Semantic Audio Engineering and Live Sound Reinforcement

Ryan Stables, DMT Lab
Birmingham City University

SEMANTIC AUDIO.CO.UK

Overview…

- Problem(s) definition
- Semantic Audio
  - Semantic mixing systems
  - Semantic audio effects
  - Live sound solutions

- The SAFE Project
  - In-DAW data retrieval
  - Initial findings
  - Plug-in tutorial

- Live sound reinforcement
  - Midas Pro 2c experiments

- The future…

Semantic Audio

Musical Semantics

- Extracting *meaning* from musical signals

- This covers a lot of different areas (e.g.)
    - ***Speech Recognition***: can we decode the vocalist's intention?
    - ***Music Informatics***: can we extract the composer's intention via some musical abstraction.
    - ***Semantic Web***: can musical data be packaged in a transferable, searchable and comparable format.
    - ***Signal Separation***: can we decompose signals into meaningful subsets?

Semantic Audio

A few applications…

- Music similarity searching and recommender systems
- Cover song identification
- Music transcription and score alignment
- Automated remixing/reproduction systems
- Digital archiving and retrieval
- Performance analysis: tutorials, simulation, etc…

SEMANTIC AUDIO.CO.UK

Semantic Audio

Musical Semantics and Audio Engineering

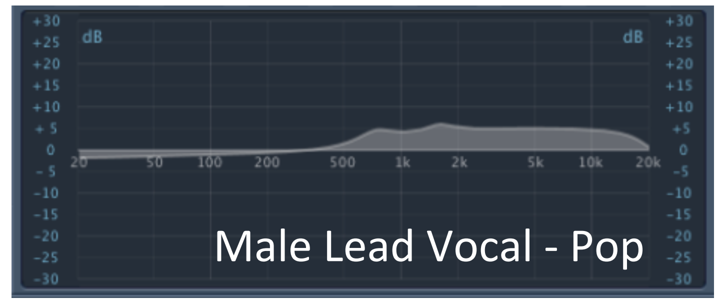
1. Engineers generally talk in a language that is hard to define computationally.
   - *The bass needs to be tighter, the toms are ringing, the vocals need to be more prominent in the mix.*

2. Parameters of music production systems generally address low-level attributes of music processing
   - Compressor threshold, relative gains, filter parameters.

There is a complex, non-linear relationship between the language in (1) and the parameters in (2).

Semantic Audio

- Context dependent/External Influences



Semantic Audio

Specifically: Live Sound Reinforcement

- Can we provide abstractions to make interfaces more intuitive?
- Can we make the music production workflow more efficient?
- *Can we extract/generate useful metadata during the process?*

Semantic Audio Engineering

Semantic Mixing

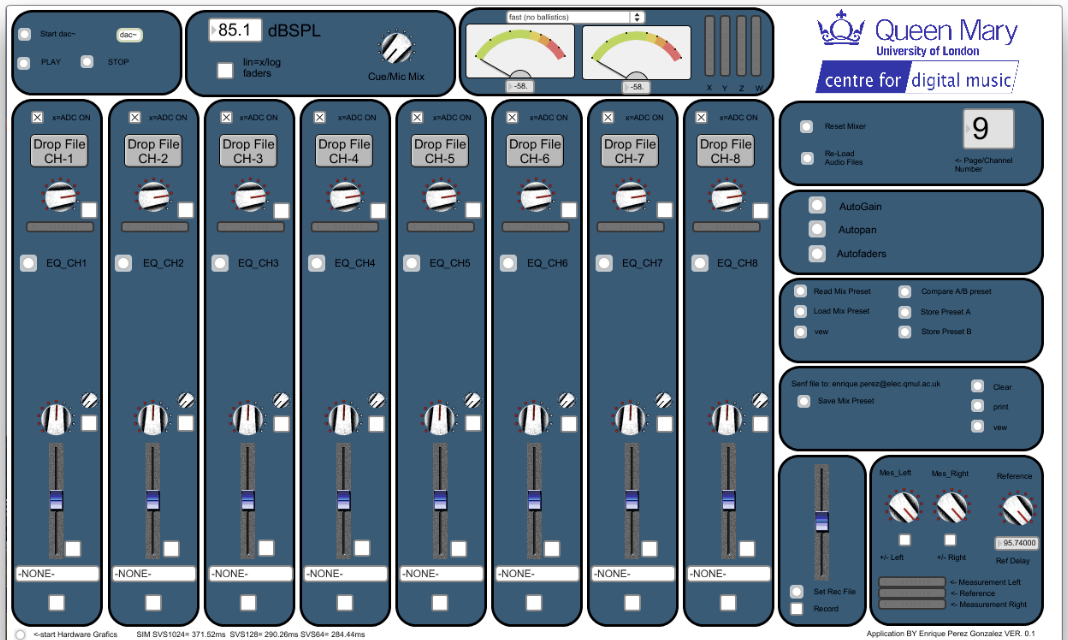
Automatic Mixing Systems

- Dugan, D. (1975)
- Campbell, E. & Whitmore, R. (1982)
- Dannenberg, R. (2007)
- Perez-Gonzalez, E. & Reiss, J. (2009)
- Ward, D. et al. (2012)

Concept: Automatically balance the faders to produce an intelligible mix across all active channels.

Semantic Mixing

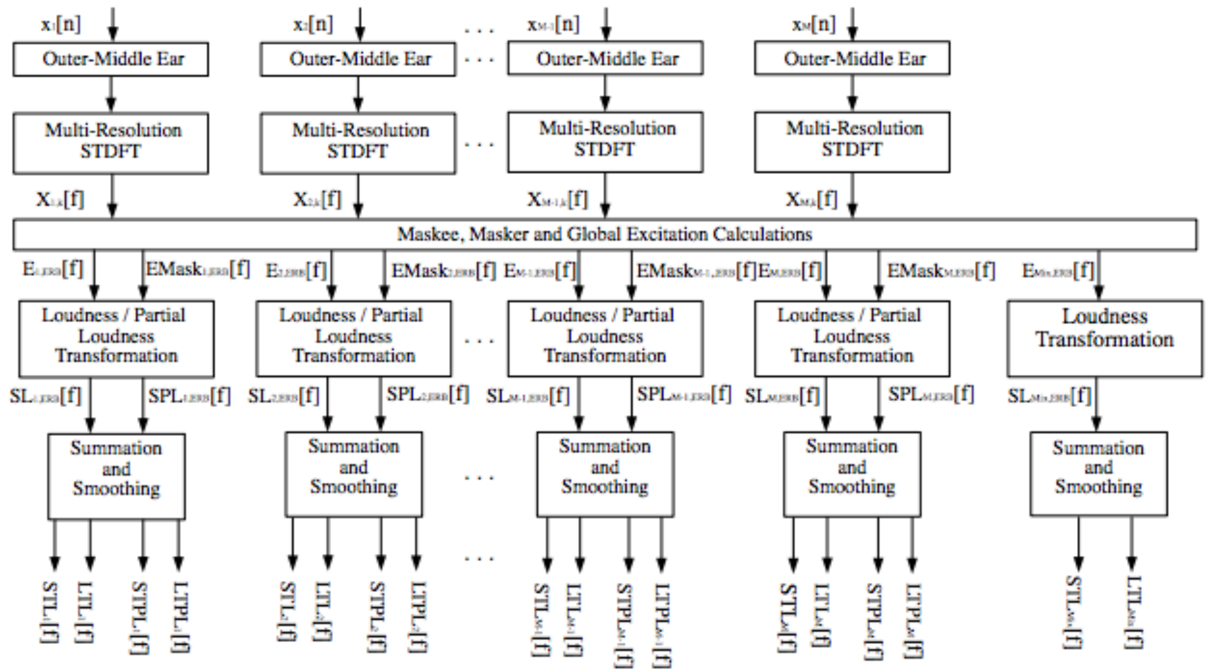
Automatic Mixer: Perez-Gonzalez, E. & Reiss, J. (2009)



- Cross-adaptively optimizes the loudness ratios between each track
- Uses inter-channel dependencies
- Capable of running in real-time, works well for live environments

Semantic Mixing

Automatic Mixer: Ward, D., Reiss, J. & Athwal, C. (2012)



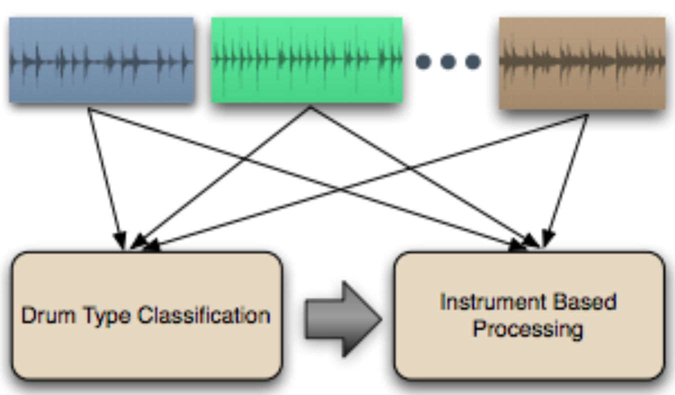
- Improvements to cross-channel intelligibility based on complex hearing models
- Optimization required for real-time application due to computational overhead

Semantic Mixing

Semantic Mixing Systems

- Scott, J. & Kim, Y. (2009)
- De Man, B. & Reiss, J. (2013)

Concept: Automatically manipulate the faders based on some embedded knowledge of the audio signals.



Instrument Class	Panning Value (θ)	Gain Values $\{\alpha, \beta, \lambda\}$
Kick Drum	0 (center)	{0.9, 1.2, 2}
Snare Drum	0 (center)	{0.9, 1.2, 2}
Toms	Spaced {-25, 25}	{0.8, 1.3, 4}
Overhead/Room	{-35, 35}	{0.8, 1.3, 4}

Scott, J. & Kim, Y. (2009)

Semantic Mixing

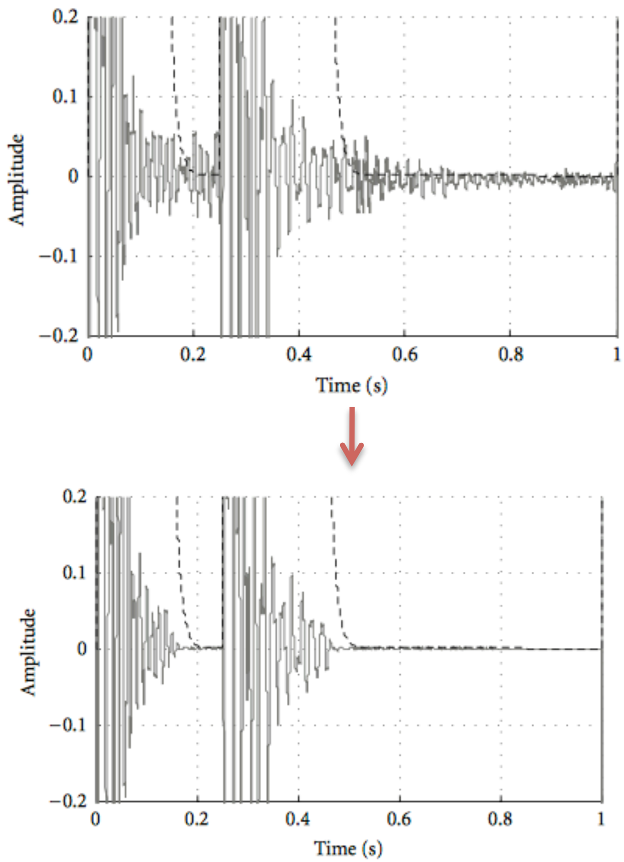
Live Sound Tools

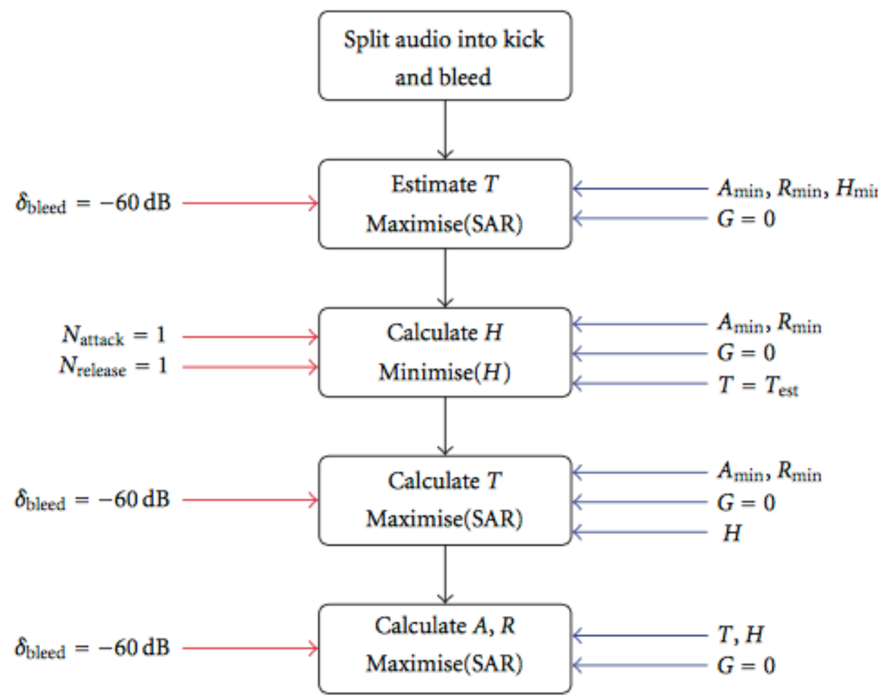
- Terrell, M. et al. (2010) – Automatic noise gating
- Perez-Gonzalez, E. & Reiss, J. (2010) – Semi-autonomous panning
- Clifford, A. & Reiss, J. (2011) – Proximity effect detection
- Clifford, A. & Reiss, J. (2013) - Comb filtering reduction
- Braun, S. (2012) – Feedback suppression

Concept: Automatically mitigate/address issues that arise in the live sound environment.

Semantic Mixing

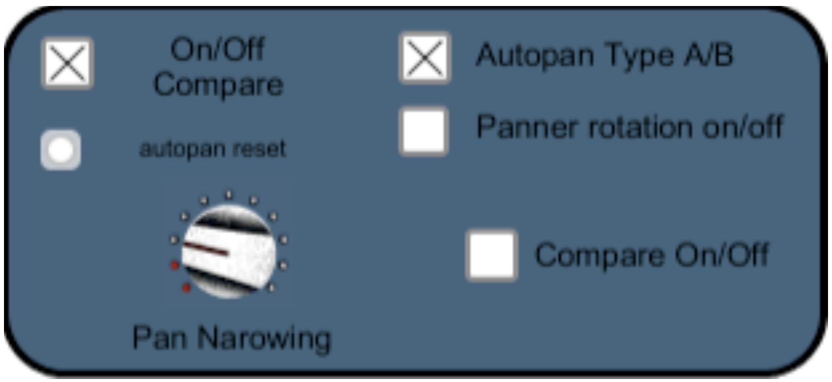
Terrell, M. et al. (2010) – Automatic noise gating (removal of bleed)

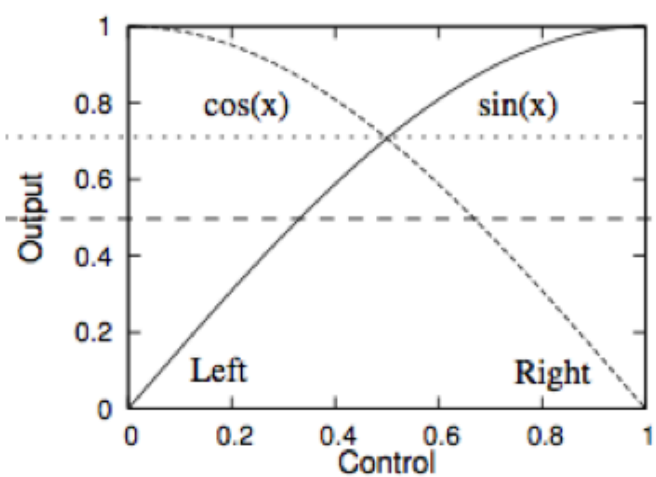


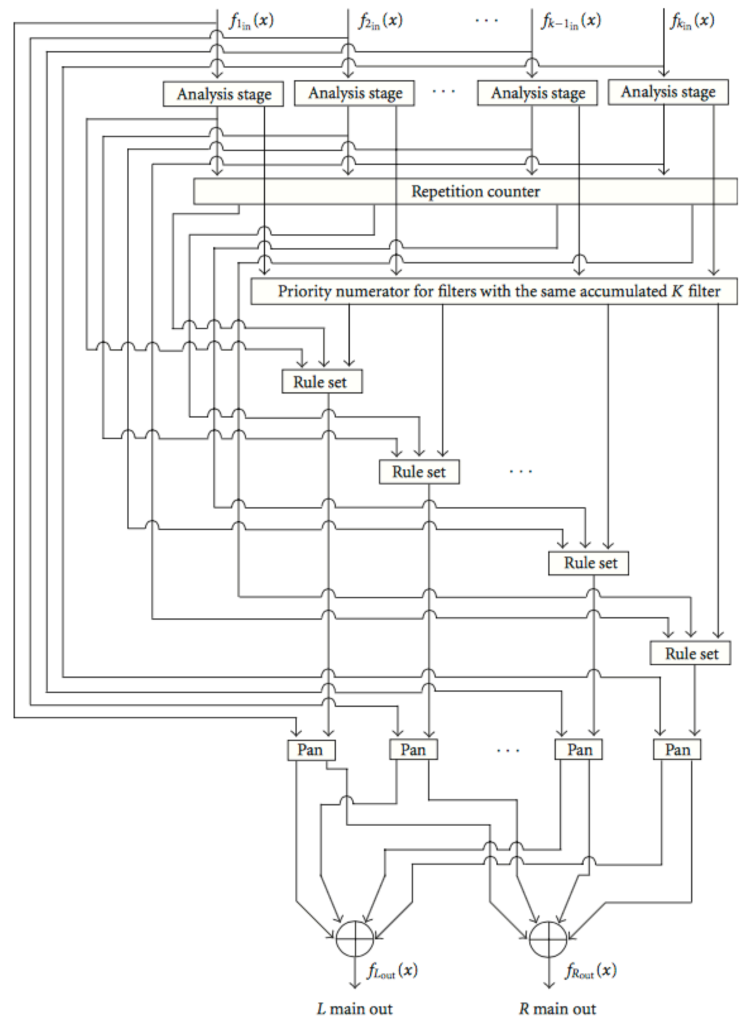


Semantic Mixing

Perez-Gonzalez, E. & Reiss, J. (2010) – Semi-autonomous panning







Semantic Mixing

Semantic Audio Effects

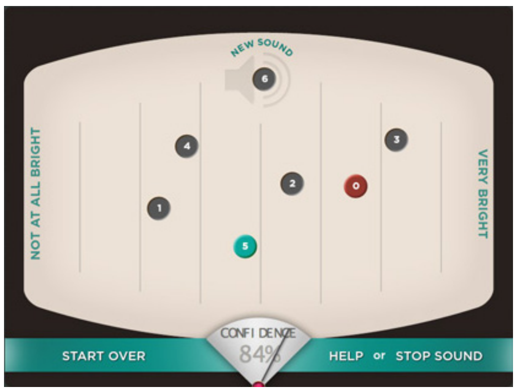
- Sabin, A. & Pardo, B. (2008) – 2DEQ
- Cartwright, M. & Pardo, B. (2013) – Social EQ
- Seetharaman, P. & Pardo, B. (2014) - Reverbalize
- Stables et al. (2014) – Semantic multi-effects
- Ma, Z. et al. (2015) – Intelligent dynamic range compression

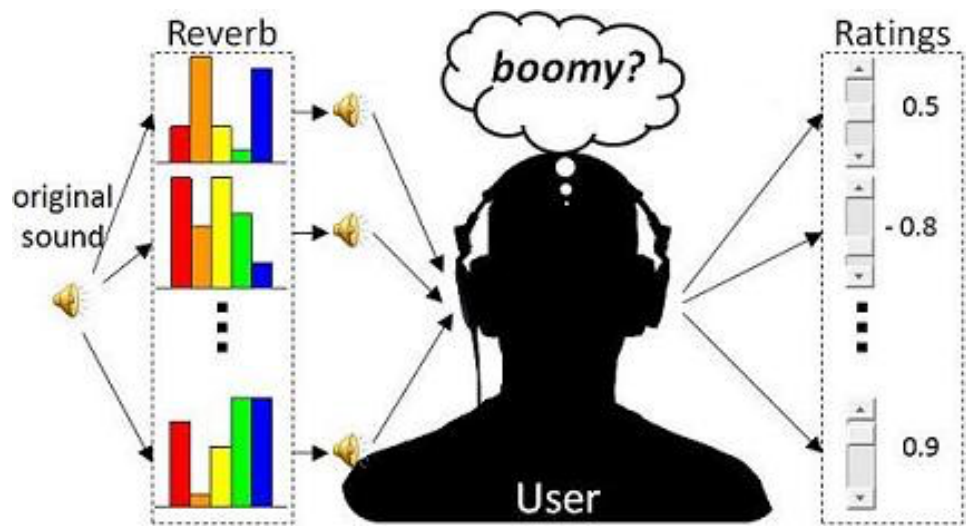
Concept: Adaptively or automatically adjust parameters of predetermined audio effects to achieve a desired context-dependent result, often via some abstraction of the parameter space.

SEMANTIC AUDIO.CO.UK

Semantic Mixing

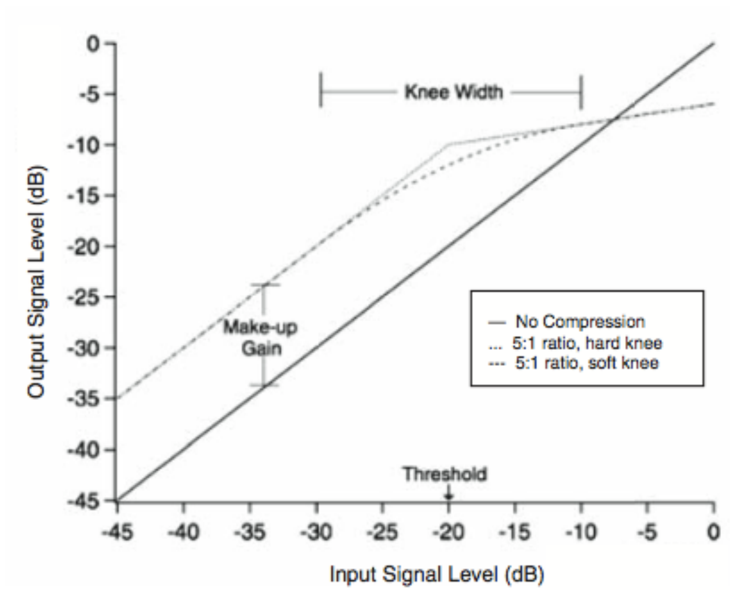
Cartwright, M. & Pardo, B. (2013) – Social EQ
Seetharaman, P. & Pardo, B. (2014) – Reverbalize

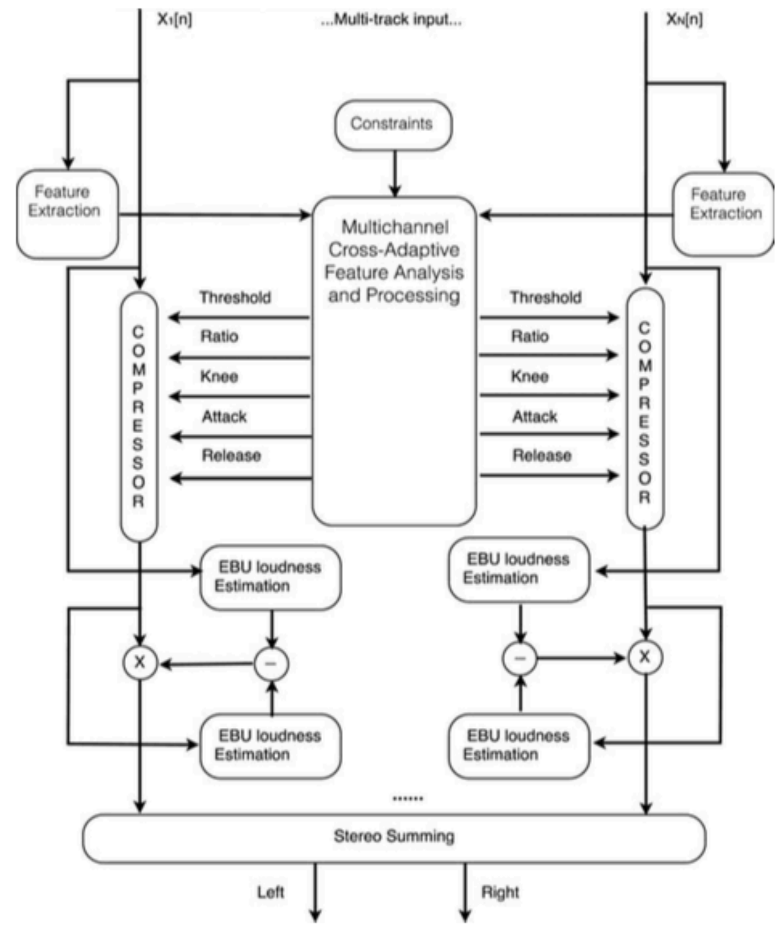




Semantic Mixing

Ma, Z., De Man, B., Pestana, P., Black, D. & Reiss, J. (2015): Intelligent Multichannel Dynamic Range Compression.





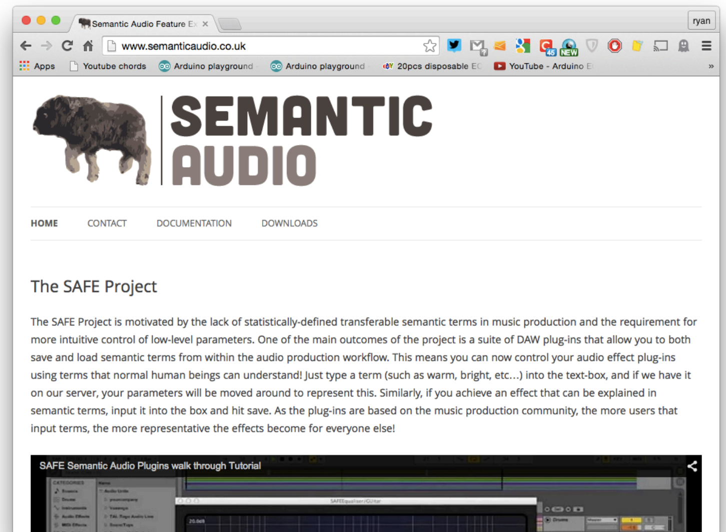
SEMANTIC AUDIO.CO.UK

The SAFE Project

The SAFE Project

SAFE: Semantic Audio Feature Extraction

- Semantic Media fund [EPSRC]



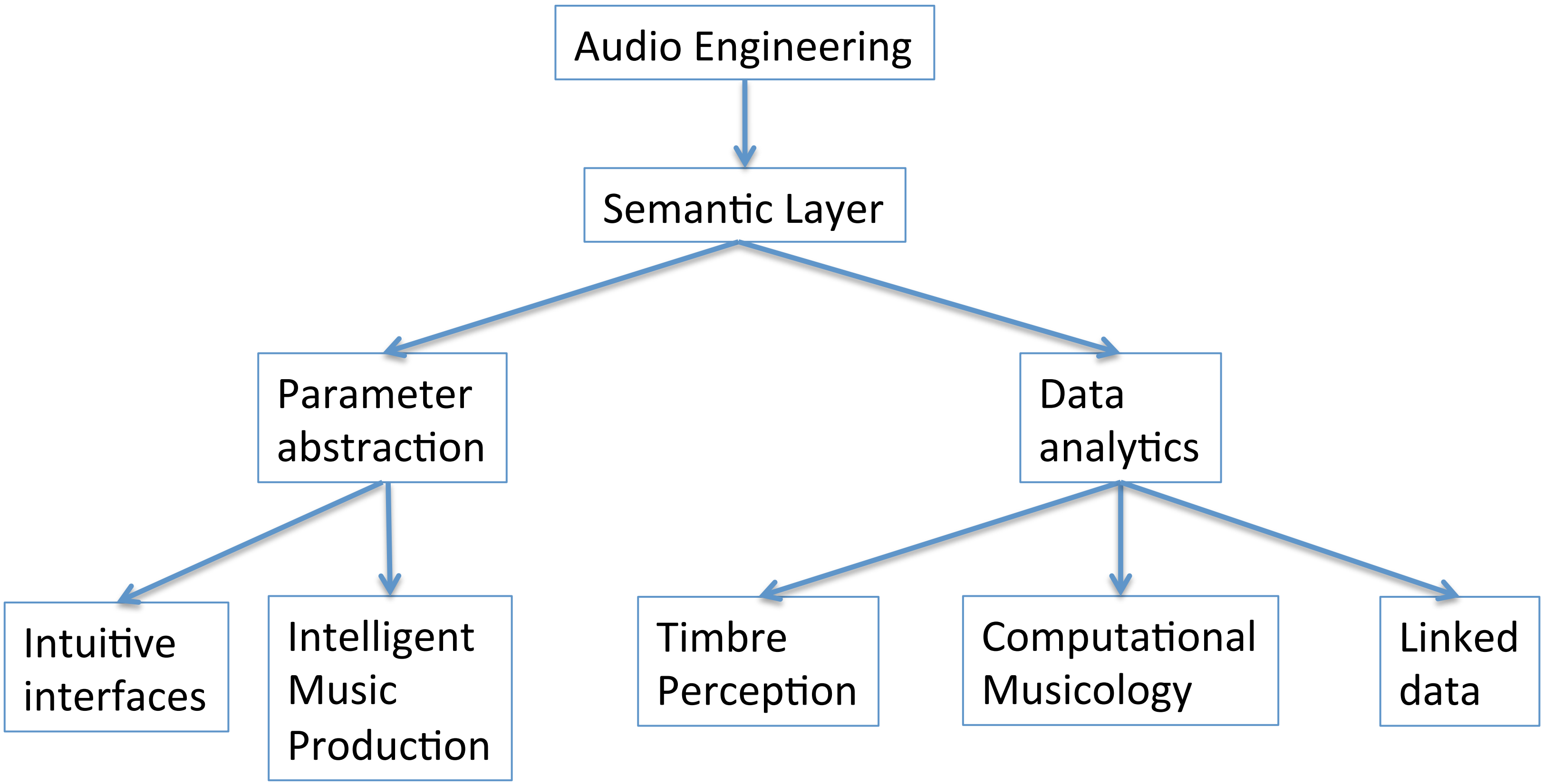
Research Question

- Can we collect representative semantic Audio Engineering data on a large scale?

Objectives:

- Extract semantically annotated metadata during the mixing process.
- Use this data to make the Audio Engineer's workflow more intuitive.

Why?

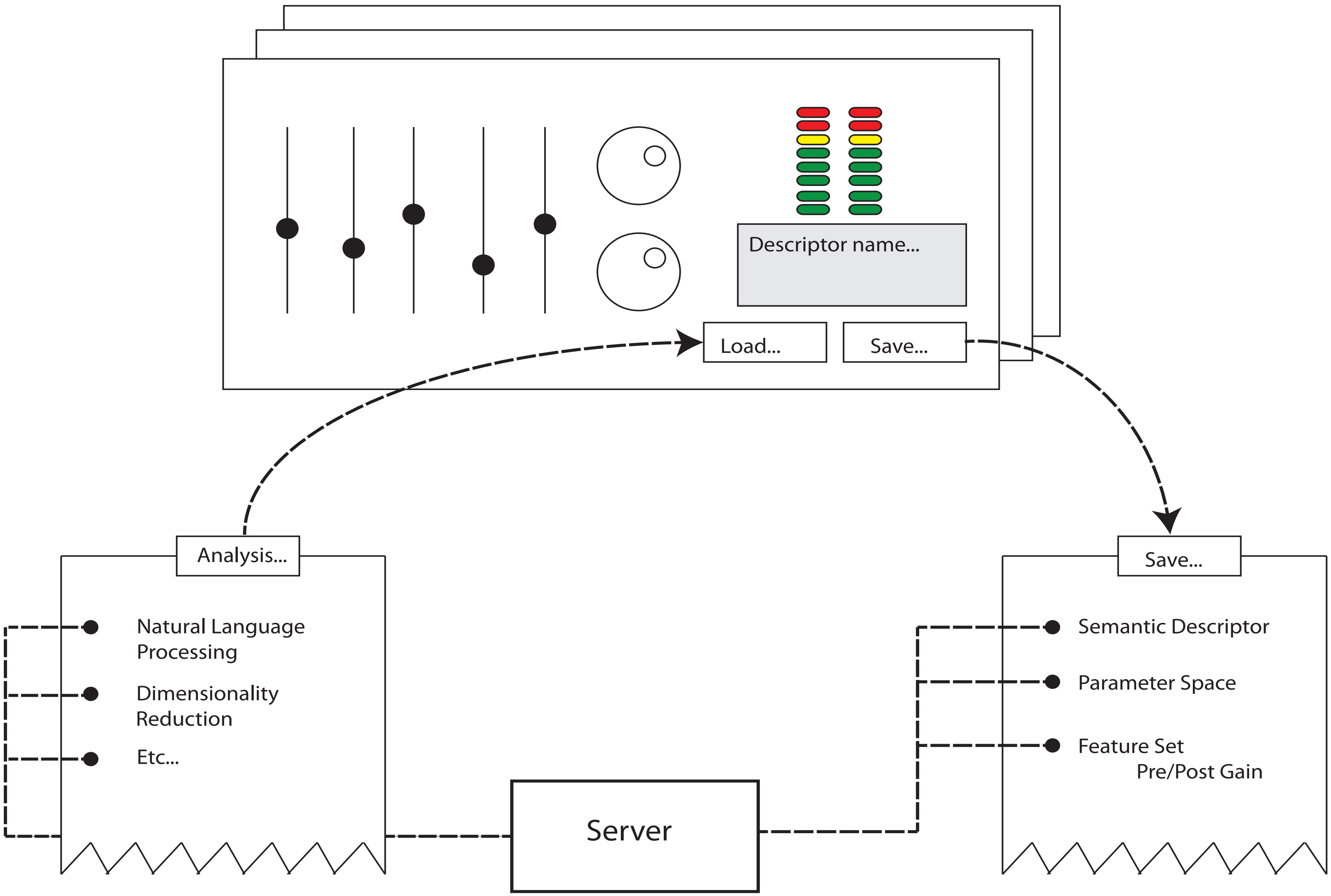

Audio Engineering
Semantic Layer
Parameter abstraction
Data analytics
Intuitive interfaces
Intelligent Music Production
Timbre Perception
Computational Musicology
Linked data

SAFE Plug-ins

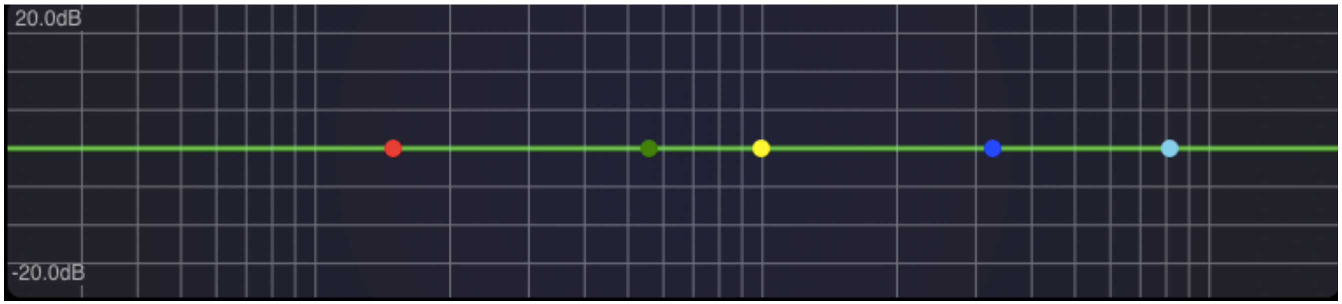
- 4 x plug-ins:
  - Parametric EQ
  - Overdrive
  - Compressor
  - Algorithmic Reverb

- Analytics
  - Semantic descriptors
  - Audio feature data
  - Parameter space
  - Metadata

SAFE: Topology



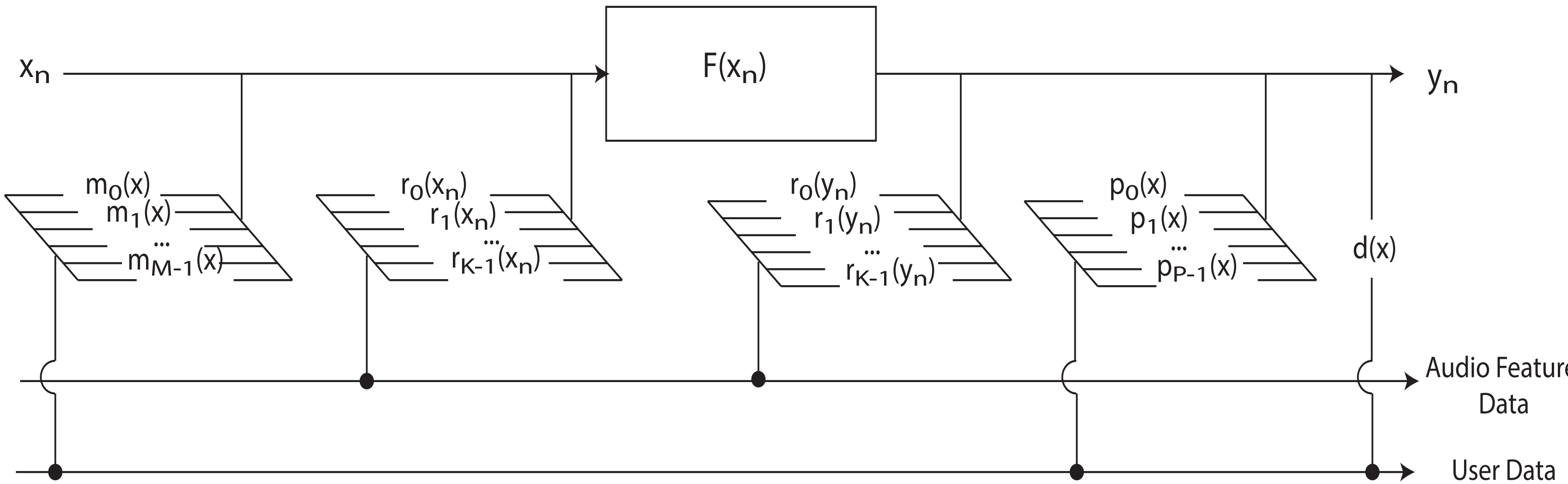
SAFE: Audio Effects



1. **Equaliser**: a five band parametric EQ with three peaking filters and two shelving filters.
2. **Distortion**: Amplitude distortion with tone control (LPF).
3. **Compressor**: a dynamic range compressor with attack and release parameters
4. **Reverb**: an algorithmic reverb based on the figure-of-eight technique proposed by Dattorro (1997)

Effects 1-3 are based on Reiss & McPherson (2014)

SAFE: Feature Extraction Architecture



- Audio features extracted per channel, per frame, per upload, before and after processing, deltas computed on server.
- Metadata extracted per-user-upload.
- 1 x Descriptor string per-upload.
- Parameter set per-upload

SAFE: Feature Extraction Architecture

```
SignalState
DataMean
Variance
Standard_Deviation
RMS_Amplitude
Zero_Crossing_Rate
Spectral_Centroid
Spectral_Variance
Spectral_Standard_Deviation
Spectral_Skewness
Spectral_Kurtosis
Irregularity_J
Irregularity_K
Fundamental
Smoothness
Spectral_Roll_Off
Spectral_Flatness
Tonality
Spectral_Crest
Spectral_Slope
```

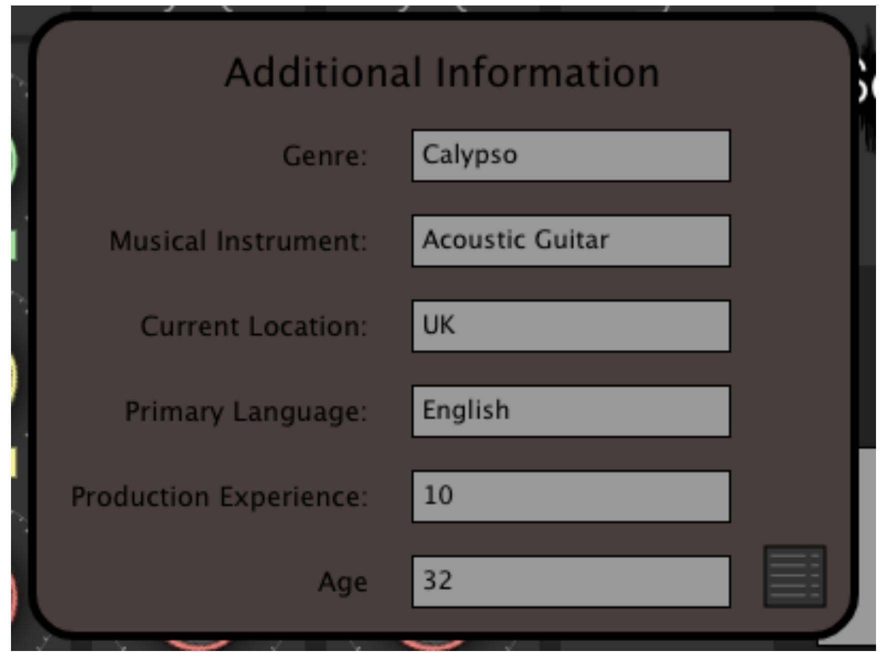
```
Spectral_Slope
Peak_Spectral_Centroid
Peak_Spectral_Variance
Peak_Spectral_Standard_Deviation
Peak_Spectral_Skewness
Peak_Spectral_Kurtosis
Peak_Irregularity_J
Peak_Irregularity_K
Peak_Tristimulus_1
Peak_Tristimulus_2
Peak_Tristimulus_3
Inharmonicity
Harmonic_Spectral_Centroid
Harmonic_Spectral_Variance
Harmonic_Spectral_Standard_Deviation
Harmonic_Spectral_Skewness
Harmonic_Spectral_Kurtosis
Harmonic_Irregularity_J
Harmonic_Irregularity_K
Harmonic_Tristimulus_1
Harmonic_Tristimulus_2
Harmonic_Tristimulus_3
Noisiness
```

```
Bark_Coefficient_25
dBark_Coefficient_25
MFCC_13
dMFCC_13
Spectral_Flux
```

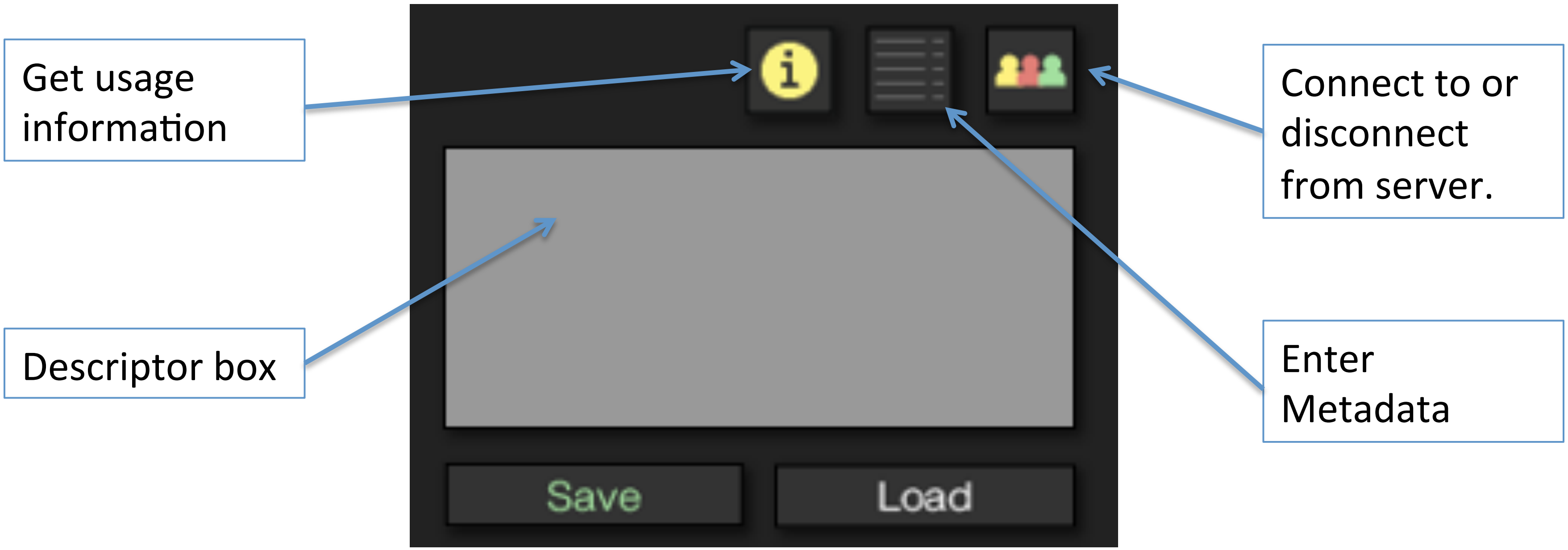
- 100+ audio features per 20ms frame of audio
- Extracted using the Libxtract C-library: Bullock (2007)
- The next version will support the VAMP framework.

SAFE: Feature Extraction Architecture

- Metadata is encouraged on start-up and available through a UI panel.
- Some fields are stored from previous sessions using a local xml file.
- Users can ignore the panel and fields are left blank.



SAFE: Saving Data



- Free text field, max 500 chars.
- Terms are sent to a MYSQL database via cURL.
- Local copies can be stored in XML format.
- Audio needs to be playing in order to send data.

SAFE: Loading Data

- An incentive for plug-in users to provide data is the load descriptor functionality.
- A descriptor-list is updated in near real-time from the database.

SEMANTIC AUDIO.CO.UK

SAFE: Code

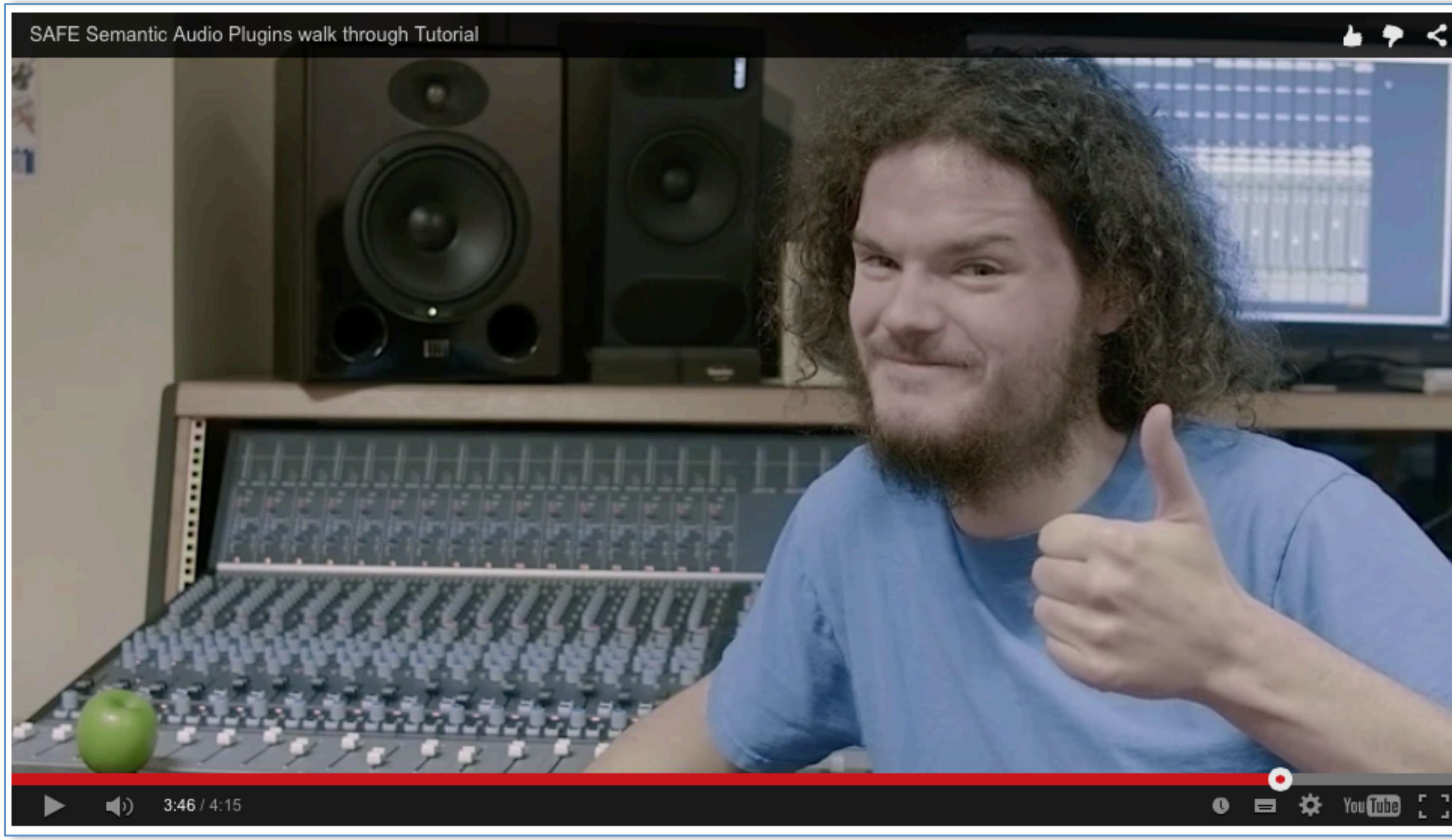
- All Plug-ins are written in JUCE, available in VST & AU formats.
- Dedicated SAFE JUCE module, available at: https://github.com/semanticaudio

- SAFE_juce_module
  - PluginUtils
    - CurlHolder.h
    - LibXtractHolder.cpp
    - LibXtractHolder.h
    - SAFEAudioProcessor.cpp
    - SAFEAudioProcessor.h
    - SAFEAudioProcessorEditor.cpp
    - SAFEAudioProcessorEditor.h
    - SAFEFeatureExtractor.cpp
    - SAFEFeatureExtractor.h
    - SAFEParameter.cpp
    - SAFEParameter.h
    - SAFEWarnings.h
  - Filters
    - AllPassFilter.cpp
    - AllPassFilter.h
    - BrechtsIIRFilter.cpp
    - BrechtsIIRFilter.h
    - QuadratureFilter.cpp
    - QuadratureFilter.h
    - Resampler.cpp
    - Resampler.h

- UIComponents
  - DynamicProcessorGraph.cpp
  - DynamicProcessorGraph.h
  - FilterGraph.cpp
  - FilterGraph.h
  - FilterInfo.cpp
  - FilterInfo.h
  - FunctionGraph.cpp
  - FunctionGraph.h
  - LevelMeter.cpp
  - LevelMeter.h
  - LevelMeterTimer.cpp
  - LevelMeterTimer.h
  - SAFEButton.cpp
  - SAFEButton.h
  - SAFEDescriptorLoadScreen.cpp
  - SAFEDescriptorLoadScreen.h
  - SAFEMetaDataScreen.cpp
  - SAFEMetaDataScreen.h
  - SAFESlider.cpp
  - SAFESlider.h

- LookAndFeel
  - SAFEColours.cpp
  - SAFEColours.h
  - SAFEImages.cpp
  - SAFEImages.h
  - SAFELookAndFeel.cpp
  - SAFELookAndFeel.h
- Effects
  - MVerb.h
- juce_module_info
- SAFE_juce_module.h

SAFE: How To…



http://www.semanticaudio.co.uk/documentation/

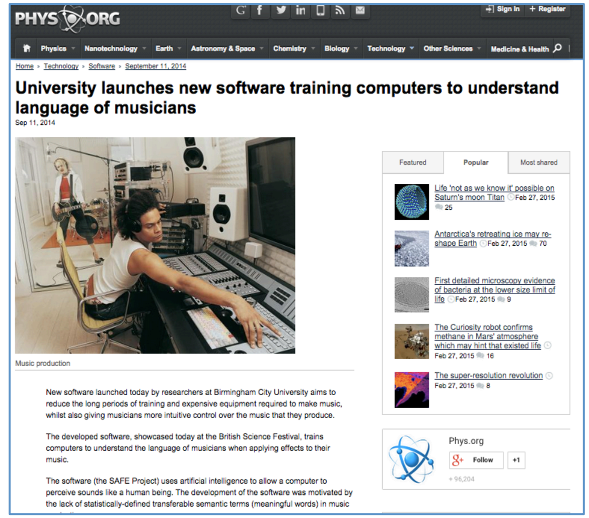
SAFE Plug-in Reception

The SAFE project teaches computers to understand your musical vocab

by Chris Velazco | @chrisvelazco | September 12th 2014 at 9:37 pm

The vocabulary we use to describe music can be tough enough for a human to grok (really, what does it mean when a guitar riff is "crunchy"?) but a team of tinkerers from Birmingham City University aren't interested in helping people understand that language. Nope -- instead, they've cooked up a way to teach your computer what you mean when you throw around words like "bright" or "fuzzy" or, yes, "crunchy" with a program they call the SAFE Project.

Spearheaded by Dr. Ryan Stables, the SAFE Project in its current form is a plugin for existing audio workstation software that lets would-be music producers apply effects by



University launches new software training computers to understand language of musicians

Sep 11, 2014

Music production

New software launched today by researchers at Birmingham City University aims to reduce the long periods of training and expensive equipment required to make music, whilst also giving musicians more intuitive control over the music that they produce.

The developed software, showcased today at the British Science Festival, trains computers to understand the language of musicians when applying effects to their music.

The software (the SAFE Project) uses artificial intelligence to allow a computer to perceive sounds like a human being. The development of the software was motivated by the lack of statistically-defined transferable semantic terms (meaningful words) in music

THE IRISH TIMES

So you want to be a music producer? Just switch on your laptop

Free software puts artificial intelligence in the mix for music producers

Researchers at Birmingham City University are able to turn the complexity of a mixing board with its hundreds of dials and settings into a simple piece of software that takes minutes to install and understand. Photograph: Birmingham City University

Michel Destrade

Wed, Sep 10, 2014, 01:02

Budding and confirmed musicians alike will soon be able to produce professional-sounding recordings of their work using a piece of free software presented yesterday at the British Science Festival in Birmingham.

BBC NEWS SCIENCE & ENVIRONMENT

10 September 2014 Last updated at 18:06

Training computers to understand the language of music

By Michael Eyre

Science Reporter, Birmingham

Researchers could make life easier for music producers

We often describe songs using terms like "warm" and "dreamy" - but do these words mean anything to a computer?

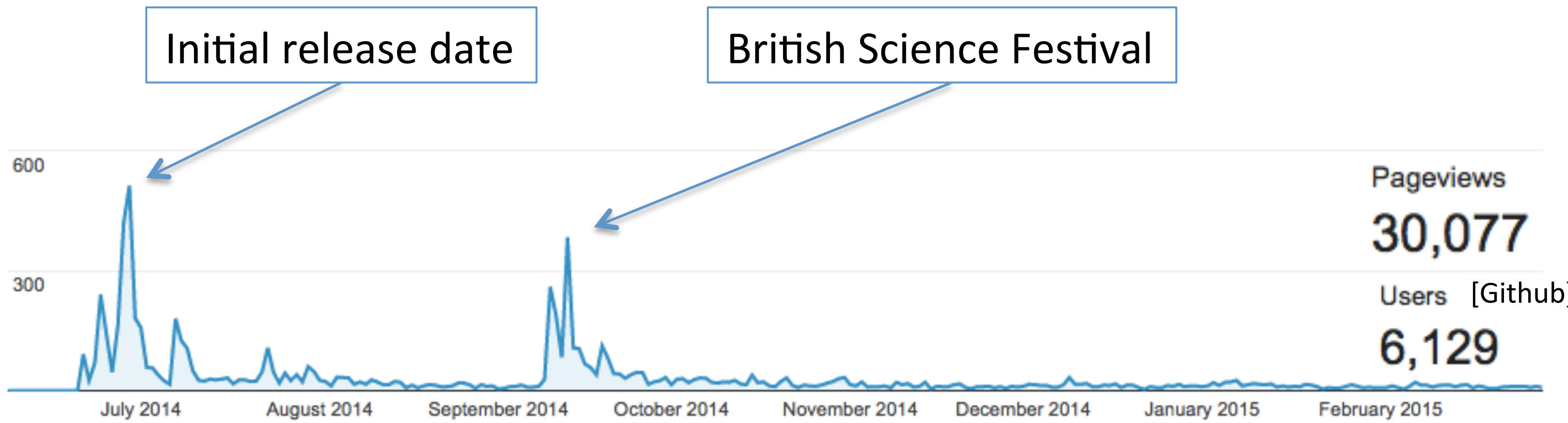
SAFE Plug-in Reception

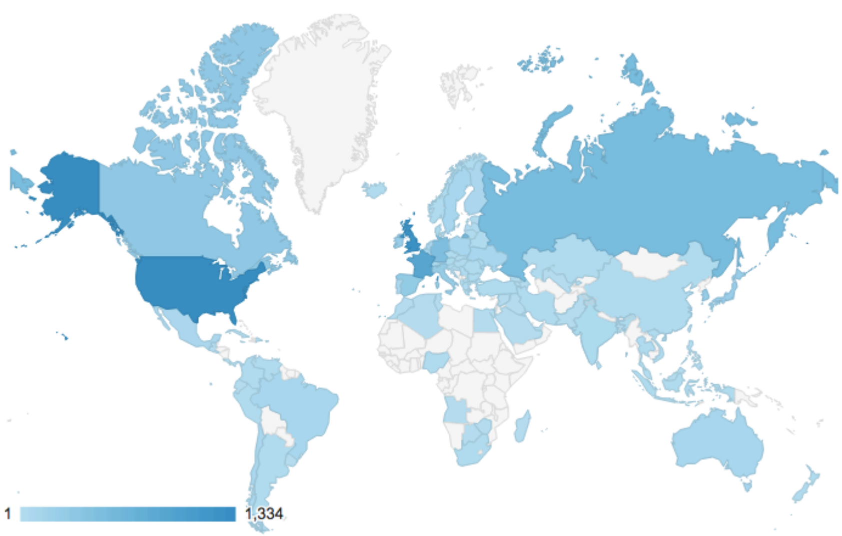
"I deleted your plugins. They cant do anything that any stock-plugin from any DAW couldnt (sic) do better. Another compressor, eq, reverb and so on ... yawn." – **Anonymous Gearslutz user**

The acronym is terrible, it sounds like an anti-bullying campaign – **M. Destrade, Irish Times.**

This idea is just as ridiculous as the idea of DJs being redundant is ridiculous. - **Anonymous Engadget commenter**

Plug-in Reception





	Country	Sessions	% Sessions
1.	United States	1,334	16.65%
2.	United Kingdom	1,244	15.53%
3.	France	888	11.08%
4.	Russia	541	6.75%
5.	Germany	465	5.80%
6.	Canada	305	3.81%
7.	Spain	290	3.62%
8.	South Korea	272	3.40%
9.	Italy	263	3.28%
10.	Japan	242	3.02%

SAFE: Word Frequency Analysis

Compressor
Unique terms: **137**
syllables-per-term: **1.64**
term length: **6.03 chars**
Top 3: **Punchy, Smooth, Vocal**

Word	Occurrences	Frequency	Rank
punchy	15	5.9%	1
smooth	14	5.5%	2
vocal	13	5.1%	3
nice	12	4.7%	4
squashed	6	2.4%	5
warm	6	2.4%	5
soft	5	2%	6
comp	5	2%	6
crushed	5	2%	6
master	4	1.6%	7

Overdrive
Unique terms: **106**
syllables-per-term: **1.63**
term length: **6.31 chars**
Top 3: **Warm, Crunch, Crunchy**

Word	Occurrences	Frequency	
warm	18	9.7%	1
crunch	9	4.9%	2
crunchy	7	3.8%	3
fuzz	7	3.8%	3
fuzzy	6	3.2%	4
raspy	4	2.2%	5
cha	3	1.6%	6
harsh	3	1.6%	6
bass	3	1.6%	6
smooth	3	1.6%	6

SAFE: Word Frequency Analysis

Equaliser

Unique terms: **188**

syllables-per-term: **1.48**

term length: **6.55** chars

Top 3: **Warm, Bright, Clear**

Word	Occurrences	Frequency	Rank
warm	48	11%	1
bright	38	8.7%	2
clear	11	2.5%	3
thin	10	2.3%	4
boomy	8	1.8%	5
clean	8	1.8%	5
and	8	1.8%	5
airy	6	1.4%	6
tinny	6	1.4%	6
	6	1.4%	6

Reverb

Unique terms: **117**

syllables-per-term: **1.73**

term length: **5.60** chars

Top 3: **Room, Small, Hall**

Word	Occurrences	Frequency	Rank
room	14	5.9%	1
small	7	3%	2
hall	7	3%	2
subtle	7	3%	2
dreamy	6	2.5%	3
natural	6	2.5%	3
drum	6	2.5%	3
damp	6	2.5%	3
echoy	5	2.1%	4
roomy	5	2.1%	4

SAFE: Word Frequency Analysis

Word	Occurrences	Frequency	Rank
warm	45	5.5%	1
bright	33	4.1%	2
smooth	17	2.1%	3
room	17	2.1%	3
punchy	15	1.8%	4
vocal	15	1.8%	4
nice	15	1.8%	4
subtle	11	1.4%	5
clear	10	1.2%	6
clean	9	1.1%	7

<u>All Plug-ins</u>

Unique terms: **362** syllables-per-term: **1.65**

Total terms: **848** Top 3: **Warm, Bright, Smooth**

SAFE: Term Highlights

Compressor

- “tickled”, “phat”, “kissing”, “die hard”
- “you need to write something in the box, fool!”

Distortion

- “beastly”, “destroyed”, “underwater”, “menacing”
- “dramatic tangs”, “el cruncho”, “caramel”

SAFE: Term Highlights

<u>Reverb</u>

- “DP-SAR-1”, “DIDI 1”, “fluffy”, “creamy”,
- “wholesome”, “150215”, “J”

<u>Equaliser</u>

- “bastard”, “drum 1”, “drum 2”, “drum 3”
- “pianombient”, “sticks out in the mix a bit more”
- “we call it bass”

Experiments in:
Live Sound

Session Analytics

- Collaboration with Music Group Research (Midas, Klark-Technik, Behringer, etc)
- Data captured in real-time from a modified **Midas** console with **SAFE** architecture.

Session Analytics

Midas Pro2C

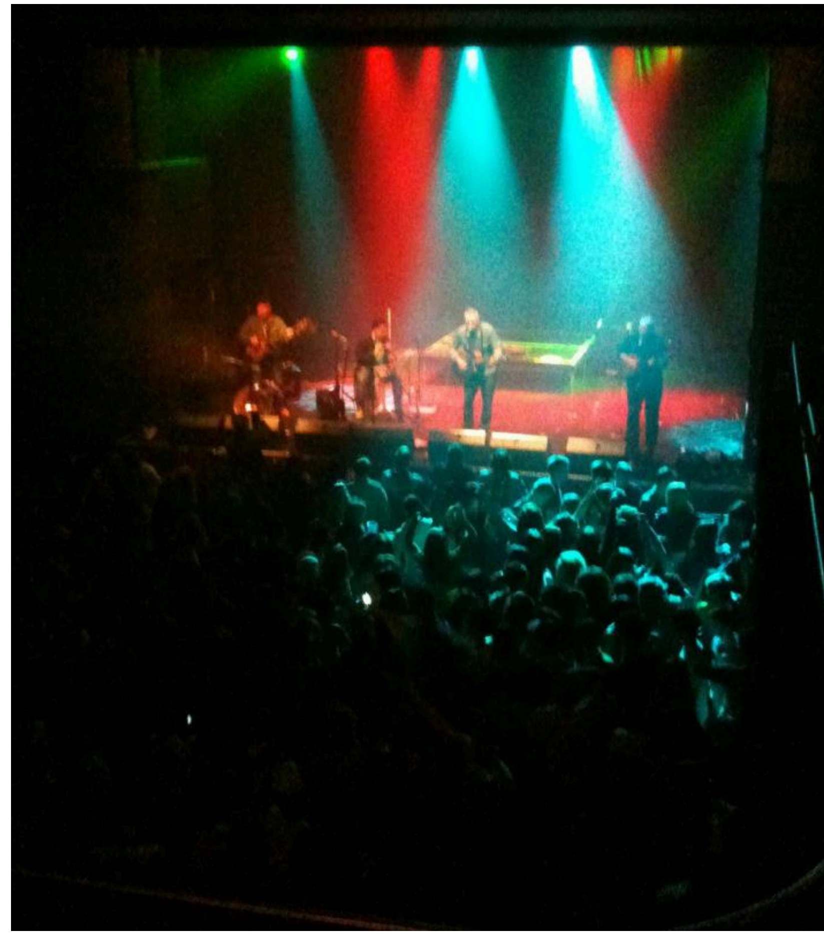
- 56 mic/line-ins
- Assignable faders and groups.
- 12 multichannel FX engines
- Dedicated effects parameter space with digital display.

Modified to print session data to console and receive state commands.

Data

- 15 live UK music festival recordings
- 32 channels per each recording
- All bands from a given stage across a full day of music
- Recordings include preliminary material and sound checks

- 30+ Engineers asked to mix as if they were at the festival



Objectives

1. **Automatic metadata derivation**
   - Via console-parameter analysis
   - Via audio-feature analysis
2. **Workflow optimization**
   - Provide intuitive interface abstractions

1. Metadata Derivation

This can be derived from:

- parameter information
- Audio feature data

Instruments/ Genre/ etc…

Engineer background/

Microphones being tested

End of Soudcheck

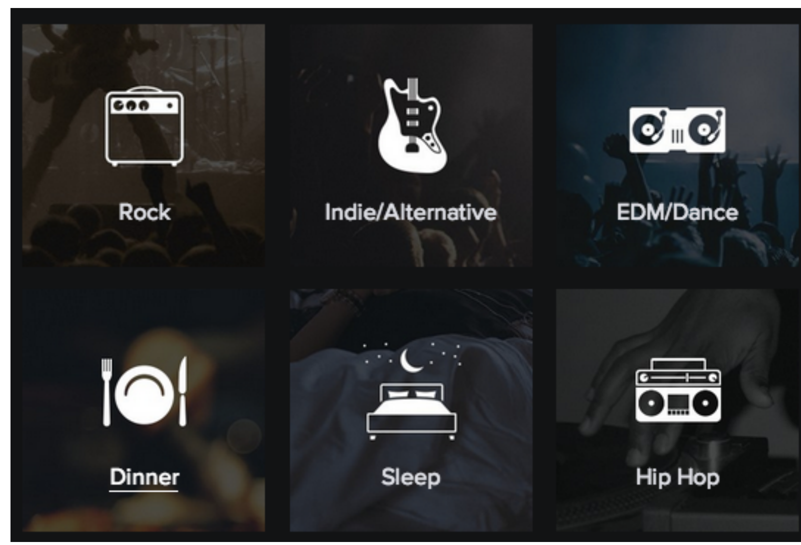
Feedback caused By microphone 1

Compression applied

Metadata Derivation

- Using machine learning, we are able to estimate instrument classes:
  - Essid et al. (2006): SVM, octave-band intensities [93%]
  - Tjoa & Liu (2010): NMF-derived temporal features and SVM [92%]
- And genre tags…
  - Tzanetakis & Cook (2002): Low/high level audio features with GMM/K-NN [61%]
  - Ezzaidi & Rouat (2006): MFCCs with GMMs [73-99%]

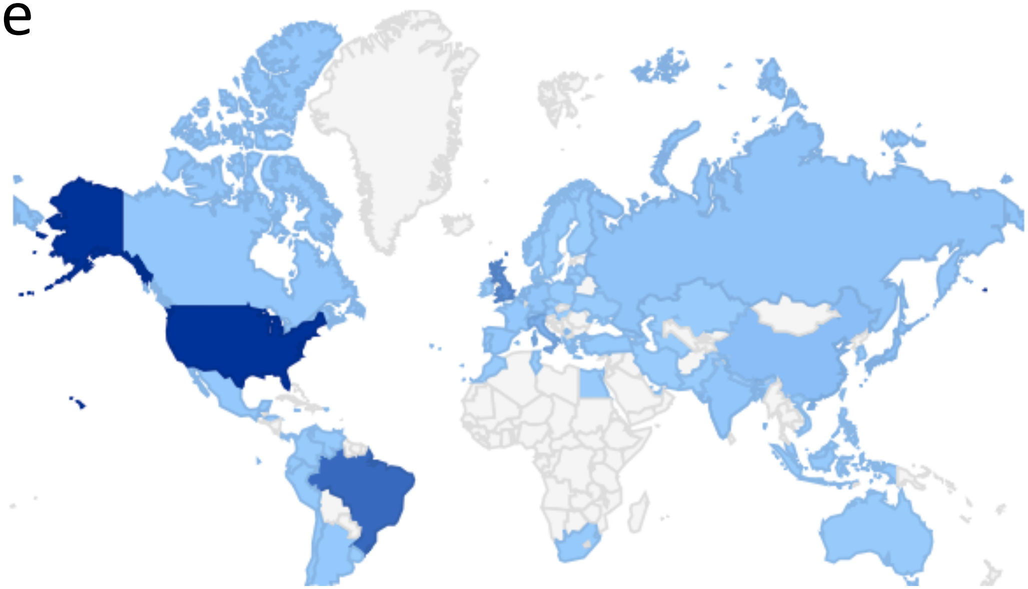
Genre, **Instrument**, Current Location, Language, Experience, Age



Metadata Derivation

- It’s relatively easy to estimate a user’s location based on their IP address, and the physical location of their ISP.
- Language is slightly more complex, but can be estimated using a language model and IP.

Genre, Instrument, **Current Location**, Language, Experience, Age



SEMANTIC AUDIO.CO.UK

Metadata Derivation

- Production experience classifiers are less common.
  - This information is useful for: weighting descriptor terms, partitioning data, user analytics.
  - We need to arrange specific experiments due to anonymous data.
  - We can use both audio feature data and parameter-space data

Genre, Instrument, Current Location, Language, **Experience,** Age

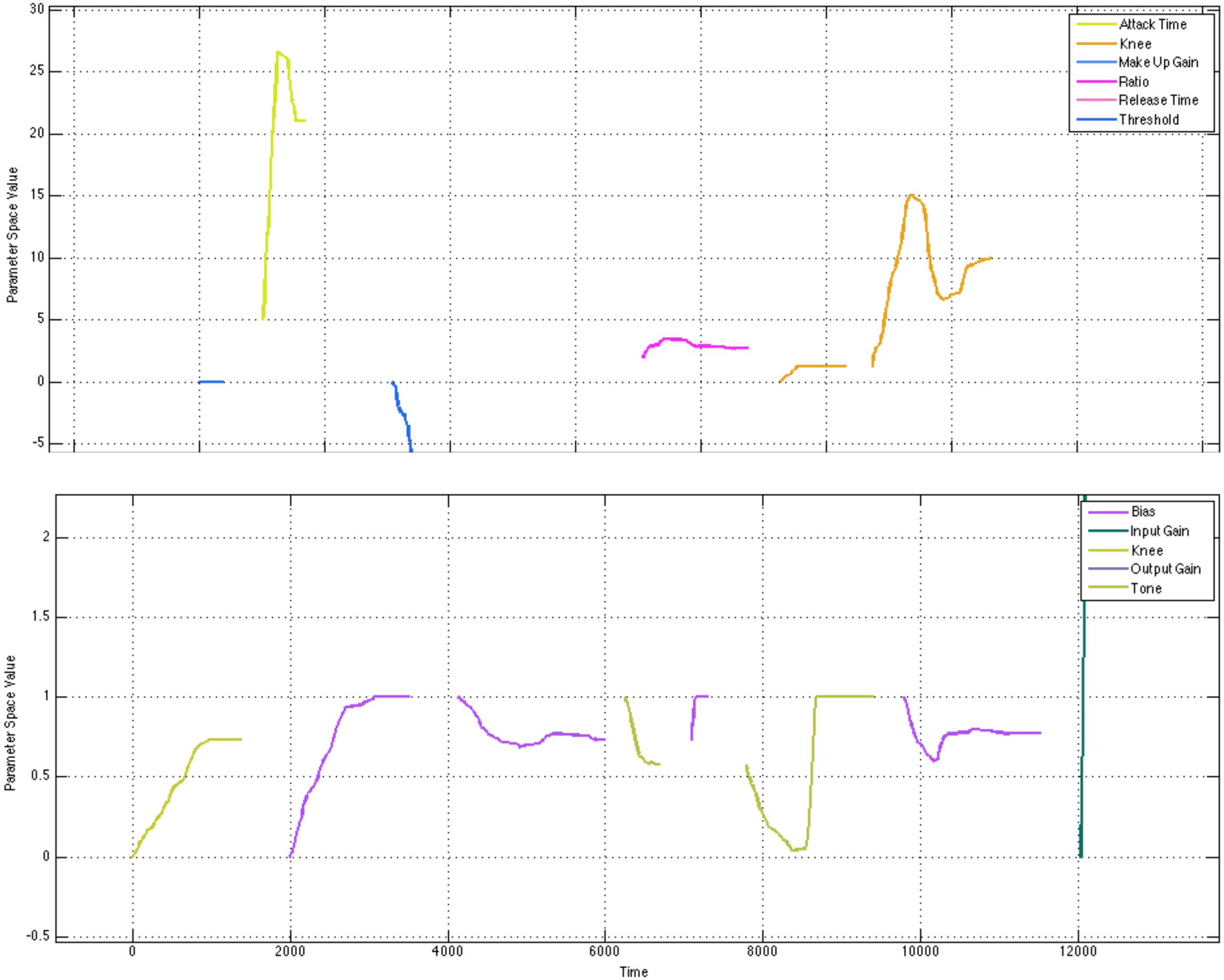
Parameter Tracking


```
1417782569678, 2, "Threshold", -32.04, 1
1417782569742, 2, "Threshold", -32.24, 1
1417782569854, 2, "Threshold", -32.44, 1
1417782569867, 2, "Threshold", -32.64, 1
1417782569887, 2, "Threshold", -32.85, 1
1417782569905, 2, "Threshold", -33.06, 1
1417782569949, 2, "Threshold", -33.27, 1
1417782570014, 2, "Threshold", -33.41, 1
1417782570040, 2, "Threshold", -33.62, 1
1417782570057, 2, "Threshold", -33.83, 1
1417782570158, 2, "Threshold", -34.04, 1
1417782570870, 2, "Threshold", -34.24, 1
1417782570894, 2, "Threshold", -34.44, 1
1417782570913, 2, "Threshold", -34.65, 1
1417782570954, 2, "Threshold", -34.86, 1
1417782570965, 2, "Threshold", -35.06, 1
1417782571271, 2, "Threshold", -35.06, 2

1417782572021, 2, "Ratio", 2, 0
1417782572033, 2, "Ratio", 2.02, 1
1417782572055, 2, "Ratio", 2.17, 1
1417782572073, 2, "Ratio", 2.44, 1
1417782572095, 2, "Ratio", 2.71, 1
1417782572106, 2, "Ratio", 2.86, 1
1417782572129, 2, "Ratio", 3.01, 1
1417782572140, 2, "Ratio", 3.22, 1
1417782572160, 2, "Ratio", 3.33, 1
1417782572179, 2, "Ratio", 3.5, 1
1417782572188, 2, "Ratio", 3.66, 1
```

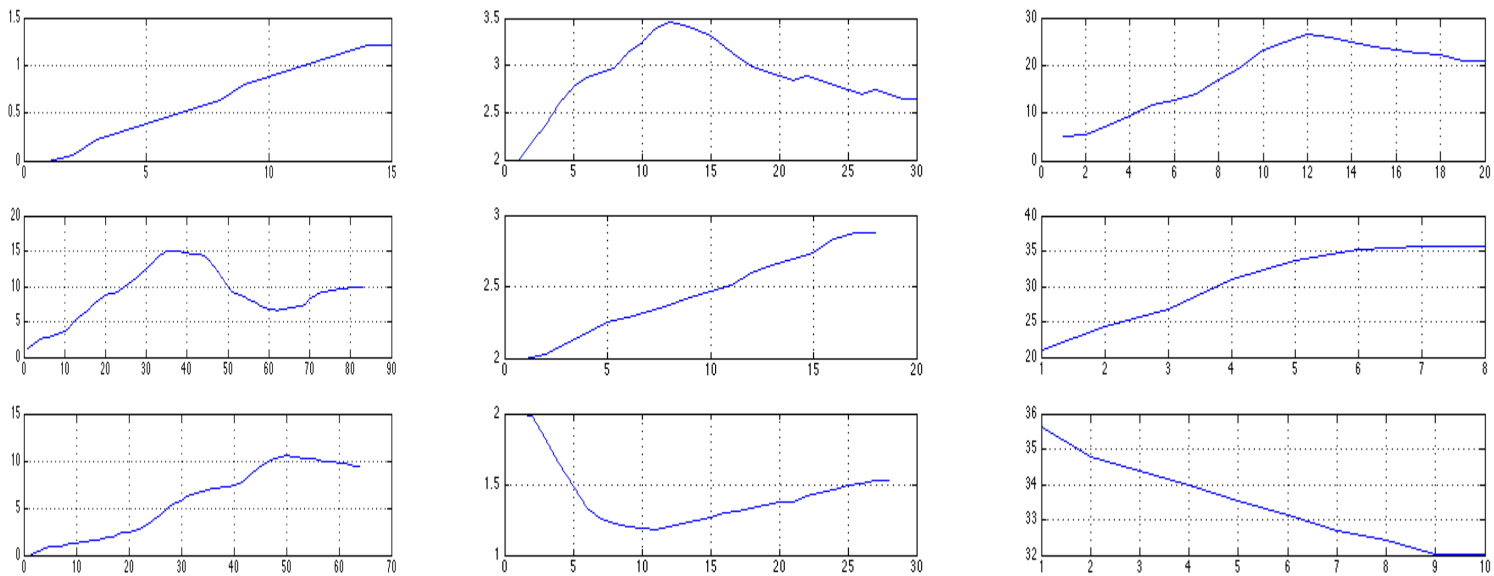

- Parameter modulation data was gathered during a series of live-simulation production sessions.
- Subjects with varying production experience were given recorded mixing tasks to perform.
- Tasks were selected based on *creative* and *corrective* procedures.
- e.g. sec 1. make channel n warmer/brighter. Sec 2. route input channels to stereo bus.

Parameter Tracking



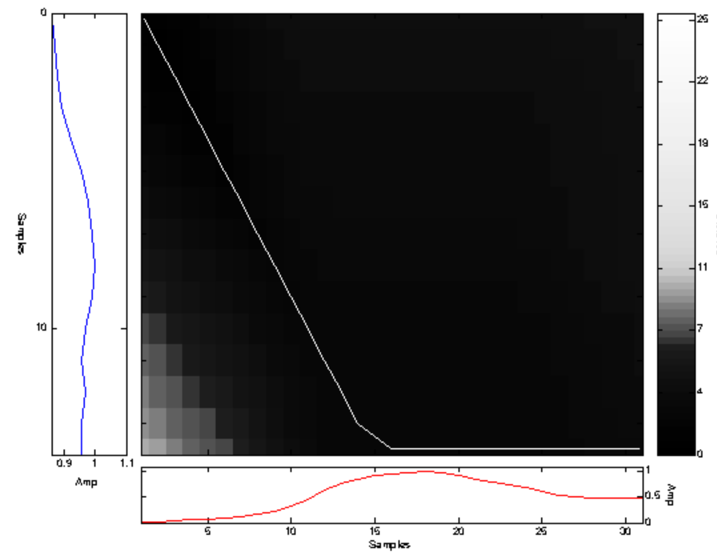
Modulation Patterns

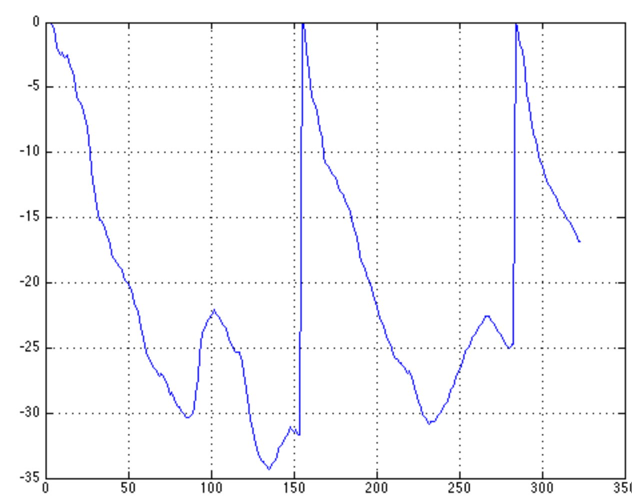
- Initial findings show range of common basis functions.
- These tend to be correlated with level of production (defined through listening tests).



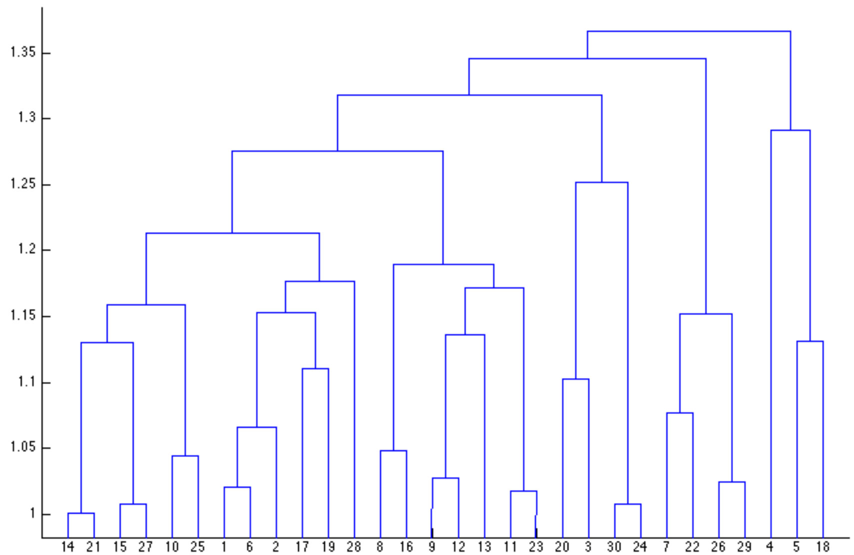
Dynamic Time Warping

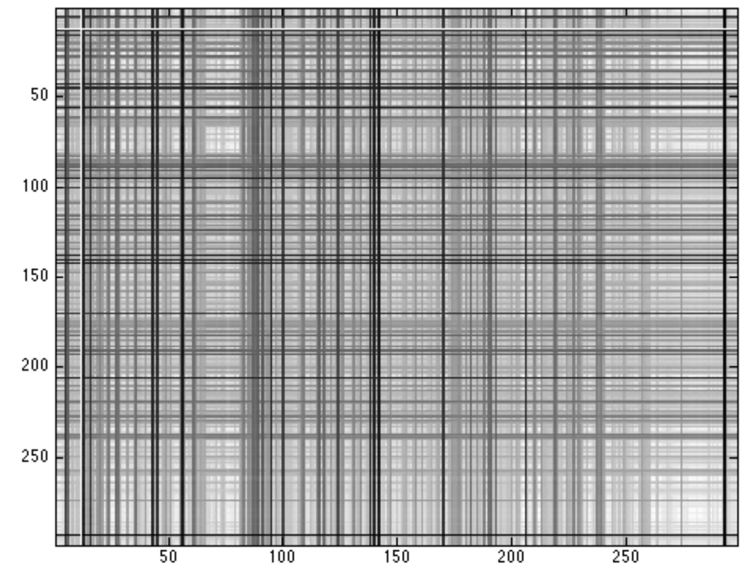
- We extract temporal features from the tracks, including modulation spectra features, and DTW-features.
- Various modulation periods are used:
  - single adjustments,
  - descriptor-wise adjustments
  - track-wise adjustments
- Clustering is then applied to identify regions of similarity





Clustering





- Agglomerative/hierarchical clustering applied to modulation patterns
- Clusters compared to labeled targets

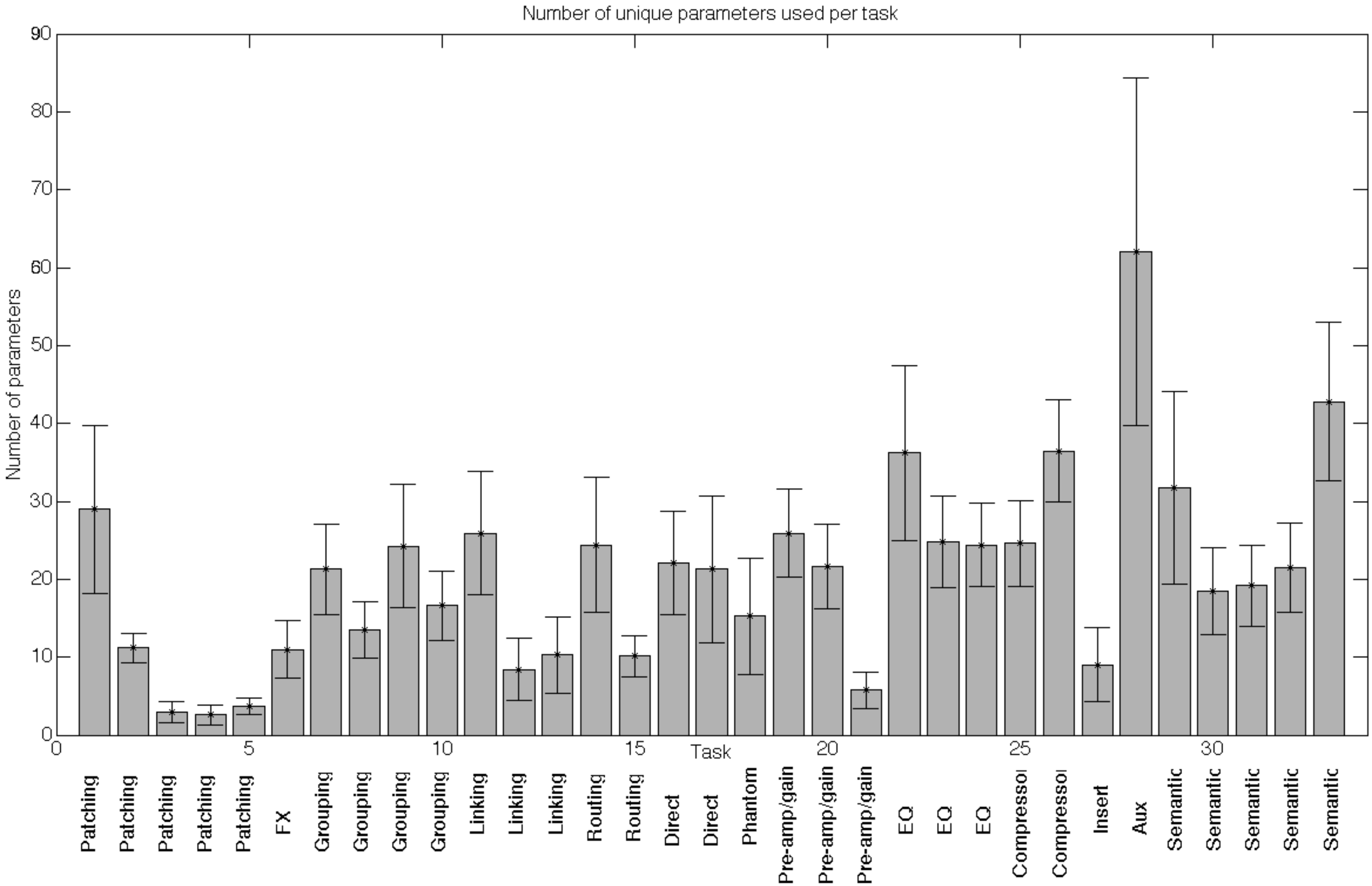
SEMANTIC AUDIO.co.uk

Workflow Optimization

How do engineers use the consoles in a live setting?

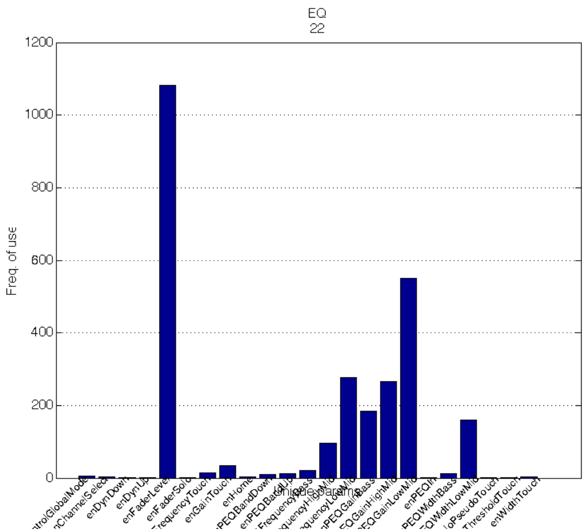
Using data can we develop novel/intuitive interfaces for live mixing?

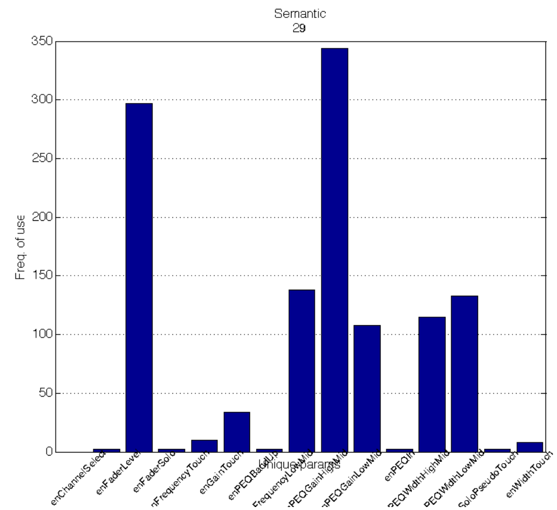
Workflow Optimization

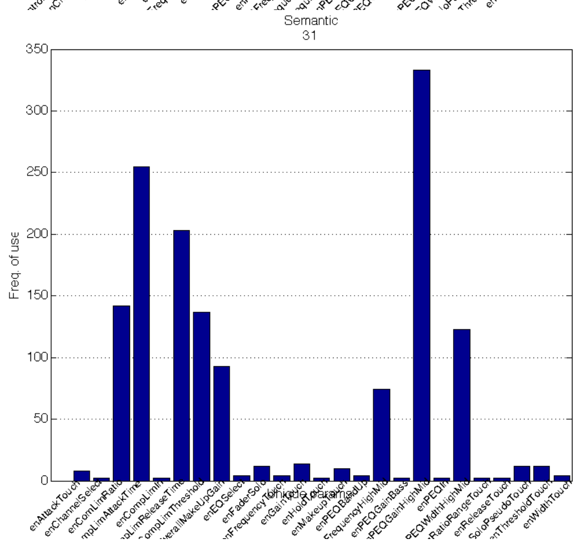


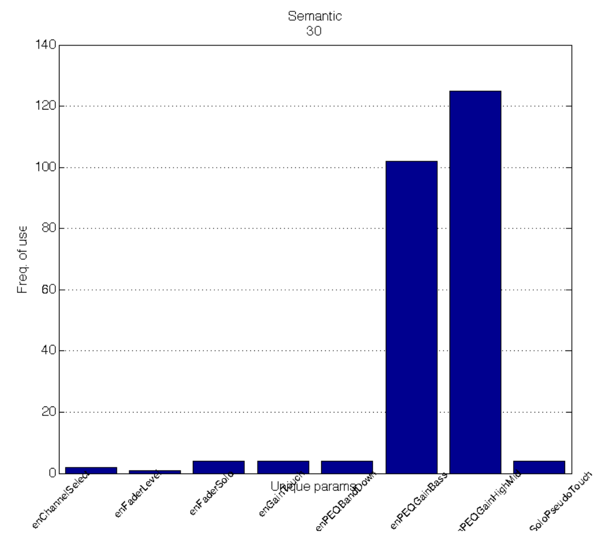
SEMANTIC AUDIO.CO.UK

Workflow Optimization





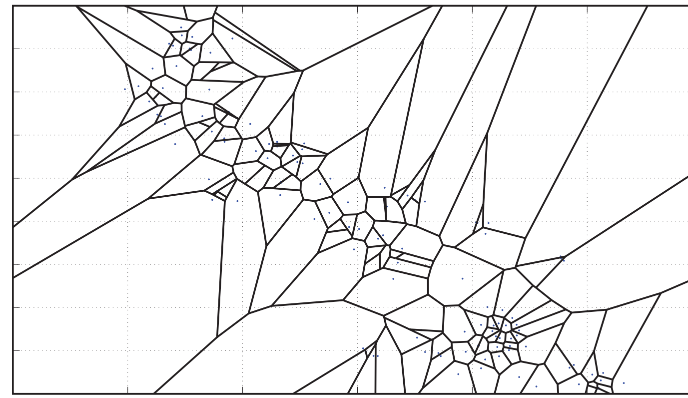


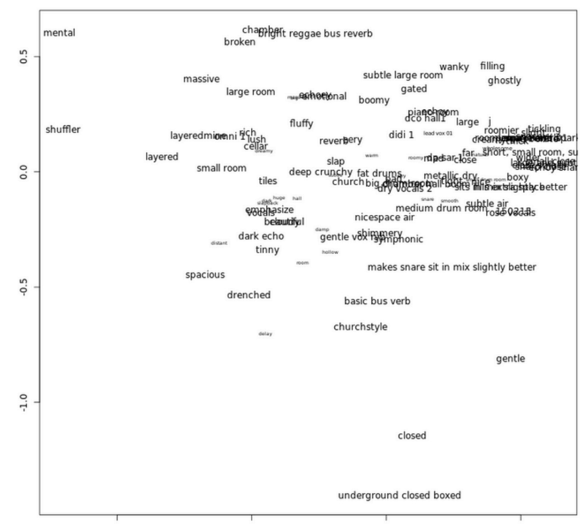


Workflow Optimization: Parameter Abstraction

- Can we control complex audio effects with low-dimensional spaces?
- Can we use both audio feature data and parameter space data.

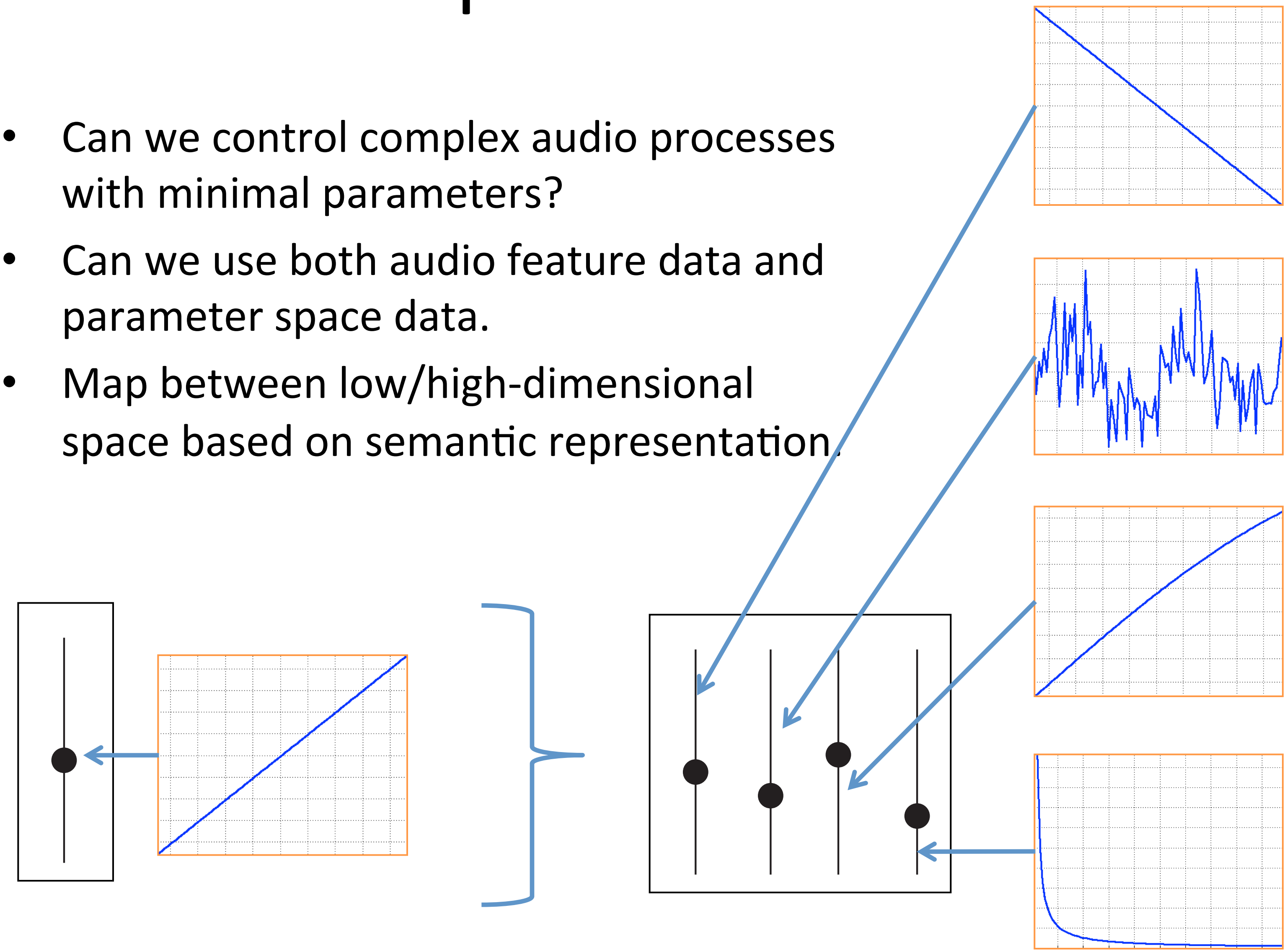
- Currently testing:
  - Dimensionality Reduction algorithms
    - *PCA, MDS, tSNE, etc*
  - Visualisation techniques
    - *Voronoi, Lloyd's relaxation*
  - Parameter space representation
    - Linear, n-D, object/path-based





Semantic Representation

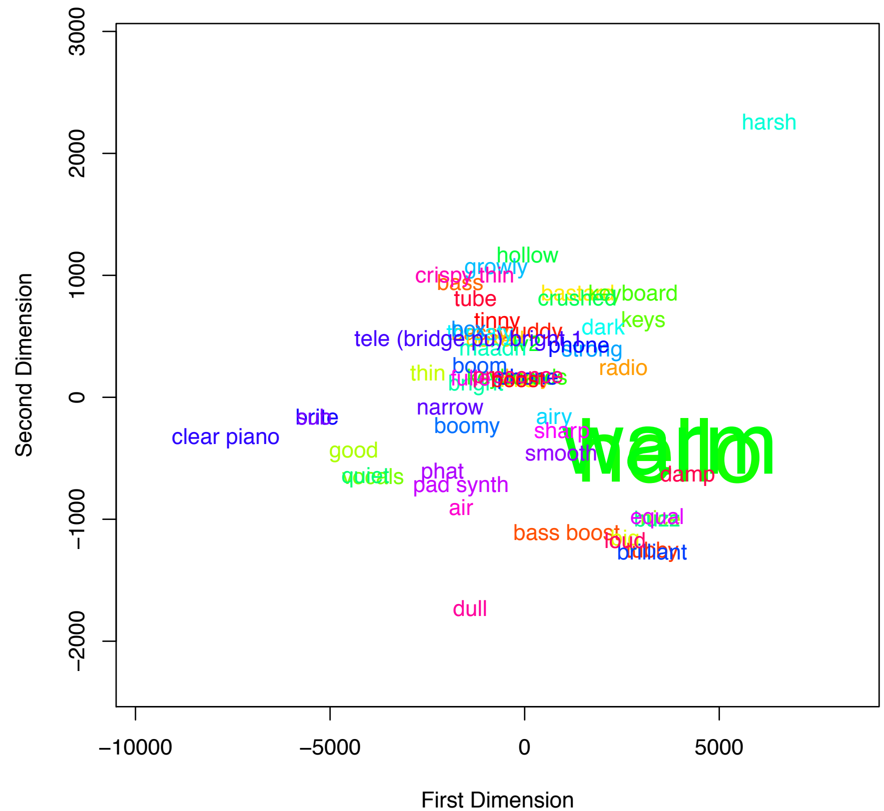
- Can we control complex audio processes with minimal parameters?
- Can we use both audio feature data and parameter space data.
- Map between low/high-dimensional space based on semantic representation.



Reduced Dimensionality Interfaces

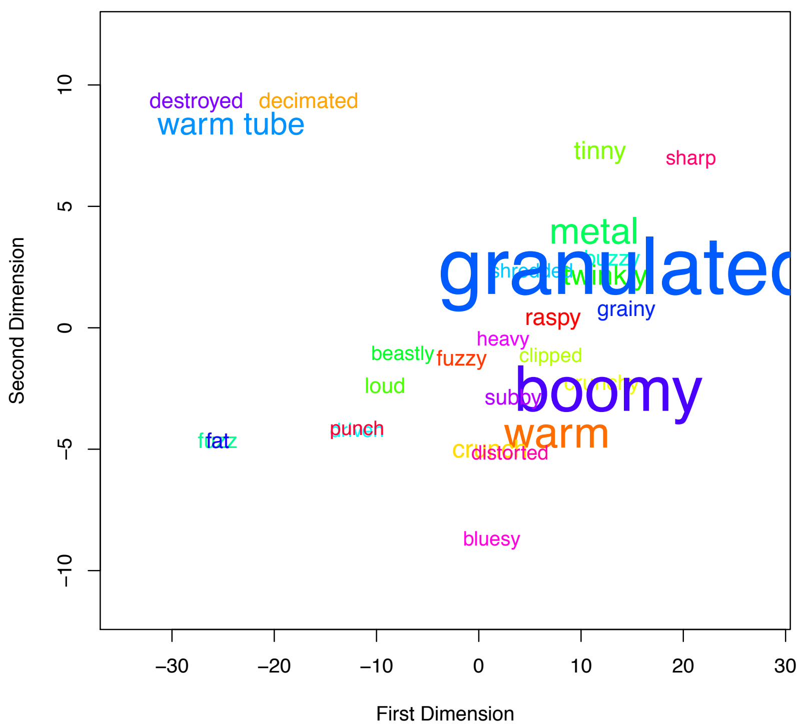
- Visualisations are available online
- Figures are updated in *near*-real-time using MDS (with the R package:).
- Two sources: params and audio features.
- Terms are placed on coordinates of 2-D map.
- Term size is determined by confidence score (sum of variance):

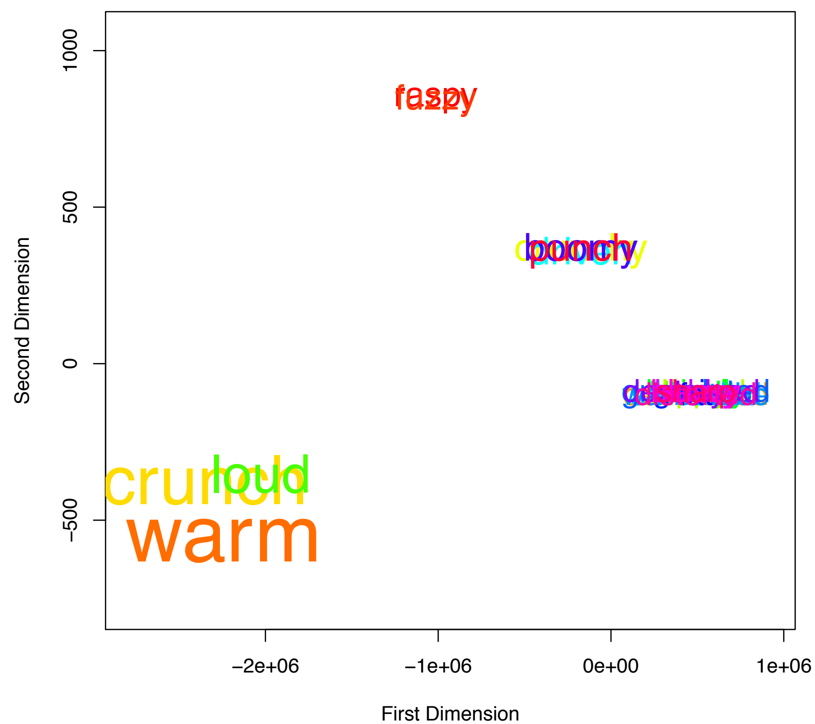
$$Conf = \frac{1}{N}\sum_{k=0}^{K-1}\sum_{n=0}^{N-1}(x_{k,n} - \mu_n)^2$$



Reduced Dimensionality Interfaces

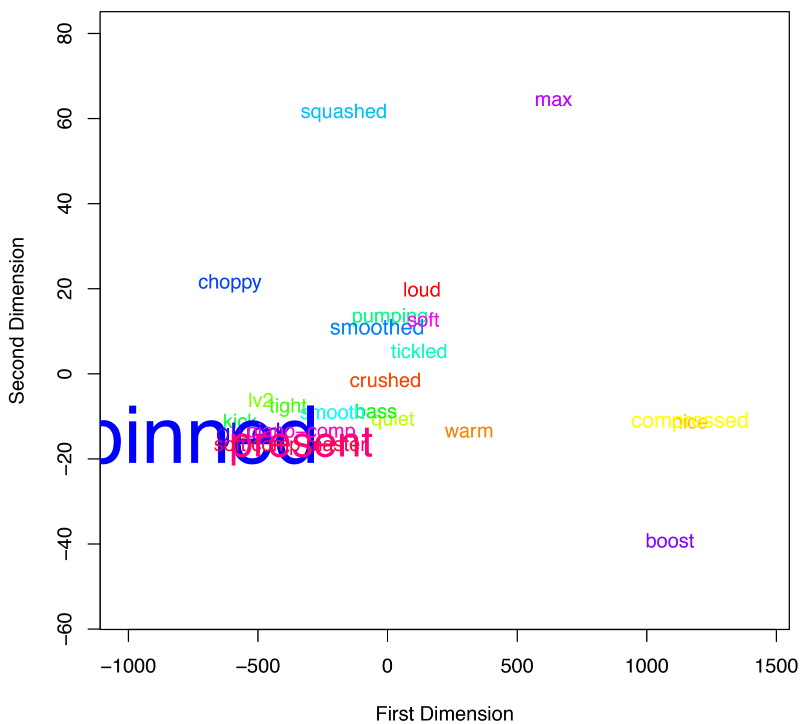
Distortion…

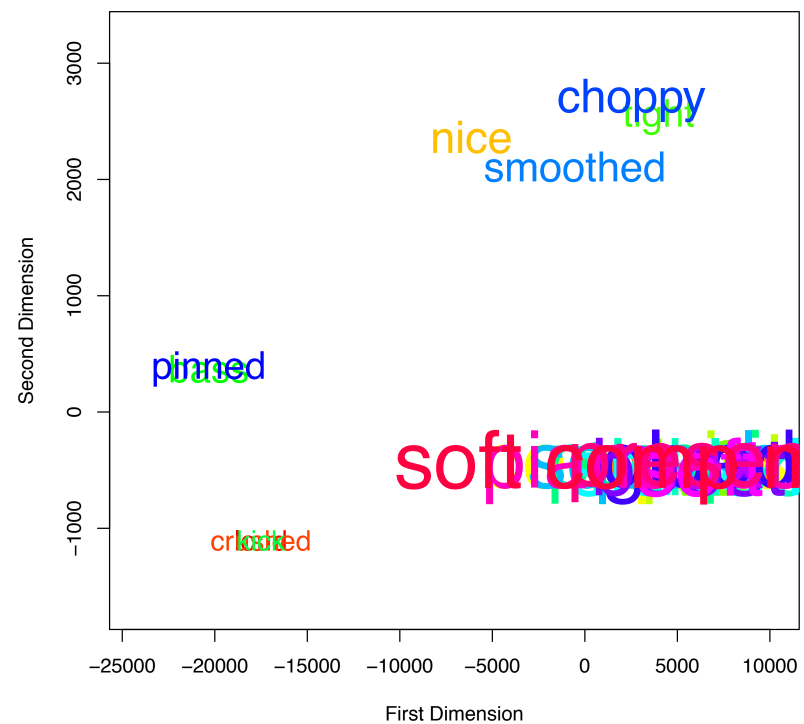




Reduced Dimensionality Interfaces

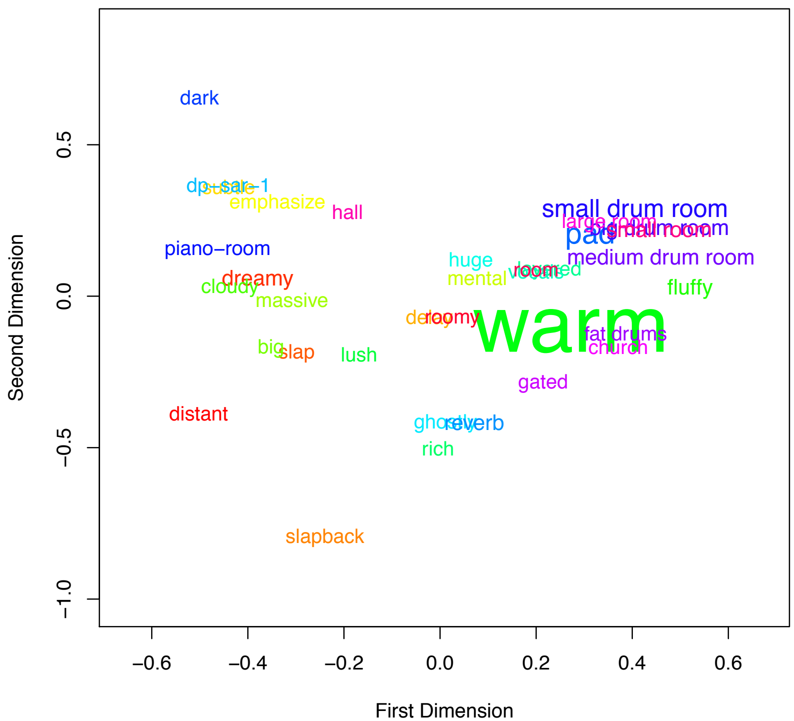
Compressor…

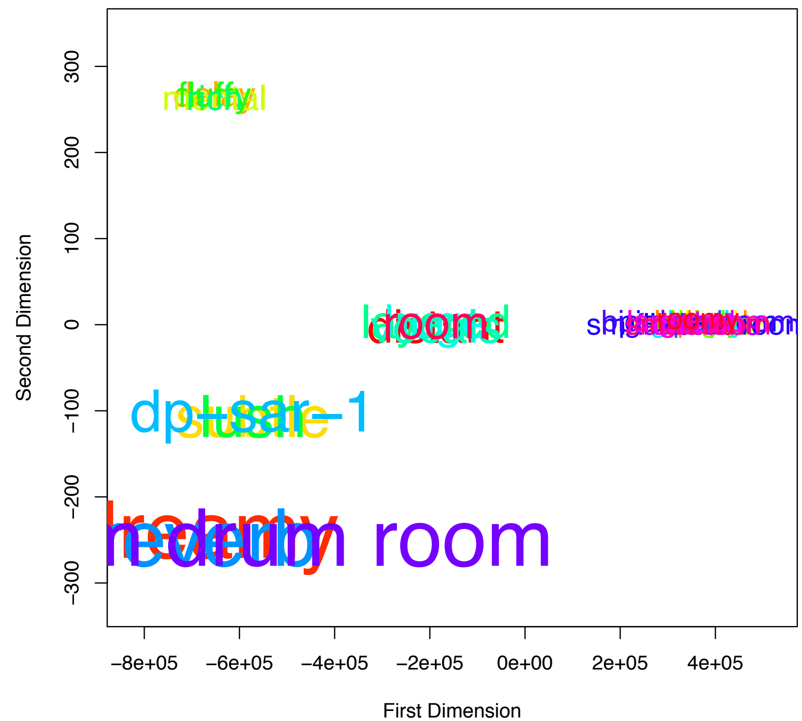




Reduced Dimensionality Interfaces

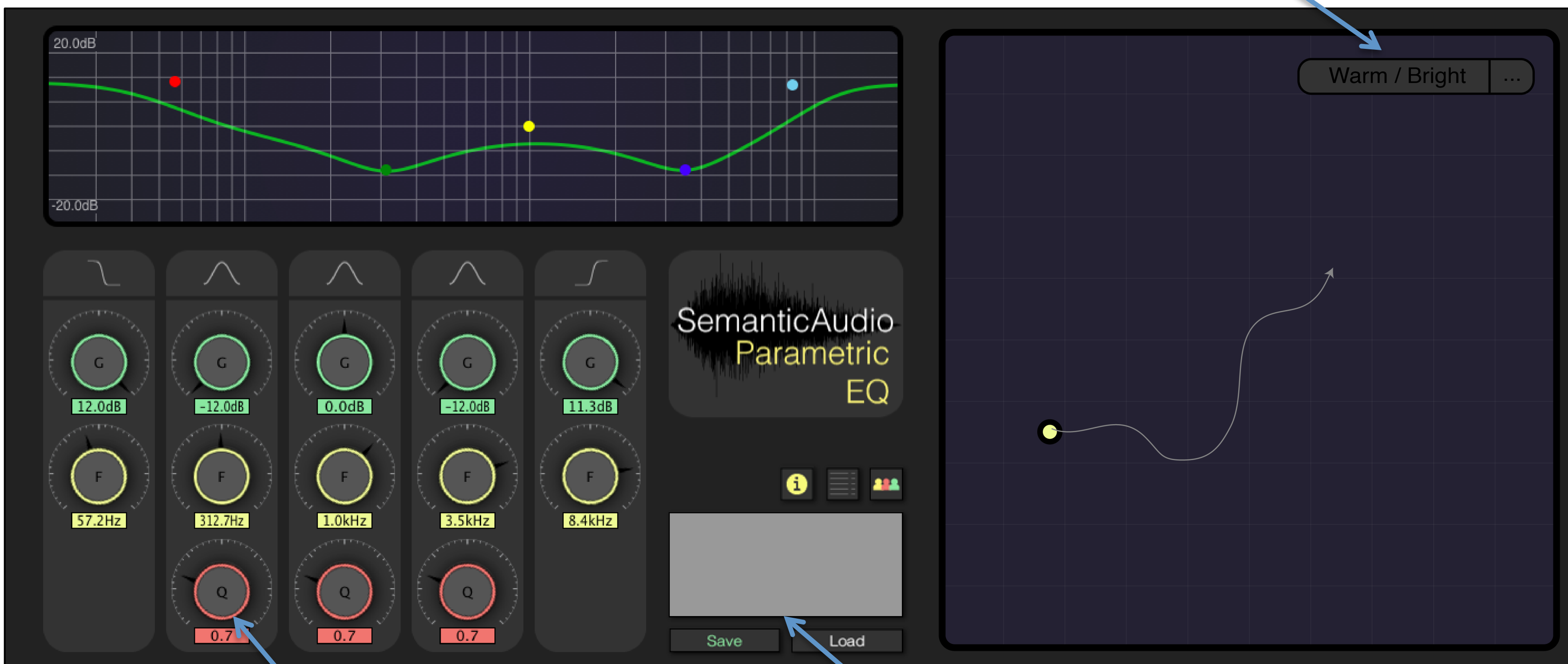
Reverb…





P2. PC-EQ-2D

Dimension assignment



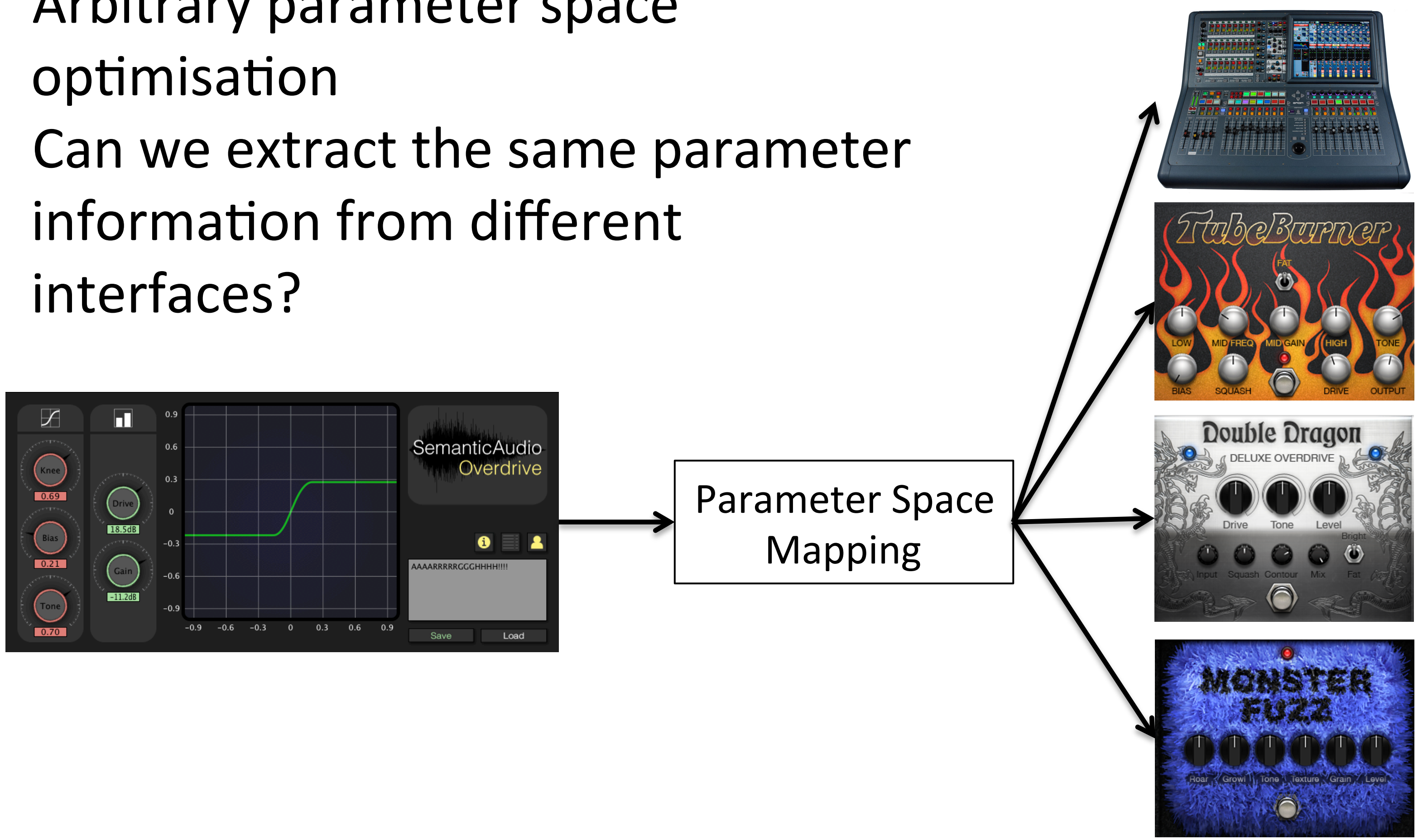
Real-time param modulation

Descriptor suggestions

The Future…

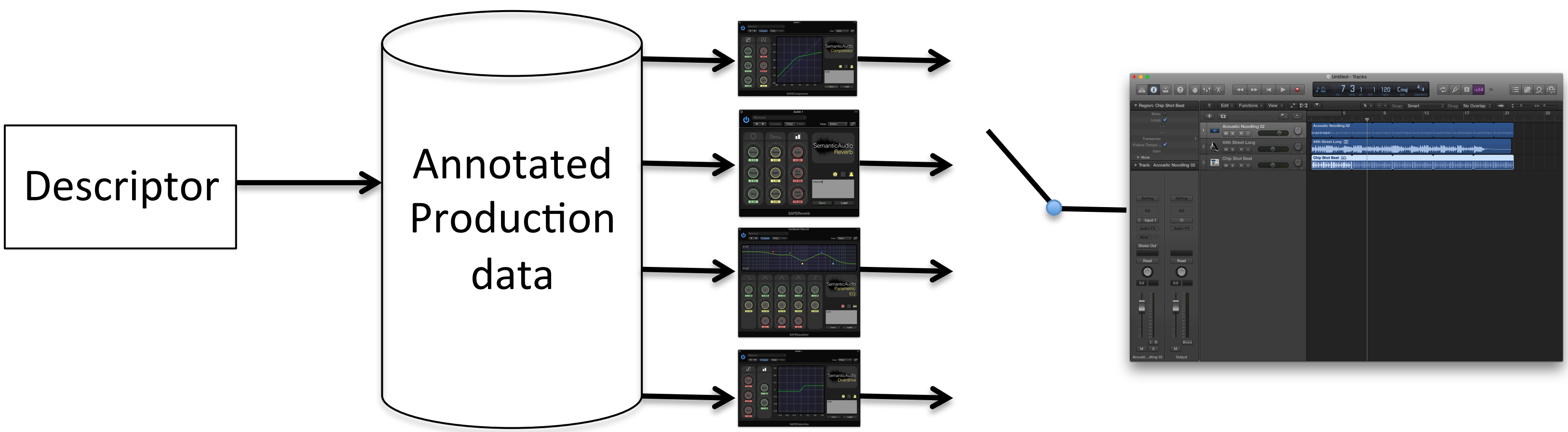
P3. Semantic API

- SAFE API:
  - Arbitrary parameter space optimisation
  - Can we extract the same parameter information from different interfaces?



P3. Interface Abstraction

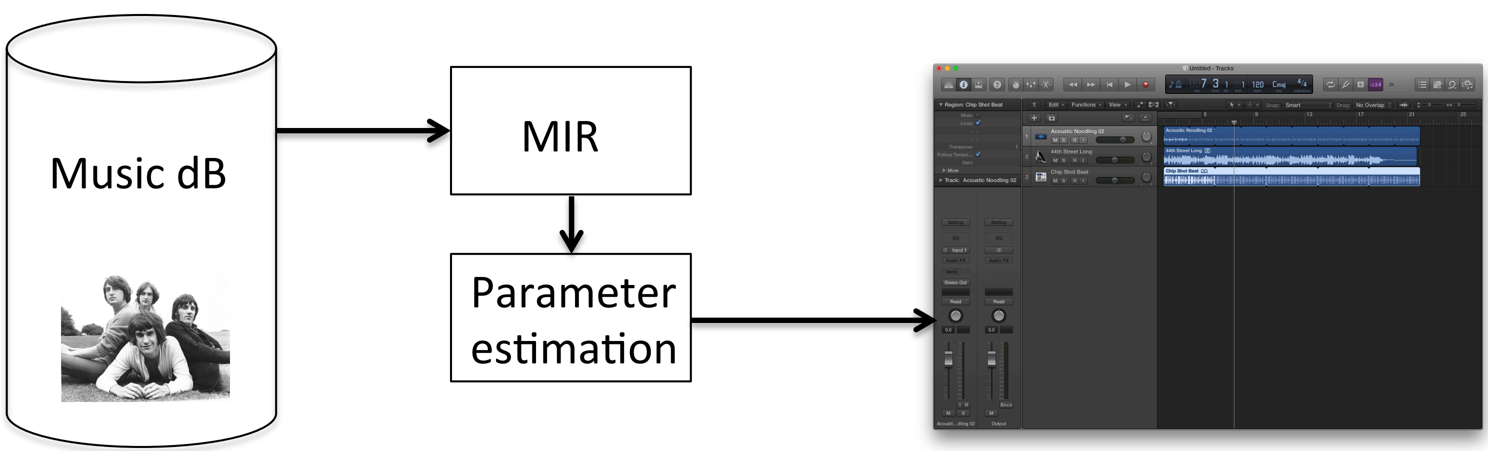
- Auto-loading plug-ins
  - what do I need to make something warm?
  - Can we load combinations of audio effects based on production data



P3. Content Reproduction

Content-based semantic music production:

- Parameter learning: *Make my guitar sound like Waterloo Sunset by The Kinks*
- Automatic remixing: *Make the Blowin' in the Wind sound like it was produced by Rick Rubin*



Anyway…

Please get involved!!!!

Plug-ins and data available from: www.semanticaudio.co.uk

ryan.stables@bcu.ac.uk

@otmiv

References

Dugan, D. "Automatic microphone mixing." Journal of the Audio Engineering Society 23.6 (1975): 442-449.

Campbell, E. and Whittemore R. T., "Automatic microphone mixing apparatus " in European Patent Office, US 4357492 (A) 1982.

Dannenberg, Roger B. "An intelligent multi-track audio editor." In Proceedings of International Computer Music Conference (ICMC), vol. 2, pp. 89-94. 2007.

Perez-Gonzalez, E., & Reiss, J. Automatic gain and fader control for live mixing. In Applications of Signal Processing to Audio and Acoustics, 2009. WASPAA'09. IEEE Workshop on (pp. 1-4). IEEE.

Ward, D., Reiss, J. D., & Athwal, C. (2012, October). Multitrack mixing using a model of loudness and partial loudness. In Audio Engineering Society Convention 133. Audio Engineering Society.

References

Scott, J., & Kim, Y. E. (2013, November). Instrument Identification Informed Multi-Track Mixing. In International Society for Music Information Retrieval (ISMIR 2013) (pp. 305-310).

De Man, B. & Reiss, J. D. (2013). A Semantic Approach To Autonomous Mixing. Journal of the Art of Record Production

References

Terrell, M., Reiss, J. D., & Sandler, M. (2010). Automatic noise gate settings for drum recordings containing bleed from secondary sources. EURASIP Journal on Advances in Signal Processing, 2010, 10.

Perez-Gonzalez, E., & Reiss, J. D. (2010). A real-time semiautonomous audio panning system for music mixing. EURASIP Journal on Advances in Signal Processing, 2010, 5.

Clifford, A., & Reiss, J. (2011). Proximity effect detection for directional microphones. In Audio Engineering Society Convention 131. Audio Engineering Society.

Braun, S. (2012) Evaluation of Various algorithms to Detect Acoustic Feedback. Master Thesis, IEM Graz, Austria.

References

Sabin, A. T., & Pardo, B. (2009, October). 2DEQ: an intuitive audio equalizer. In Proceedings of the seventh ACM conference on Creativity and cognition (pp. 435-436). ACM.

Cartwright, M. B., & Pardo, B. (2013, November). Social-EQ: Crowdsourcing an Equalization Descriptor Map. In ISMIR (pp. 395-400).

Seetharaman, P., & Pardo, B. (2014, November). Reverbalize: a crowdsourced reverberation controller. In Proceedings of the ACM International Conference on Multimedia (pp. 739-740). ACM.

Stables, R., Enderby, S., De Man, B., Fazekas, G., & Reiss, J. (2014, October). SAFE: A system for the extraction and retrieval of semantic audio descriptors. In 15th International Society for Music Information Retrieval Conference (ISMIR 2014).

Ma, Z., De Man, B., Pestana, P. D., Black, D. A., & Reiss, J. D. (2015). Intelligent Multitrack Dynamic Range Compression. Journal of the Audio Engineering Society, 63(6), 412-426.

References

Dattorro, J. (1997). Effect design, part 1: Reverberator and other filters. Journal of the Audio Engineering Society, 45(9), 660-684.

Reiss, J. D., & McPherson, A. (2014). Audio Effects: Theory, Implementation and Application. CRC Press.

Bullock, J. (2007). Libxtract: A lightweight library for audio feature extraction. In Proceedings of the International Computer Music Conference ICMC (Vol. 43).

Thanks…

- Matthew Cheshire, DMT Lab, BCU
- Brecht De Man, C4DM, QMUL
- Sean Enderby, DMT Lab, BCU
- Spyridon Stasis, DMT Lab, BCU

SAFE Developers

- Cham Athwal, DMT Lab, BCU
- George Fazekas, C4DM, QMUL
- Alessandro Palladini, Music Group Research.
- Josh Reiss, C4DM, QMUL
- Mark Sandler, C4DM, QMUL

Other Helpful People

- EPSRC Semantic Media Project (semanticmedia.org.uk)