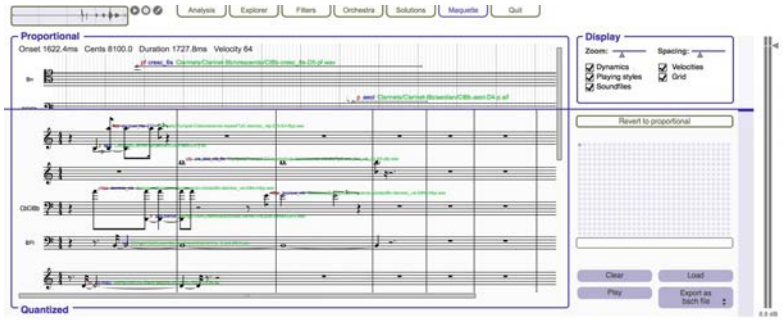


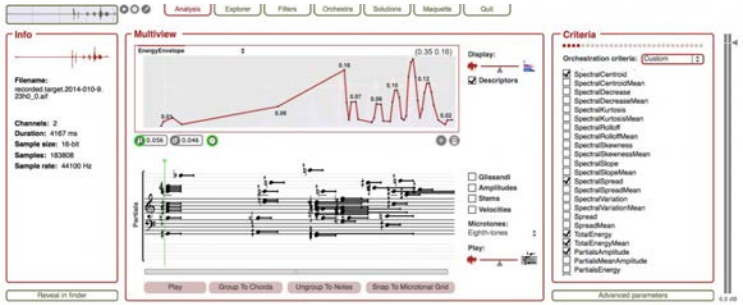
Orchids – Released and debugged

Release of Orchids in late 2014, sold on Forumnet

Already very used and proficient in musical productions (Matlab proto)

Currently GdR/GdT Orchestration every month at IRCAM





Latest version delivered on Forumnet march 2016

Improved accuracy and search heuristics

Fully multi-threaded version

Extended database

Multiple bug corrections

2 years full-time C++ developer **hired in july 2017**

Carmine Emanuele Cella work jointly at IRCAM and HEMG

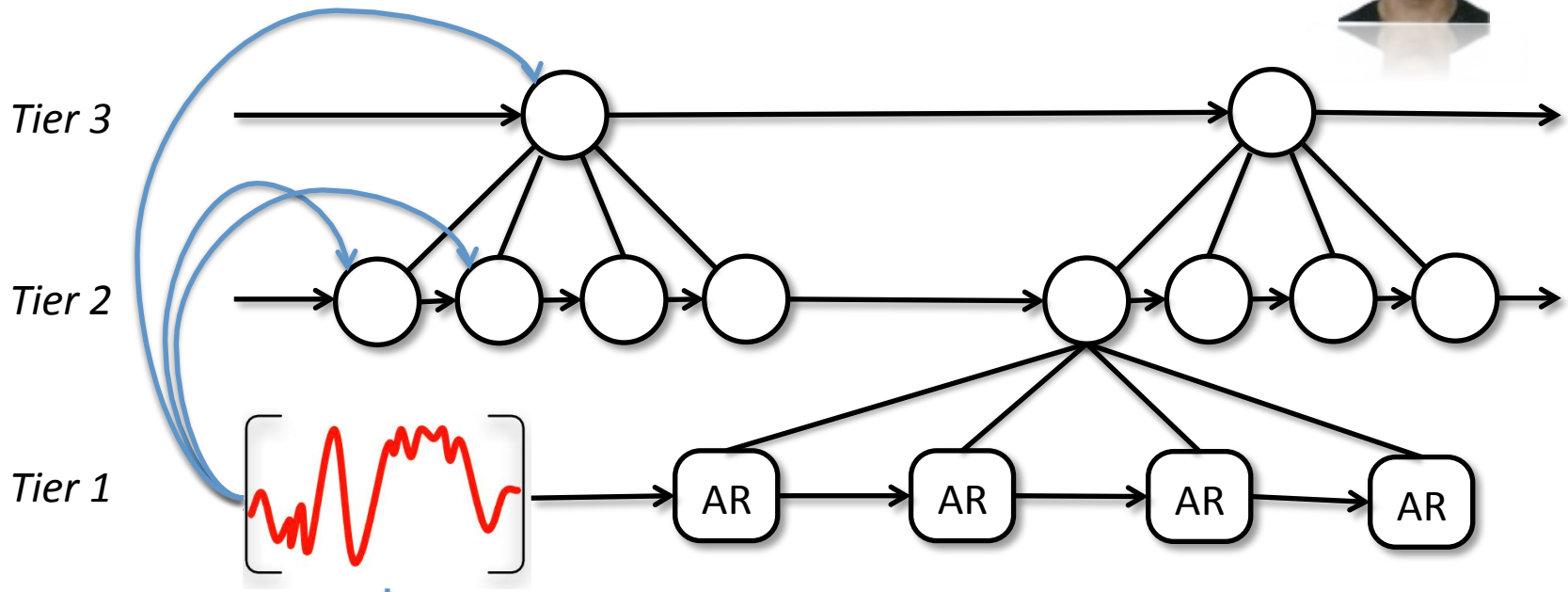
Temporal algorithms, deep learning and efficient spectral descriptors database handling

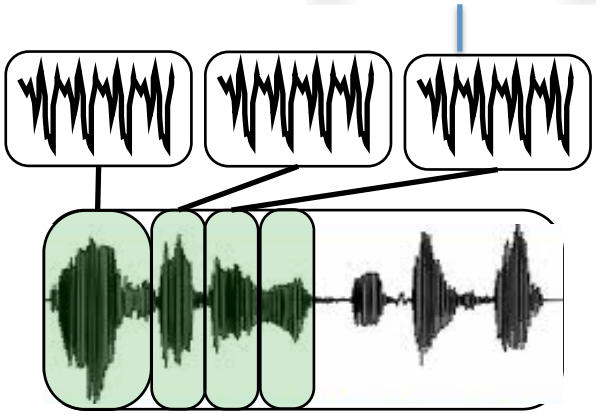
Participants : P. Esling,, D. Ghisi (residency), Y. Maresz, M. Vitorio Garcia, E. Daubresse

External collaborations : McGill, Montreal (S. McAdams, CIRMMT) – HEMG, Geneva (E. Daubresse)

SampleRNN

Daniele Ghisi, Léopold Crestel





Autoregressive behavior (order p)

$$X_t = c + \sum_{i=1}^{p} \phi_i X_{t-i} + \epsilon_t$$

$$p(\mathbf{X}) = \prod_{i=0}^{T-1} p(x_{i+1} \mid x_1, \cdots, x_i)$$

Schubert generations

- Gen. 1
- Gen. 2
- Gen. 3

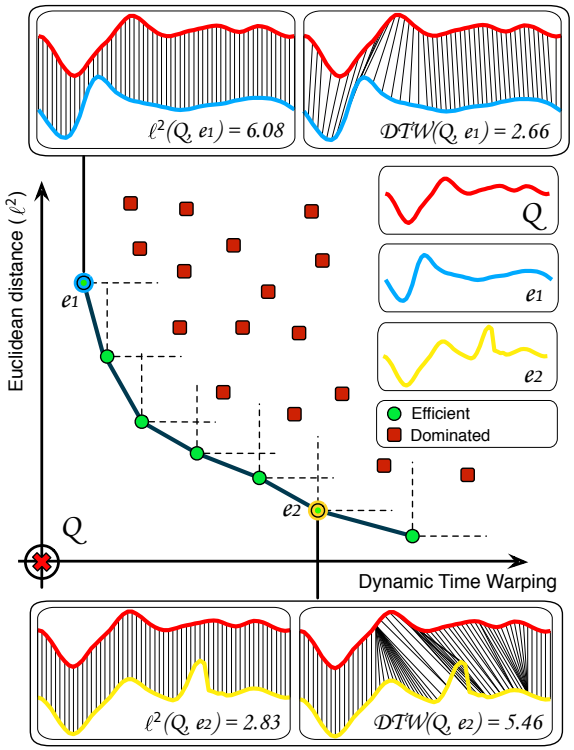
Multiple time scales deep temporal granularities learning

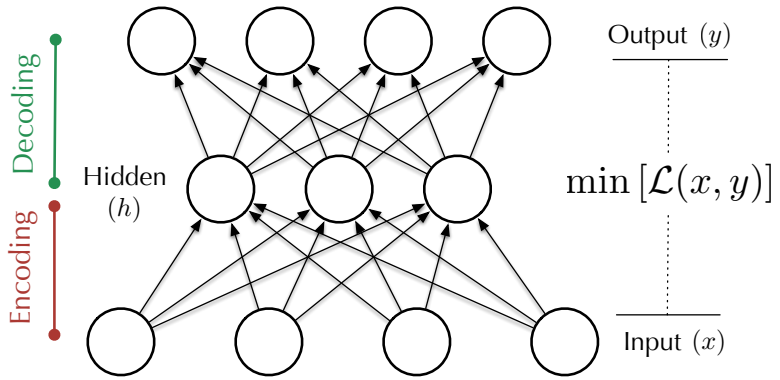
Time series at variable granularities

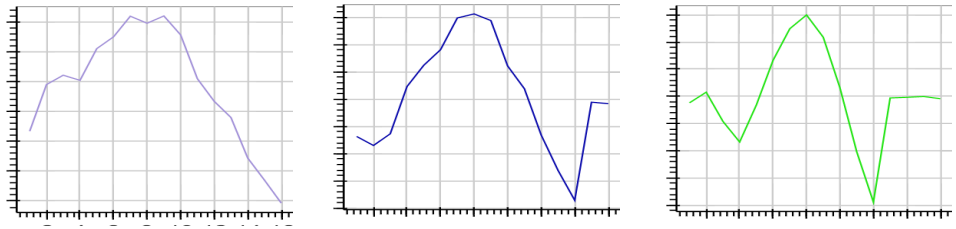
Representation of time series for learning

Deep learning and multiobjective time series

Artificial Creative Intelligence







Multiple time scales deep temporal granularities learning

Time series at variable granularities

Representation of time series for learning

Deep learning and multiobjective time series

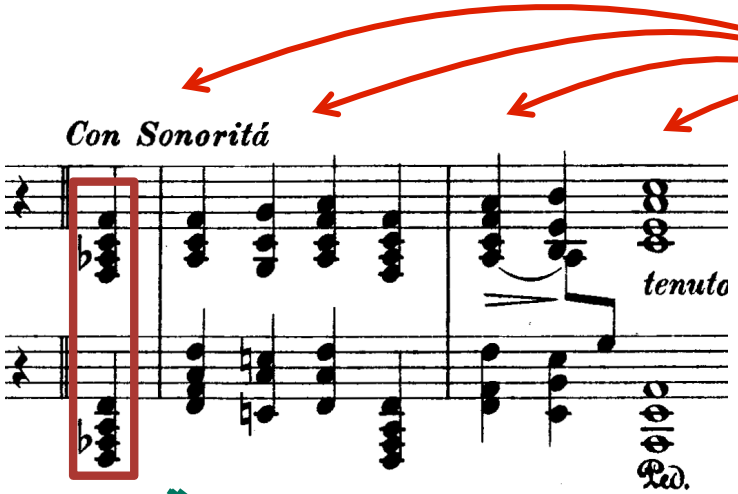
Artificial Creative Intelligence

Joint embedding spaces of signal, symbol and cognition

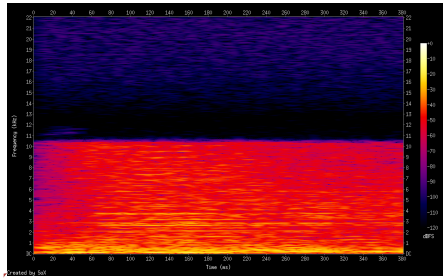
Variational space-space multimodality learning

Joint signal-symbol embedding spaces and zero-shot learning

Multimodal regularities and semantic knowledge inference



Signal



Symbol

Multiple time scales deep temporal granularities learning

Time series at variable granularities

Representation of time series for learning

Deep learning and multiobjective time series

Artificial Creative Intelligence

Joint embedding spaces of signal, symbol and cognition

Variational space-space multimodality learning

Joint signal-symbol embedding spaces and zero-shot learning

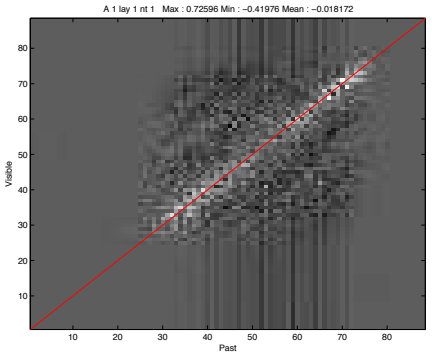
Multimodal regularities and semantic knowledge inference

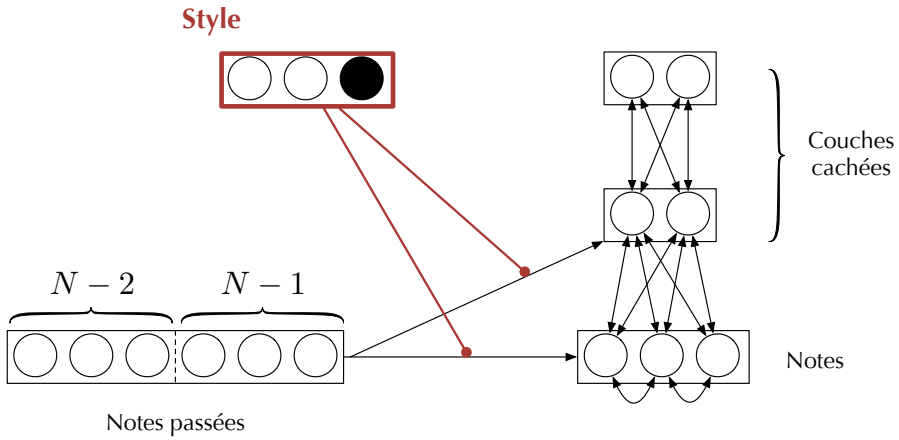
Understanding learning and knowledge extraction

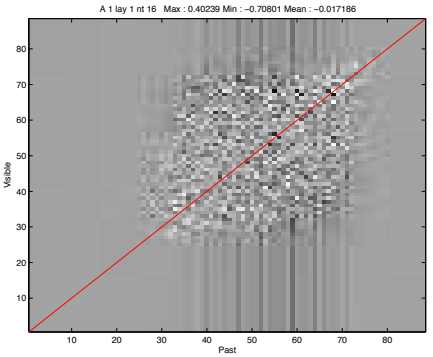
Auxiliary task transfer, indirect knowledge inference

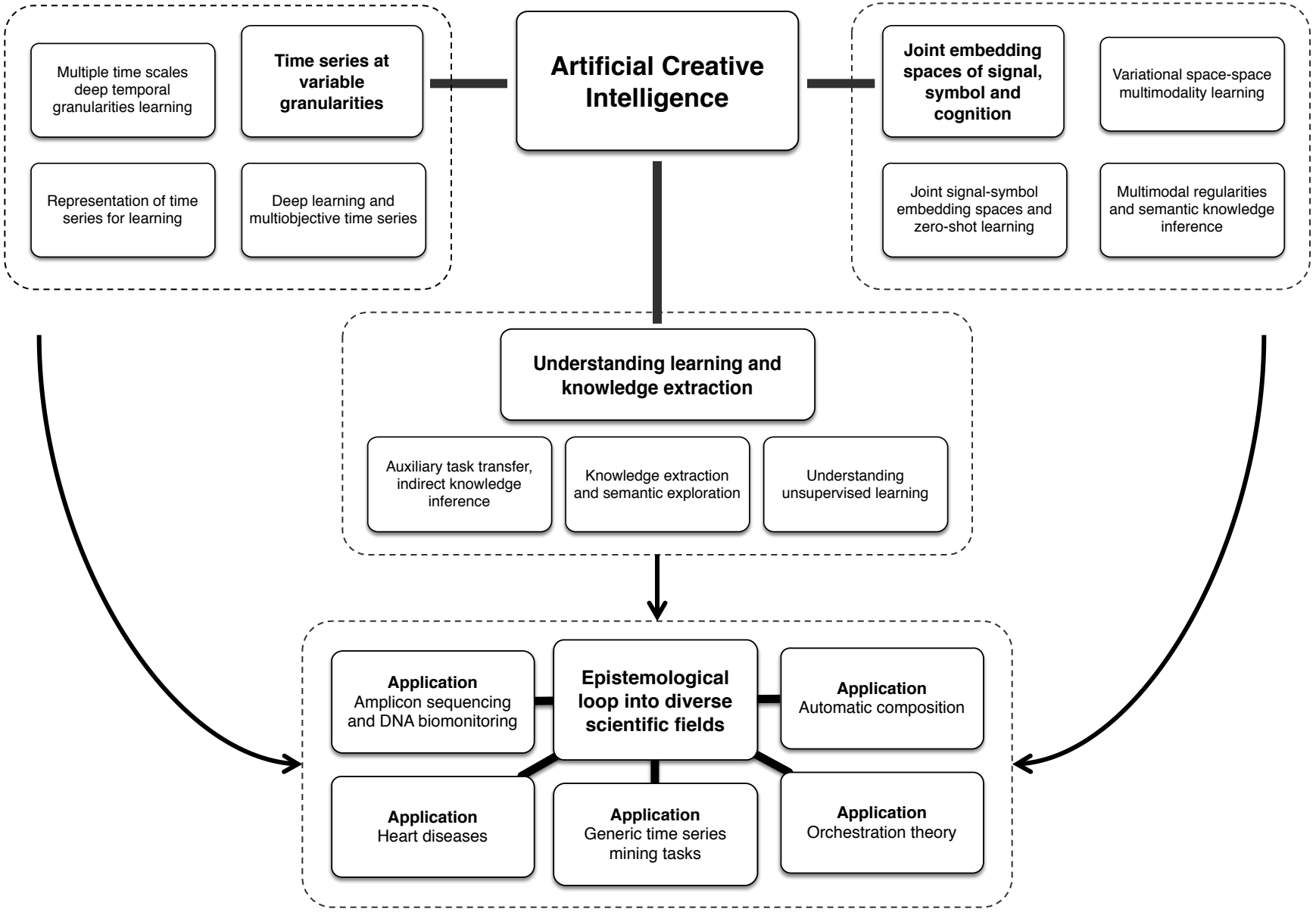
Knowledge extraction and semantic exploration

Understanding unsupervised learning








Artificial Creative Intelligence
Time series at variable granularities
Multiple time scales deep temporal granularities learning
Representation of time series for learning
Deep learning and multiobjective time series
Joint embedding spaces of signal, symbol and cognition
Variational space-space multimodality learning
Joint signal-symbol embedding spaces and zero-shot learning
Multimodal regularities and semantic knowledge inference
Understanding learning and knowledge extraction
Auxiliary task transfer, indirect knowledge inference
Knowledge extraction and semantic exploration
Understanding unsupervised learning
Epistemological loop into diverse scientific fields
Application Amplicon sequencing and DNA biomonitoring
Application Automatic composition
Application Heart diseases
Application Generic time series mining tasks
Application Orchestration theory

IRCAM – Before ACTOR

S3AM

RepMus

AnaSyn

APM

PDS

EAC

1 permanent researcher

Philippe Esling

1 PhD + 1 intern

Carsault, Crestel, Prang

Research group of composers

Y. Maresz

D. Ghisi

(and whole GdR Orchestration = ~30 composers)

ircam Centre Pompidou

IRCAM – Inside ACTOR

S3AM

RepMus

AnaSyn

APM

PDS

EAC

Patrick Susini

Olivier Warusfel
Markus Noisternig

Nicolas Donin

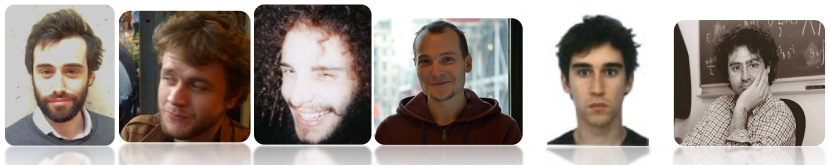
3 permanent researchers

Carlos Agon

Gerard Assayag

Philippe Esling

5 PhDs and 1 developer



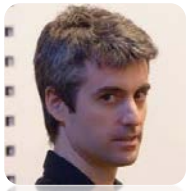
Bitton, Carsault, *Cella (Dev.)*, Chemla, Crestel, Prang

Large research group of composers

Y. Maresz

D. Ghisi

K. Haddad

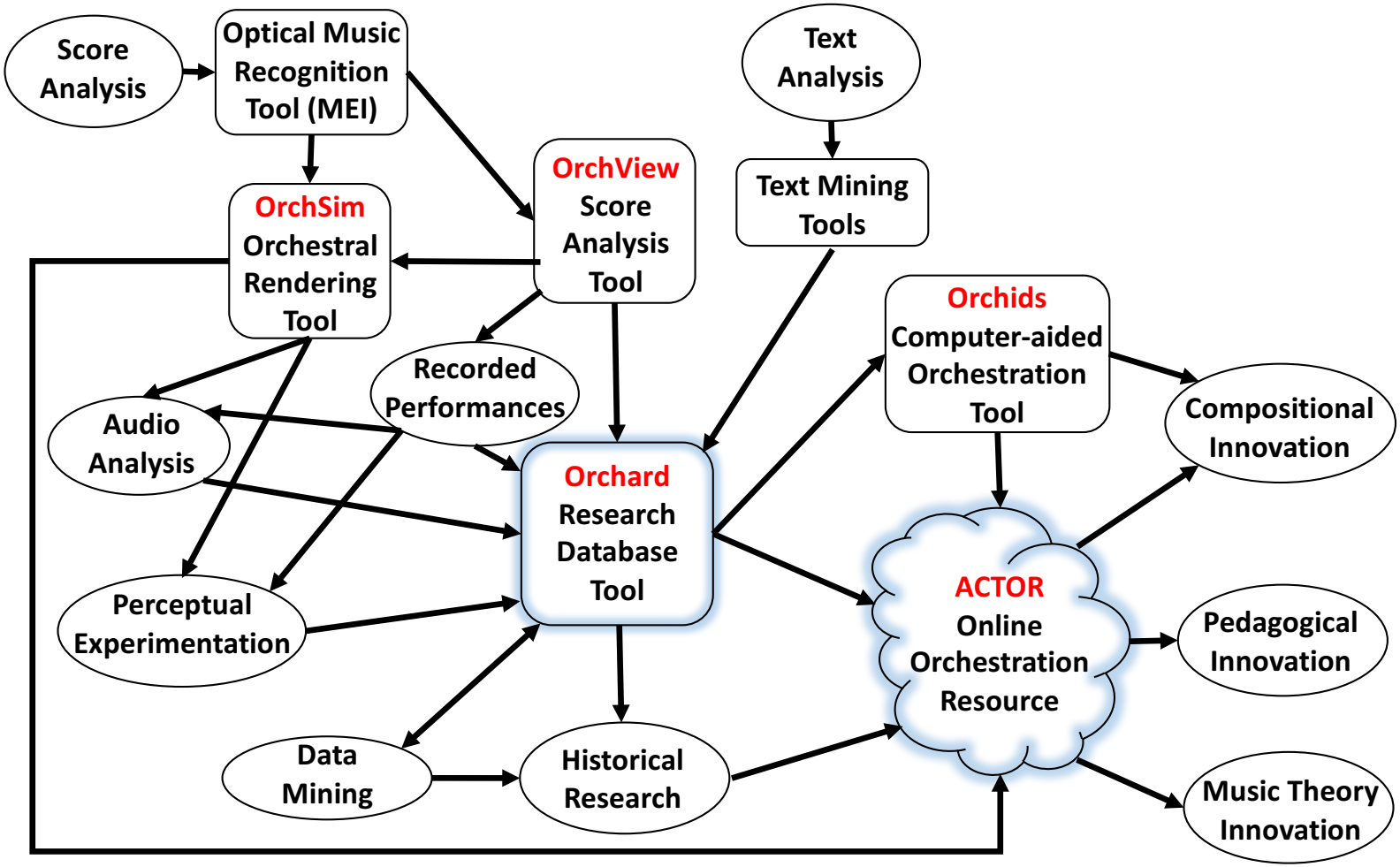


C. Castellarnau

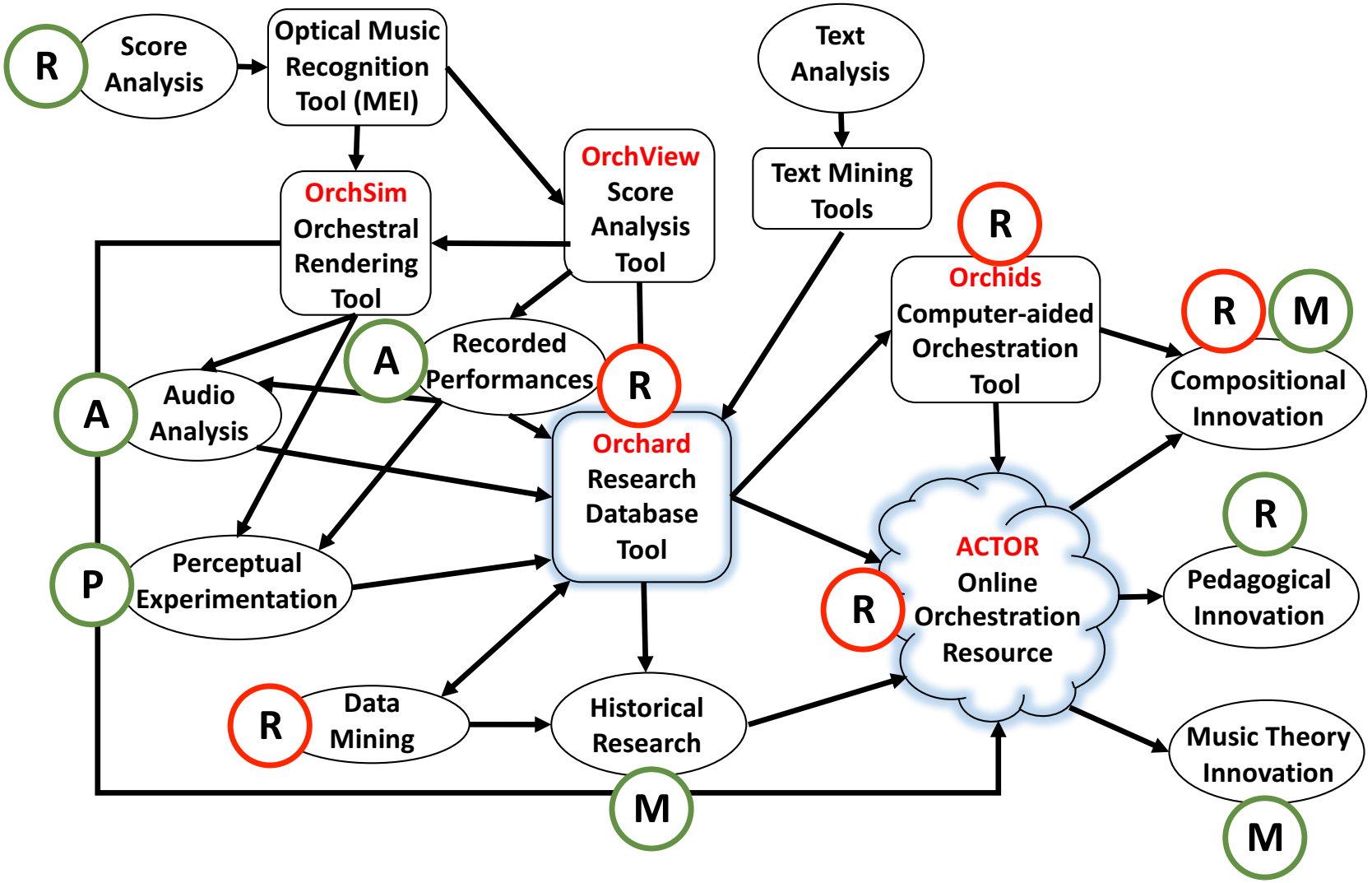
(and whole GdR Orchestration = ~30 composers)

ircam
Centre
Pompidou

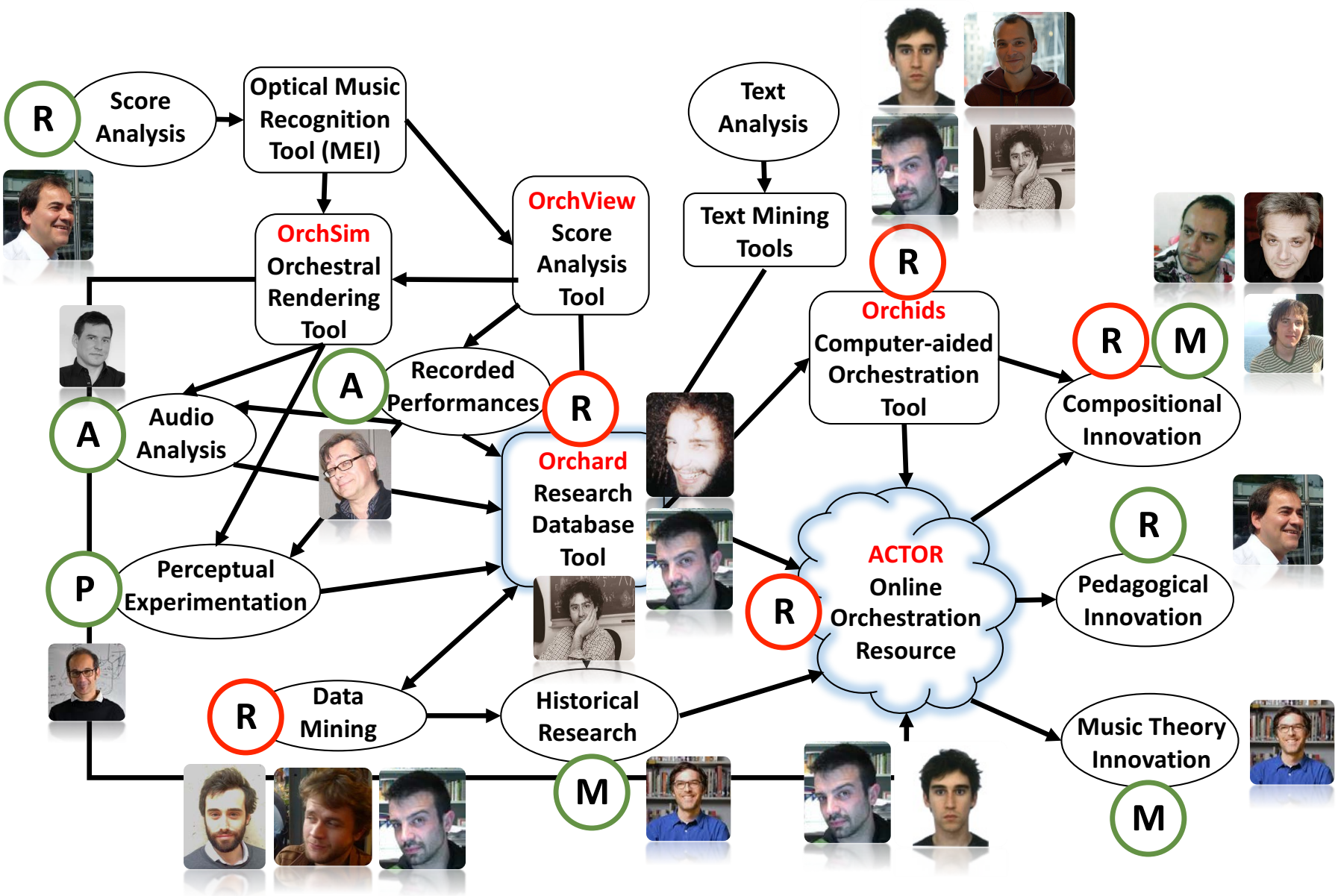
ACTOR Project Structure



ACTOR Project Structure



ACTOR Project Structure



Implications in ACTOR project

- Fully dedicated research team for ACI and orchestration
  - 1 permanent researcher
  - 5 PhDs (3 of those starting in 2017)
  - 1 C++ developer
- Additionnal research team
  - 6 permanent researchers
  - From 4 different teams at IRCAM
- Full research group in orchestration (GdR Orchestration)
- Large interest in the IRCAM composer community
- In-house collaboration with composers for high-level user feedback and assessment in high-impact artistic production
- Public conferences in the Manifeste and Forumnet
- Centre Pompidou : large exposure to public and media
- Creation of a dedicated research team

Technologies provided in ACTOR

- Sets of 1st of kind computer-aided orchestration softwares
  - Orchids: automatic orchestration from sound targets
  - Live Orchestral Piano (L.O.P) for real-time composition
- Newly developed ML algorithms
  - Orchestral embedding
  - Multimodal processing
- Future TimeDB development
- Several newly developed databases
  - Large database of classical MIDI files
  - Projective Orchestration Database (POD)
  - Currently collected multimodal database
- Access to the instrumental databases
  - StudioOnLine (SOL)
  - Future TICEL project
- All forumnet technologies

Thank you for your questions !

esling@ircam.fr