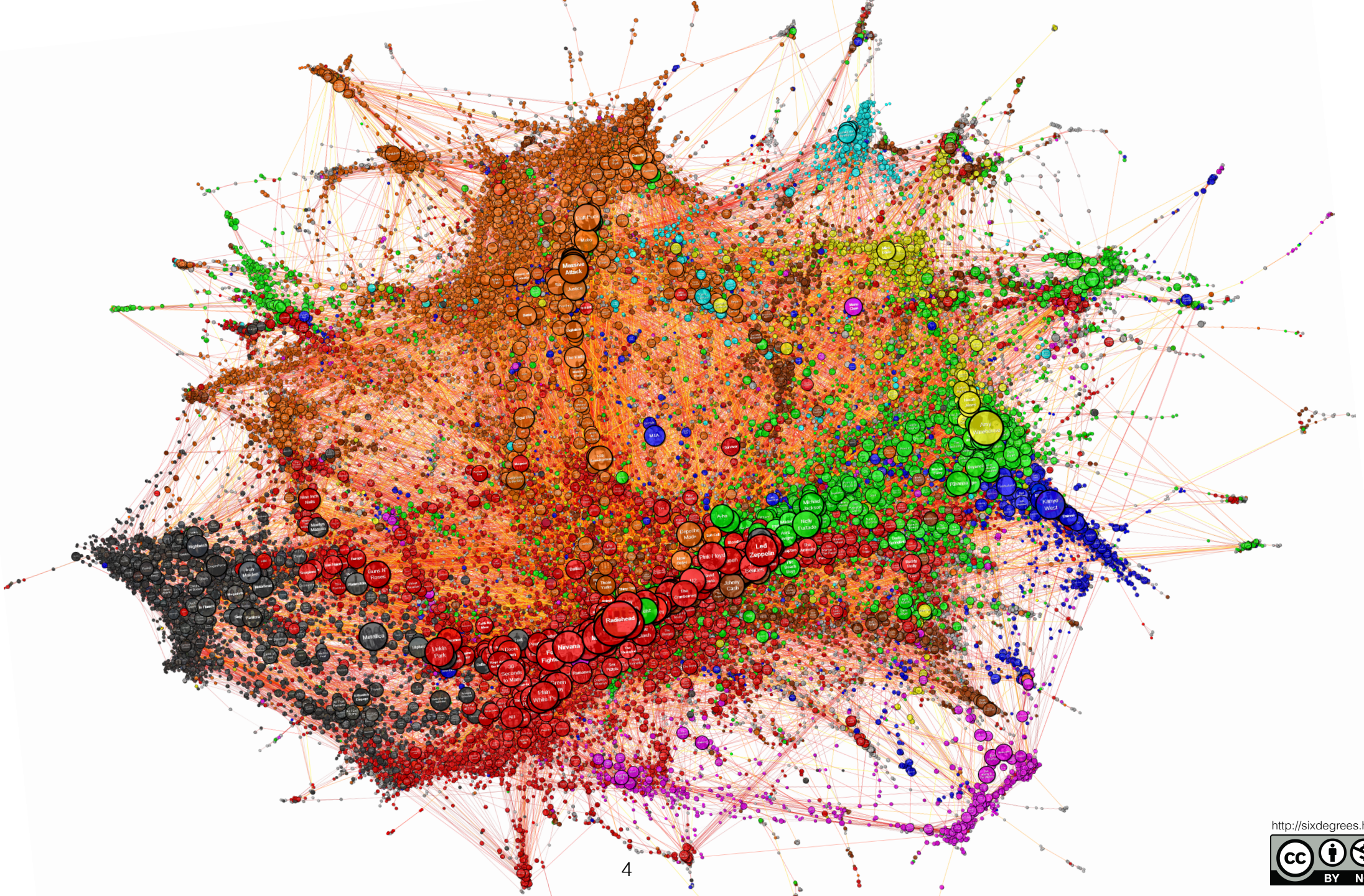
Music & the Internet
MUMT301

Gabriel Vigliensoni
Schulich School of Music
McGill University

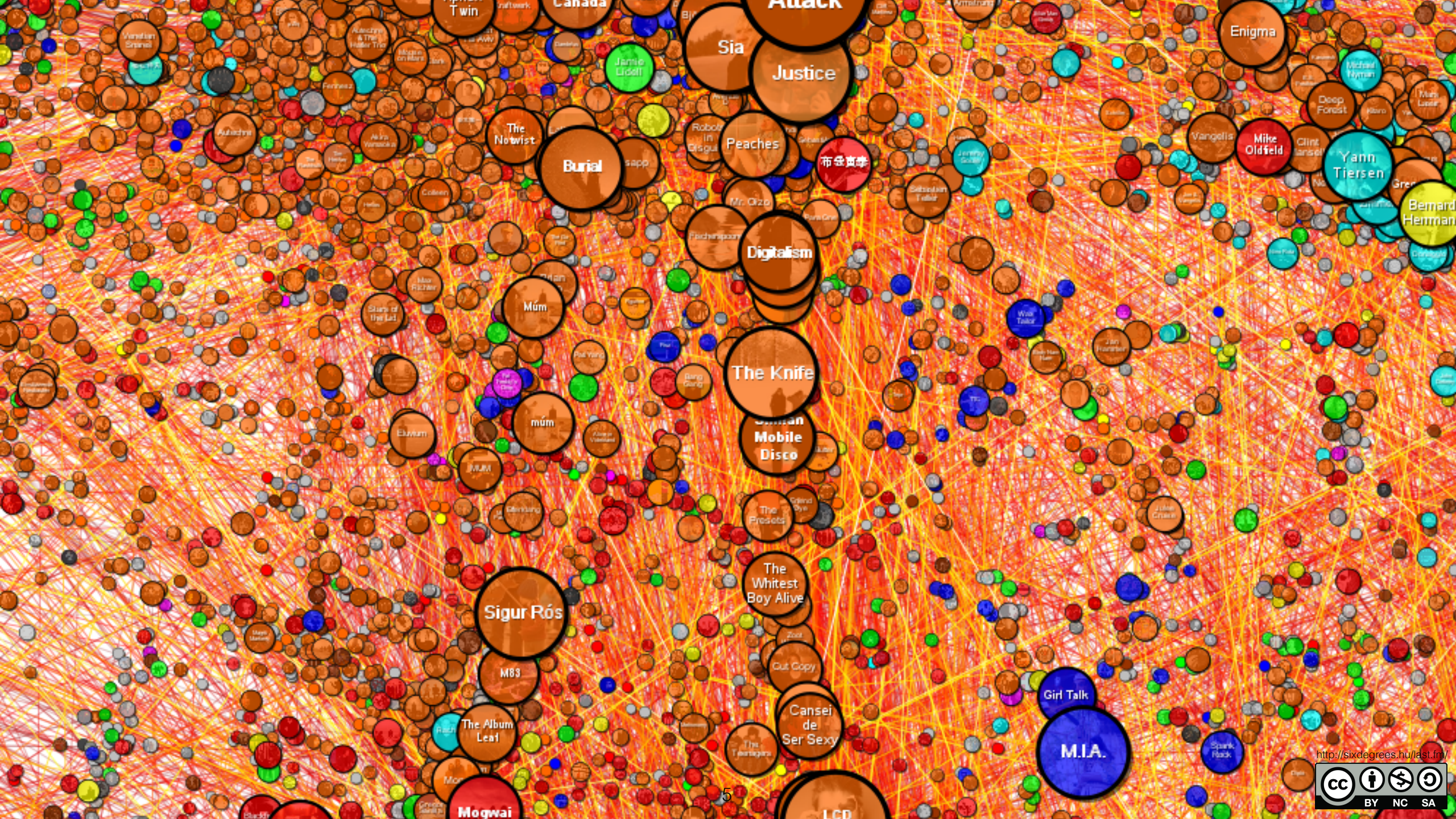
Plan

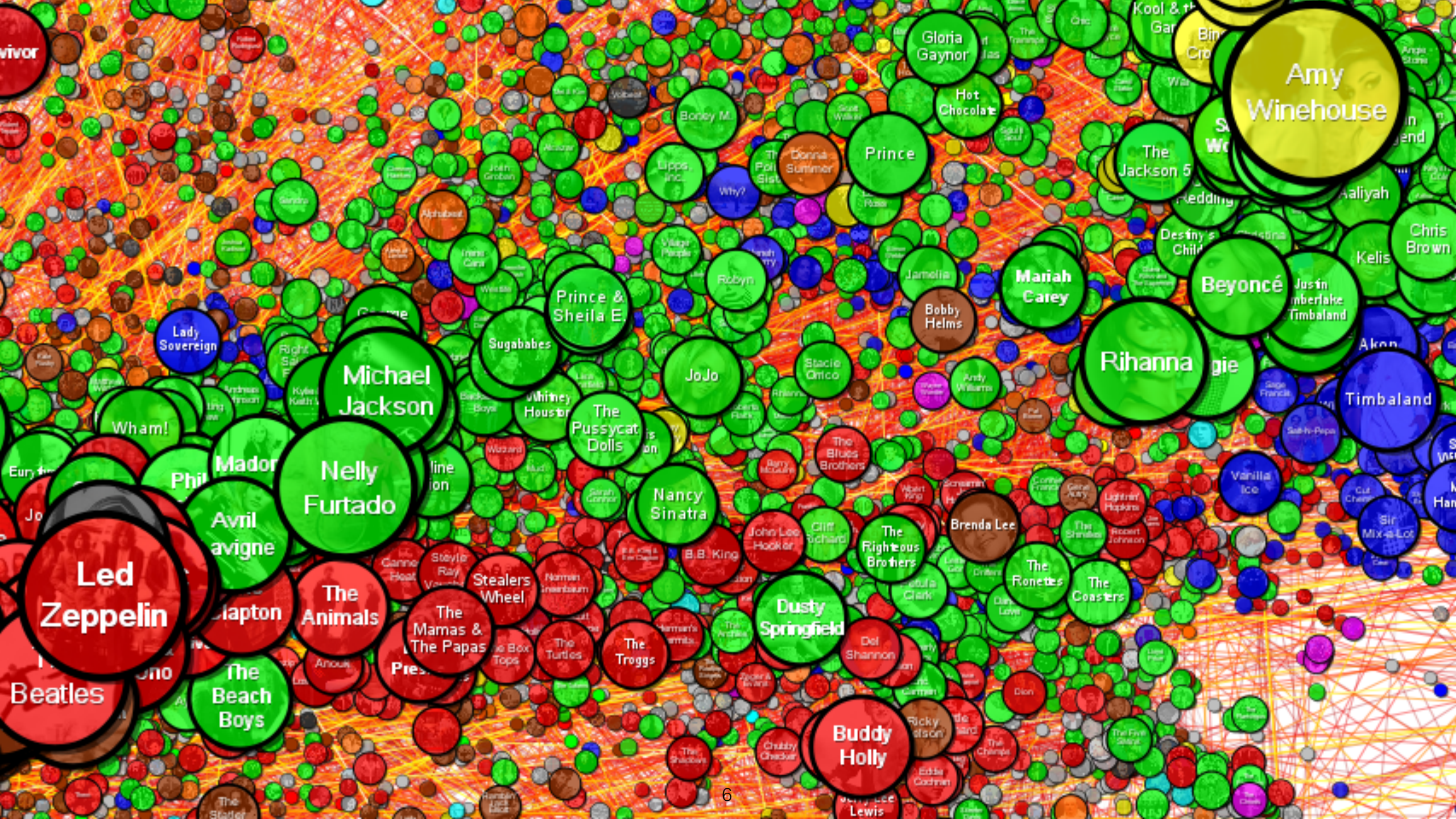
- Automatic music recommendation
- Internet radio broadcasting and streaming media software
- Intellectual Property
- Mid-term review

Automatic music recommendation systems



http://sixdegrees.hu/last.fm/


Sia
Justice
Enigma
Burial
The Notwist
Peaches
Digitalism
Yann Tiersen
Mike Oldfield
Vangelis
Deep Forest
Bernard Herman
Múm
múm
The Knife
Mobile Disco
The Presets
The Whitest Boy Alive
Sigur Rós
M83
The Album Leaf
Cut Copy
Cansei de Ser Sexy
Girl Talk
M.I.A.
Mogwai
LCD
http://sixdegrees.hu/last.fm/
BY NC SA

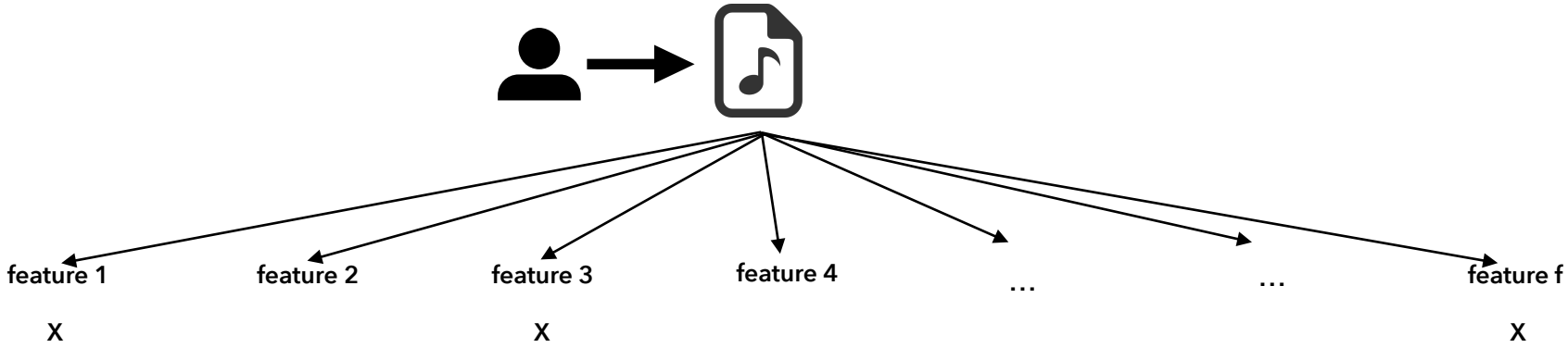

Amy Winehouse
Gloria Gaynor
Hot Chocolate
Prince
The Jackson 5
Aaliyah
Chris Brown
Kelis
Beyoncé
Mariah Carey
Bobby Helms
Rihanna
Timbaland
Prince & Sheila E.
Sugababes
Michael Jackson
Lady Sovereign
JoJo
Stacie Orrico
The Pussycat Dolls
Nelly Furtado
Avril Lavigne
Madonna
Wham!
Nancy Sinatra
Led Zeppelin
Beatles
The Animals
The Mamas & The Papas
Stealers Wheel
The Turtles
The Troggs
The Beach Boys
Dusty Springfield
Buddy Holly
Brenda Lee
The Righteous Brothers
The Ronettes
The Coasters
The Blues Brothers
Vanilla Ice
Sir Mix-a-Lot

Music recommendation approaches

- Content-based
- Collaborative filtering

Music recommendation approaches

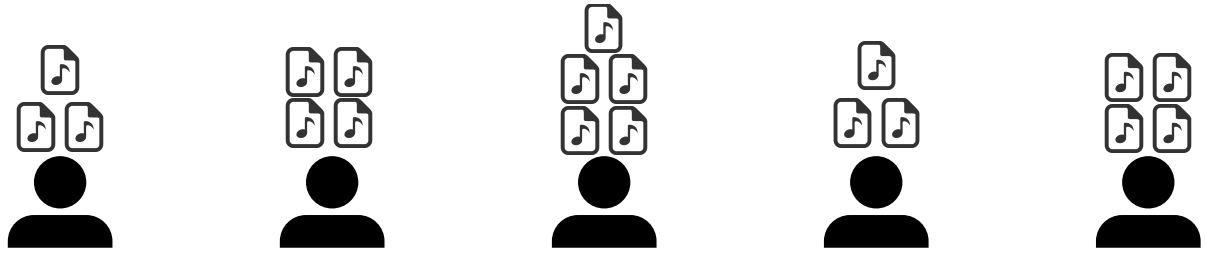
Content-based



	feature 1	feature 2	feature 3	feature 4	…	…	feature f
music item 1		X					X
music item 2	X	X			X		
music item 3		X			X		X
music item 4			X			X	
…				X			
…		X			X		
music item i	X		X				X

Music recommendation approaches

Collaborative filtering



user/artist	beatles	rolling	bsabbath	metallica	miles	coltrane	autechre	dpunk	reich	bartok
a	5	4			3	2	4		4	5
b	5	4		2				2	3	
c	4		3		3		5	4		
d	5			1	5	4				1
e	5				4		1		1	1

Music recommendation approaches

Collaborative filtering

User similarity

user/artist	beatles	rolling	bsabbath	metallica	miles	coltrane	autechre	dpunk	reich	bartok
a	5	4	3	3	3	2	4	3	4	5
b	5	4	2	2	5	3	3	2	3	2
c	4	4	3	1	3	3	5	4	5	3
d	5	4	2	1	5	4	3	2	2	1
e	5	3	2	1	4	3	1	1	1	1

Jaccard similarity
Intersection of sets

	a	b	c	d
b	0.33			
c	0.33	0.25		
d	0.50	0.25	0.25	
e	0.71	0.25	0.43	0.43

Cosine vector similarity
Angle of vectors in space

	a	b	c	d
b	0.99			
c	0.98	0.92		
d	0.82	0.98	0.99	
e	0.79	0.98	0.81	0.99

Pearson correlation
Distance of points in space

	a	b	c	d
b	0.93			
c	0.52	0.20		
d	-0.39	0.96	-0.51	
e	-0.06	1.00	-0.30	0.94

Music recommendation approaches

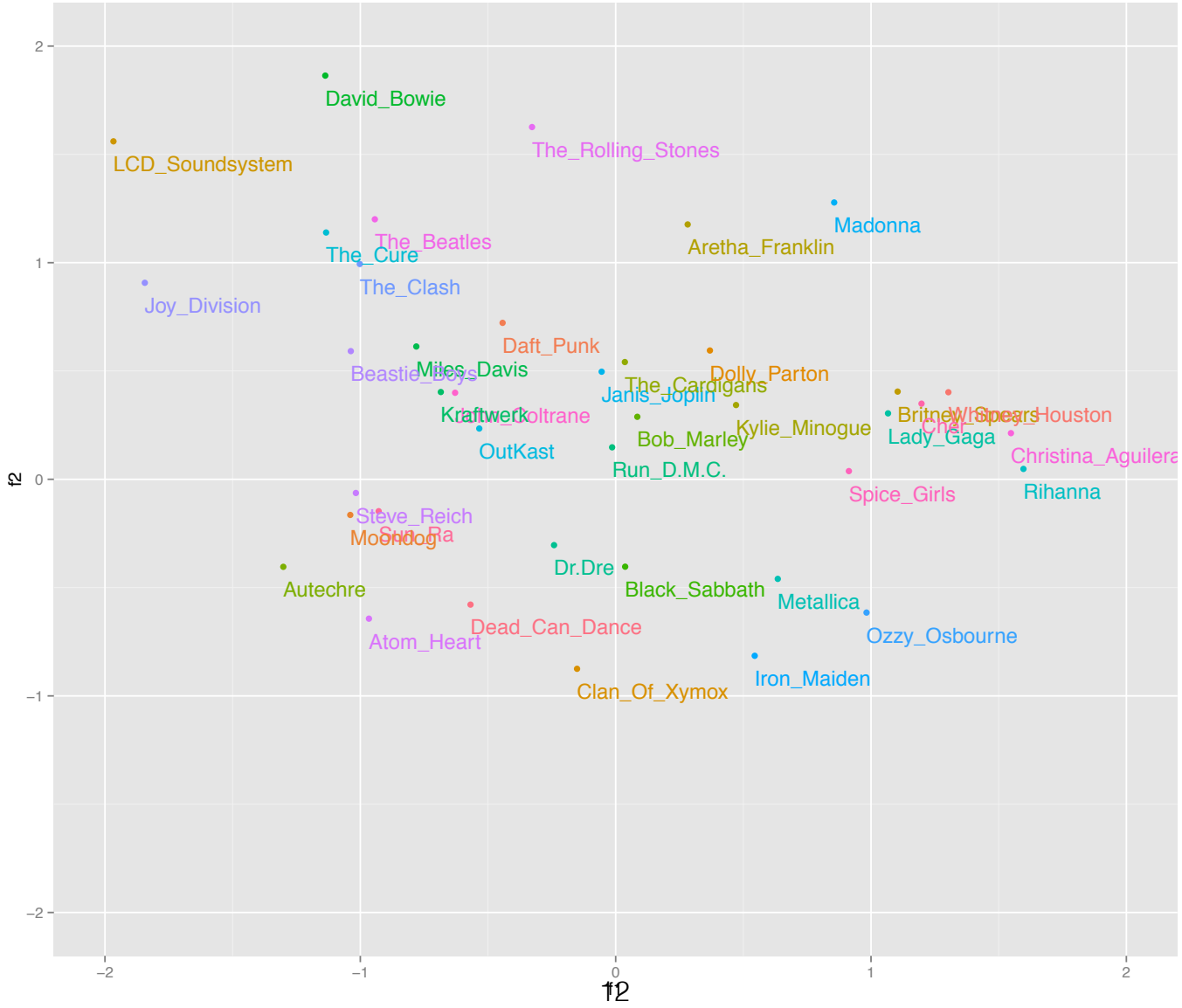
Collaborative filtering

Model-based

		factor 1	-0.188	-0.015	-0.006	0.069	-0.472	-0.353	0.238	0.078	0.168	0.481
		factor 2	-0.074	-0.001	-0.002	-0.014	-0.042	0.036	0.171	0.035	0.083	-0.186
factor 1	factor 2	**user/artist**	**beatles**	**rolling**	**bsabbath**	**metallica**	**miles**	**coltrane**	**autechre**	**dpunk**	**reich**	**bartok**
0.440	0.001	**a**	5	4	3	2	3	2	4	3	4	5
-0.009	-0.010	**b**	5	4	3	2	4	3	3	2	3	4
0.315	0.118	**c**	4	4	3	2	3	3	5	4	3	4
-0.634	0.115	**d**	5	4	3	1	5	4	3	3	2	1
-0.092	-0.229	**e**	5	3	2	1	4	3	1	2	1	1

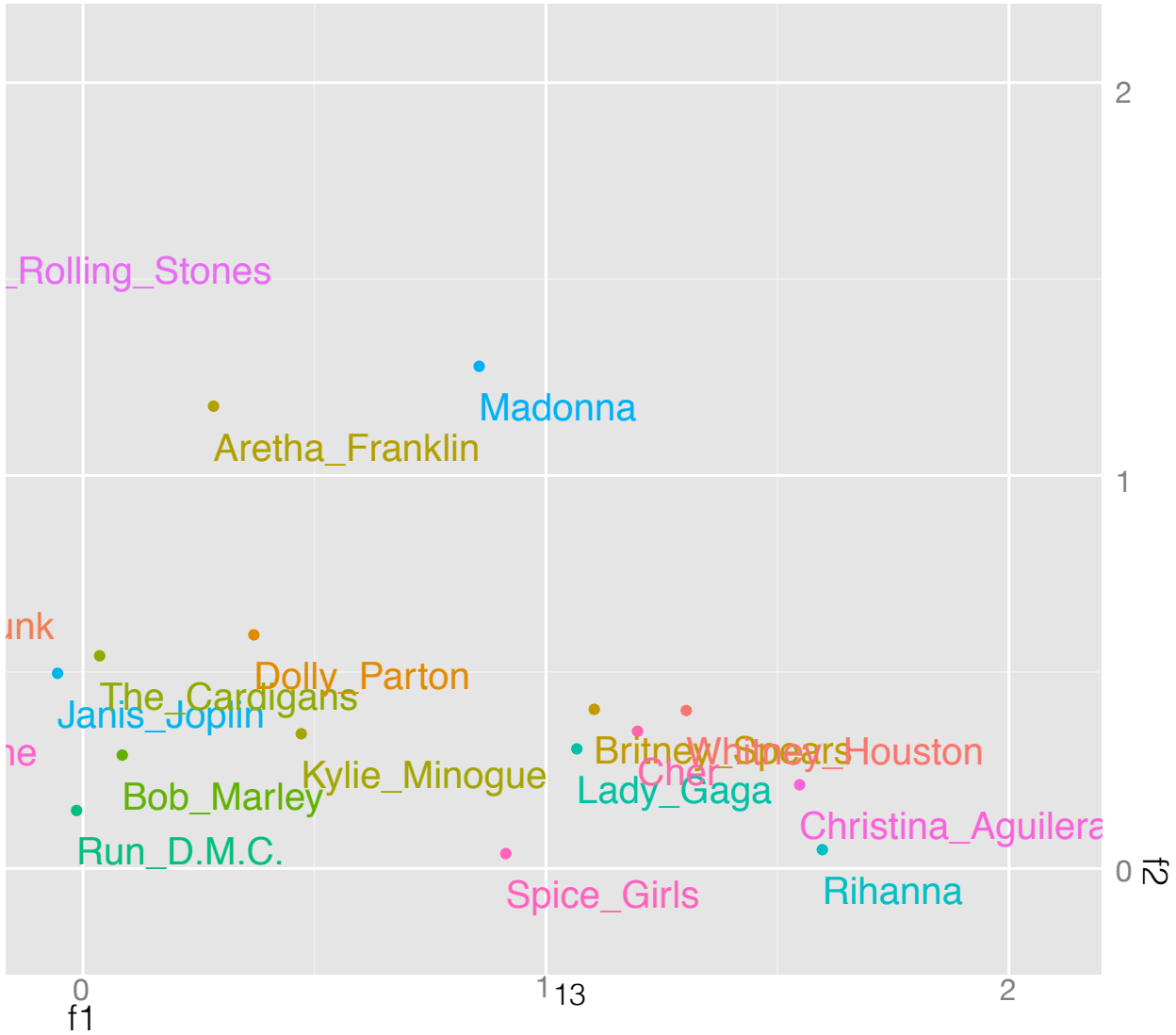
Music recommendation approaches

Latent factors models



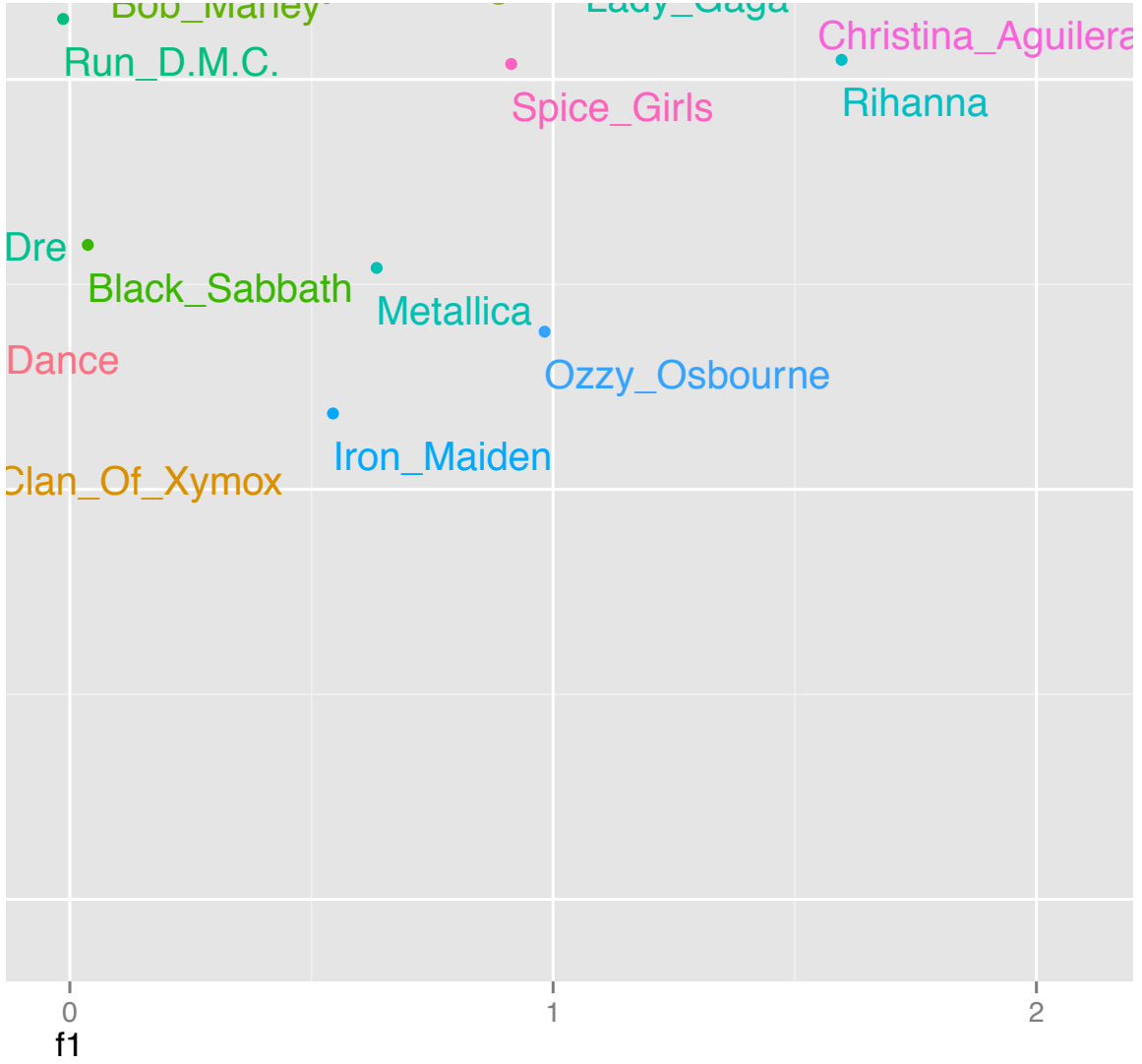
Music recommendation approaches

Latent factors models



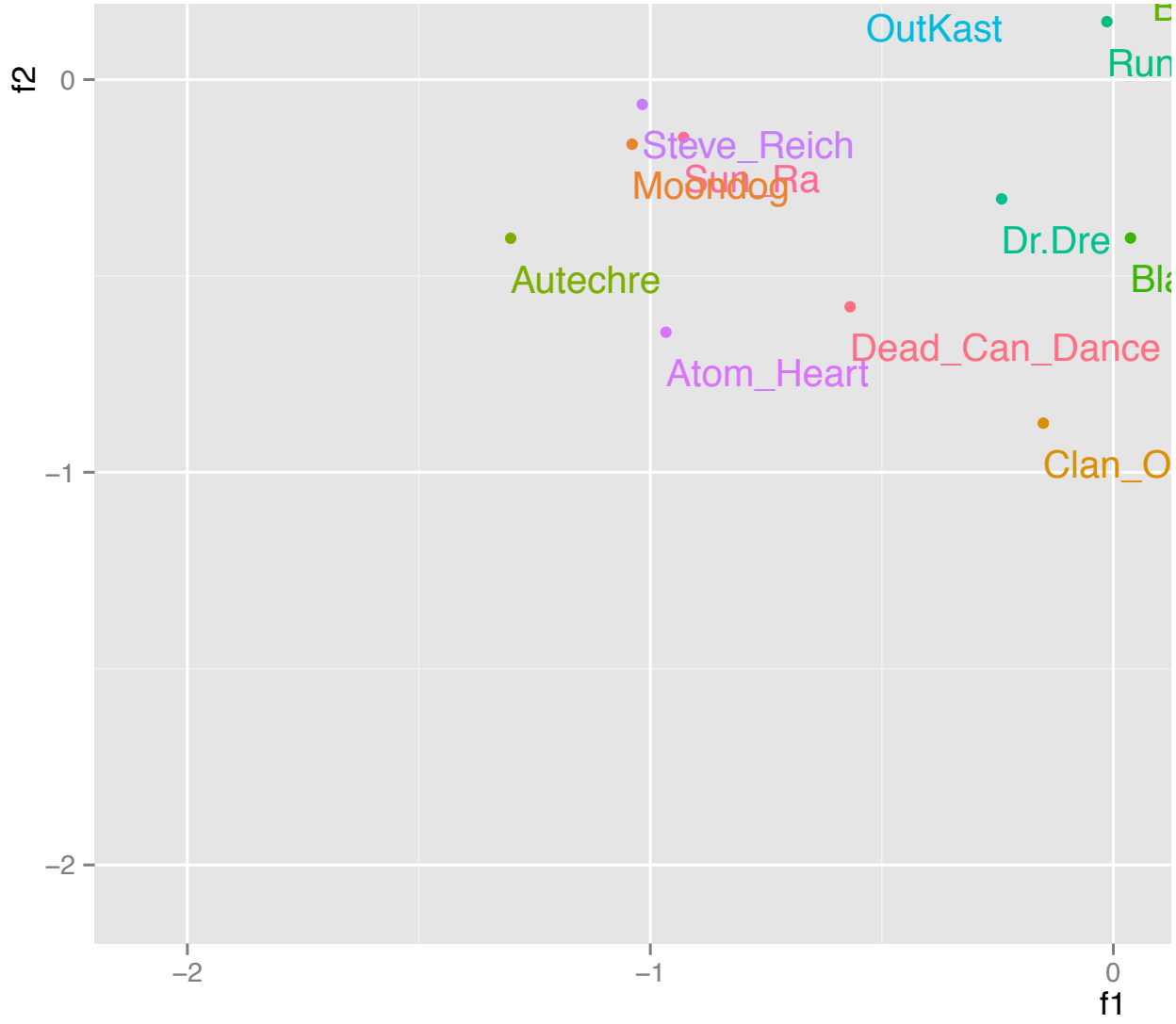
Music recommendation approaches

Latent factors models



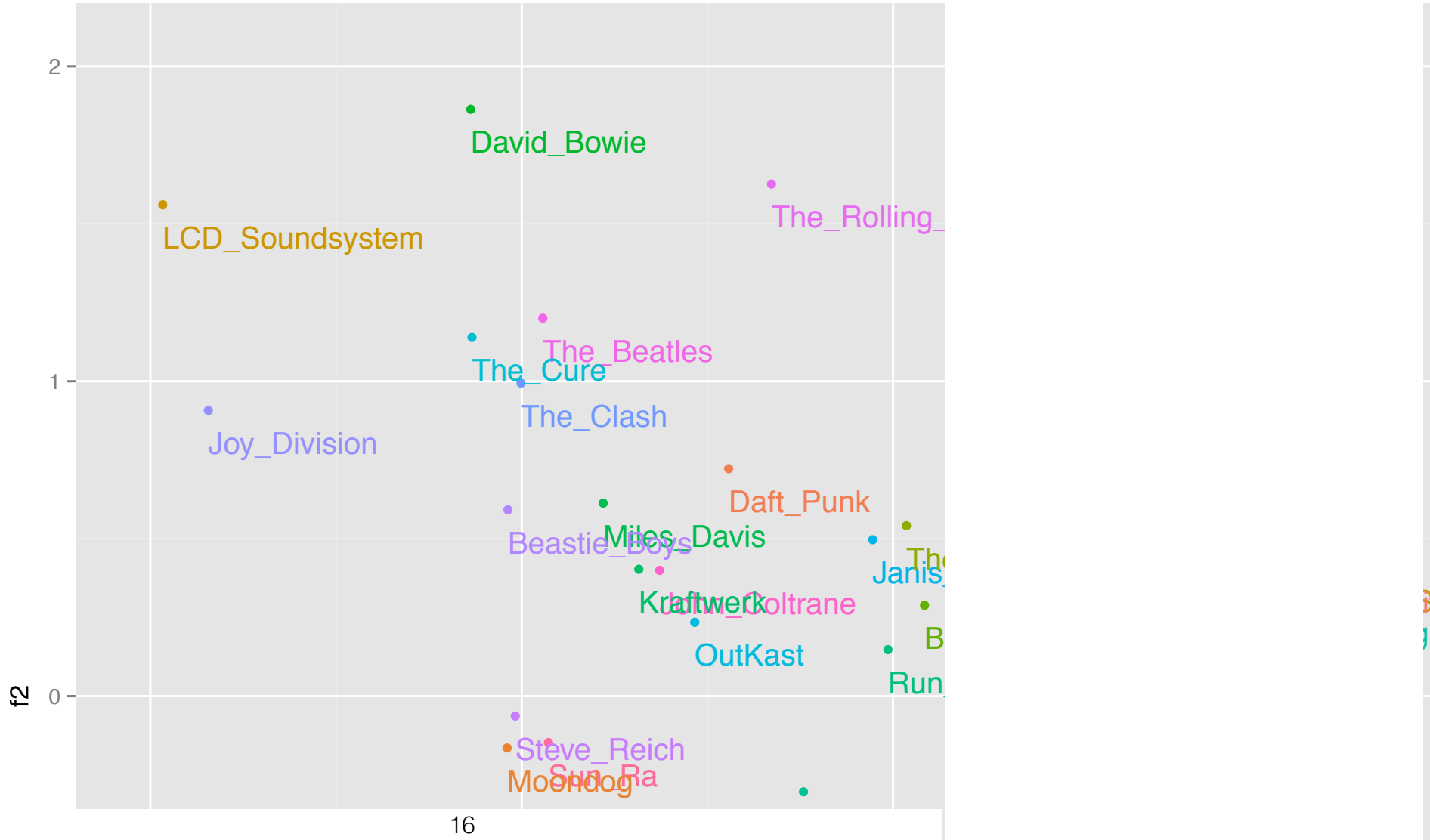
Music recommendation approaches

Latent factors models



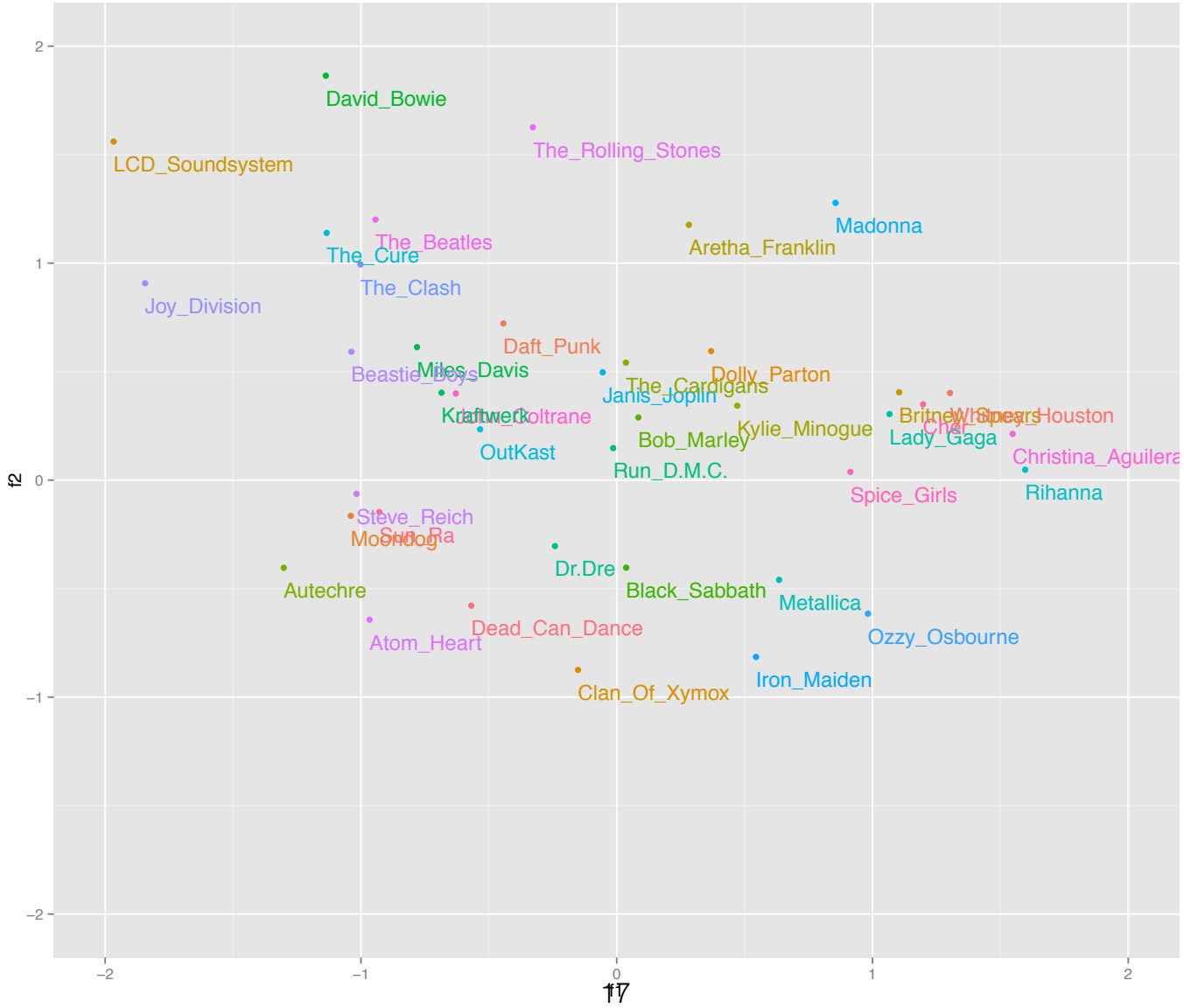
Music recommendation approaches

Latent factors models



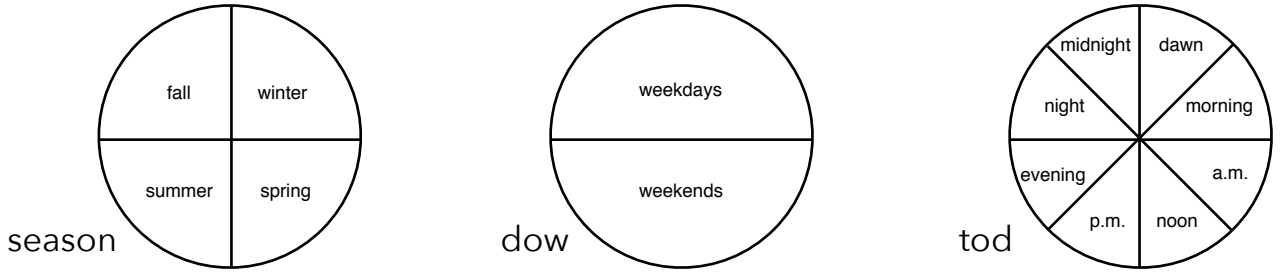
Music recommendation approaches

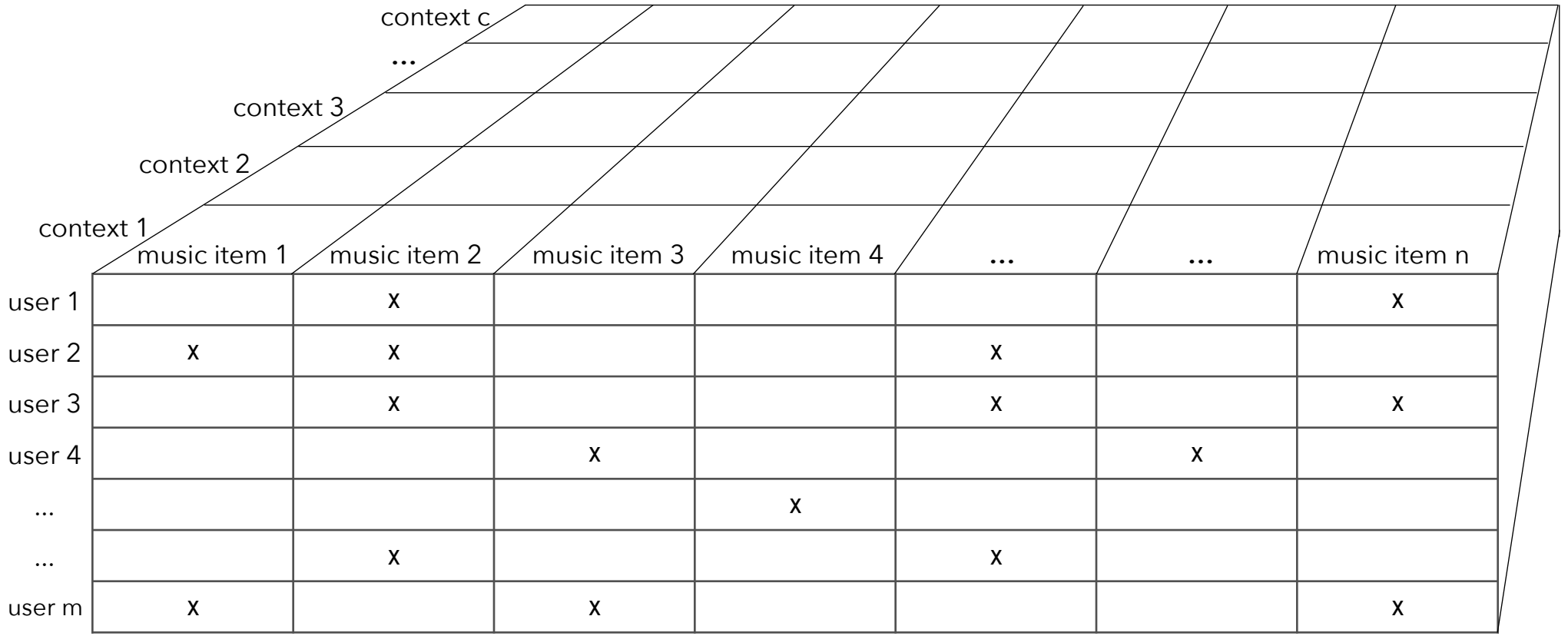
Latent factors models



Music recommendation approaches

Incorporating context into the model





Music recommendation approaches

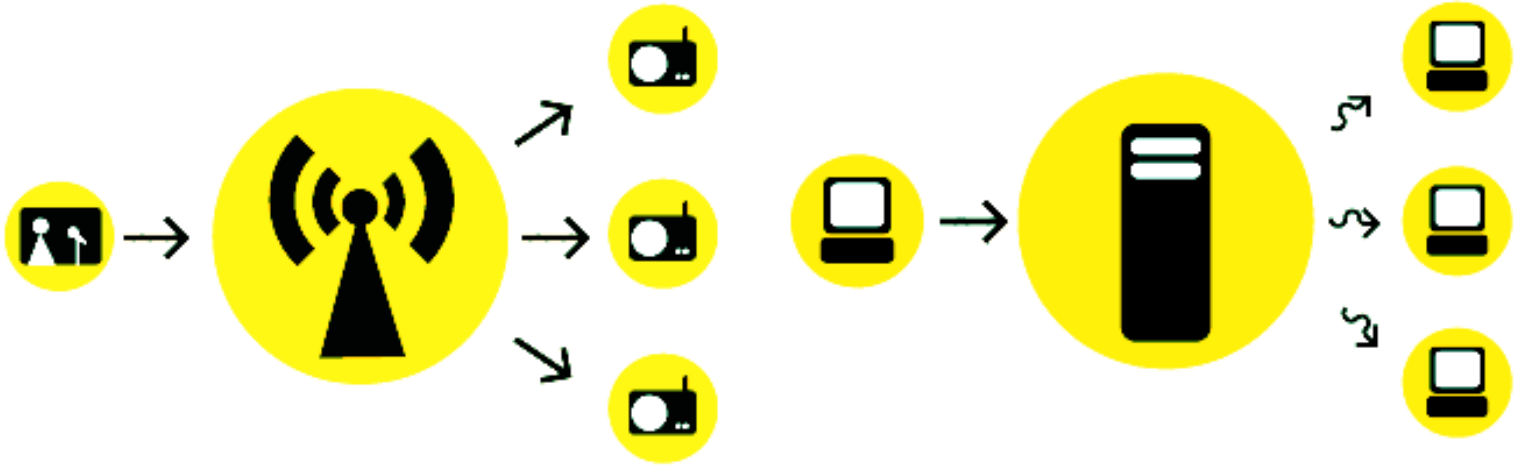
Incorporating side features into the model

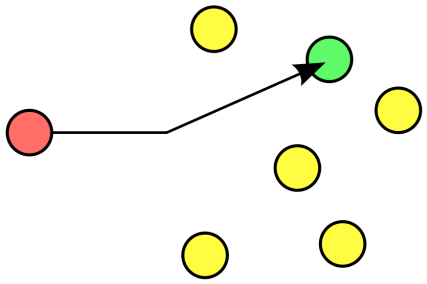
				factor1	-0.1250	-0.0758	0.2597	-0.4078	0.3469	0.4341	0.8519	0.5257	0.7062	-1.6830
				factor2	0.3663	0.2352	-0.1869	-0.3558	0.9251	0.8393	-0.8159	-1.2368	-0.8491	-0.0734
gender	**age**	**factor1**	**factor2**	**user/artist**	**beatles**	**rolling**	**bsabbath**	**metallica**	**miles**	**coltrane**	**autechre**	**dpunk**	**reich**	**bartok**
m	o	-0.7364	-0.5380	**a**	5	4	3	3	3	2	4	4	4	5
f	o	0.4071	0.3544	**b**	5	4	3	2	4	3	3	2	3	4
m	y	0.6459	-1.1424	**c**	4	3	3	1	3	2	5	4	5	1
m	y	0.8063	0.6618	**d**	5	4	3	1	5	4	4	2	4	1
f	y	-1.0800	0.7037	**e**	5	4	2	2	4	3	1	0	1	5
m	o	-0.0002	0.0000	**f**	5	4	2	2	4	3	4	3	4	4

side feature	value	factor 1	factor 2
age	o	-0.3872	-0.1737
age	y	0.1648	0.4031
gender	m	0.3986	-0.1552
gender	f	-0.3539	0.1451

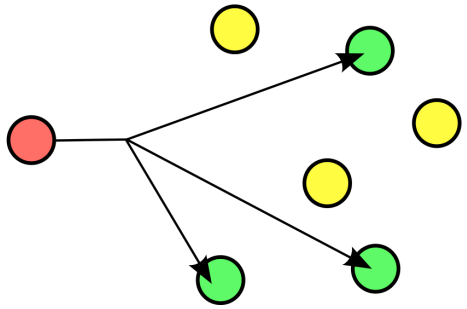
Rendle, Steffen. "Factorization machines." In IEEE 10th International Conference on Data Mining, pp. 995-1000. 2010.

Internet radio broadcasting



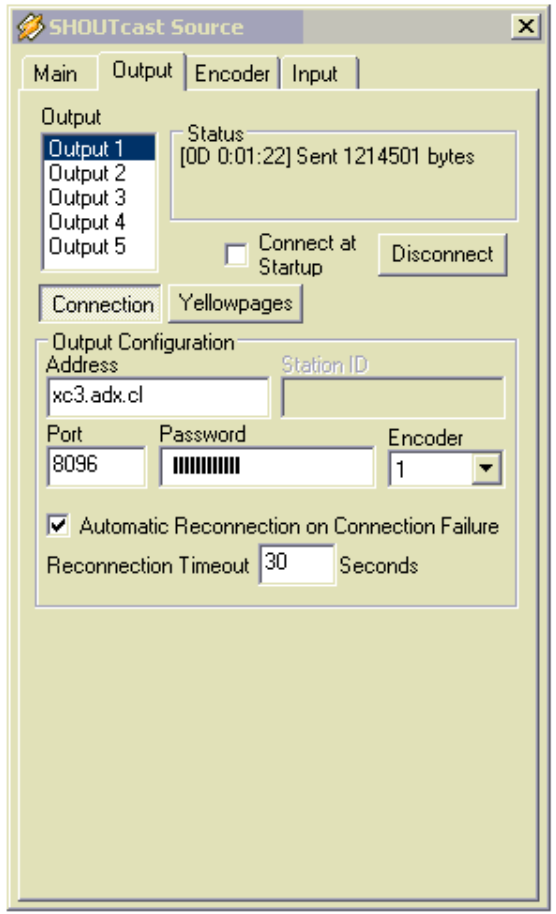


Unicast: Sends IP packets to identified recipients on a network, with added bandwidth

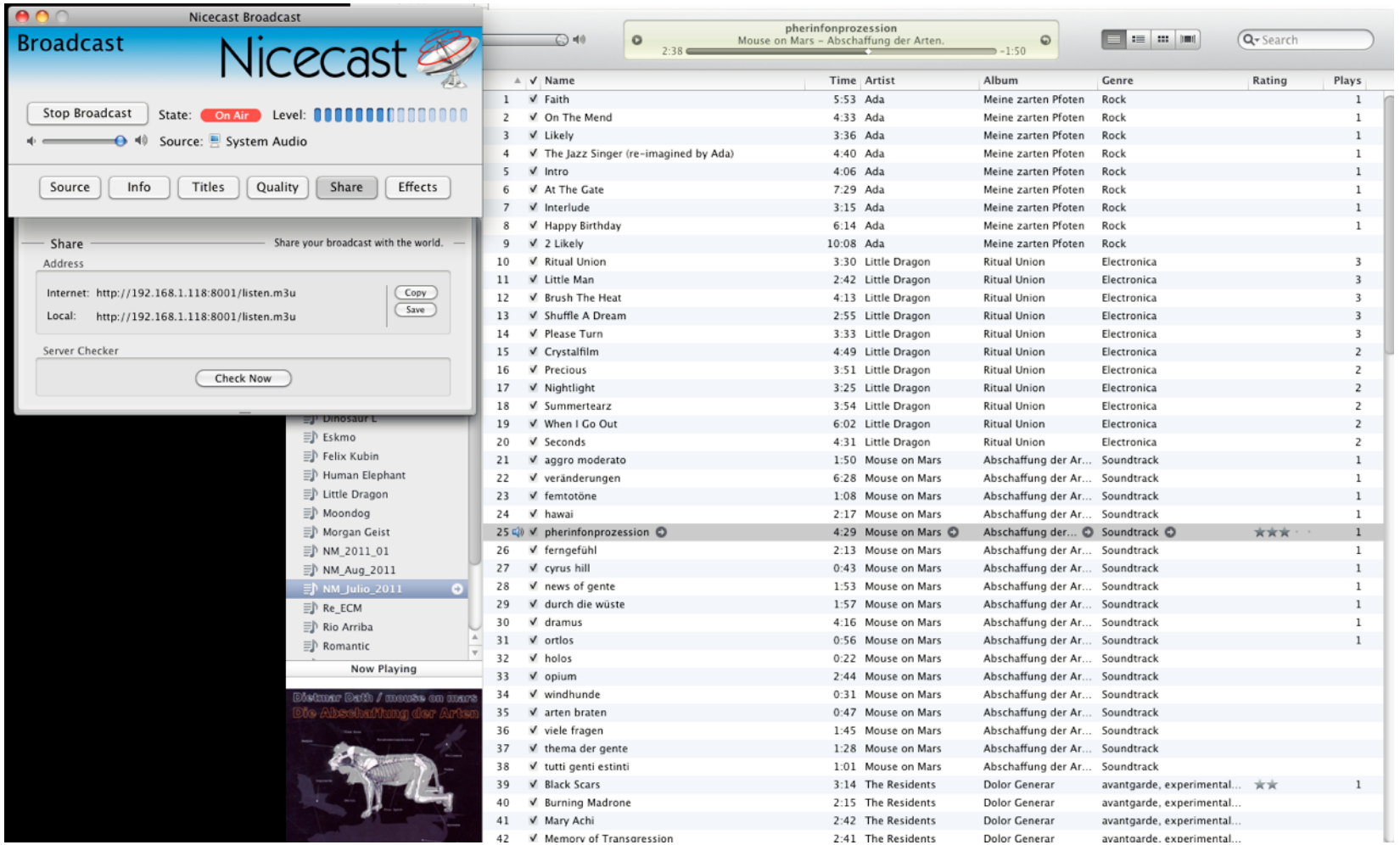


Multicast: sends IP packets to a group of hosts on a network, with no added bandwidth, and not requiring prior knowledge of who or how many receivers there are.

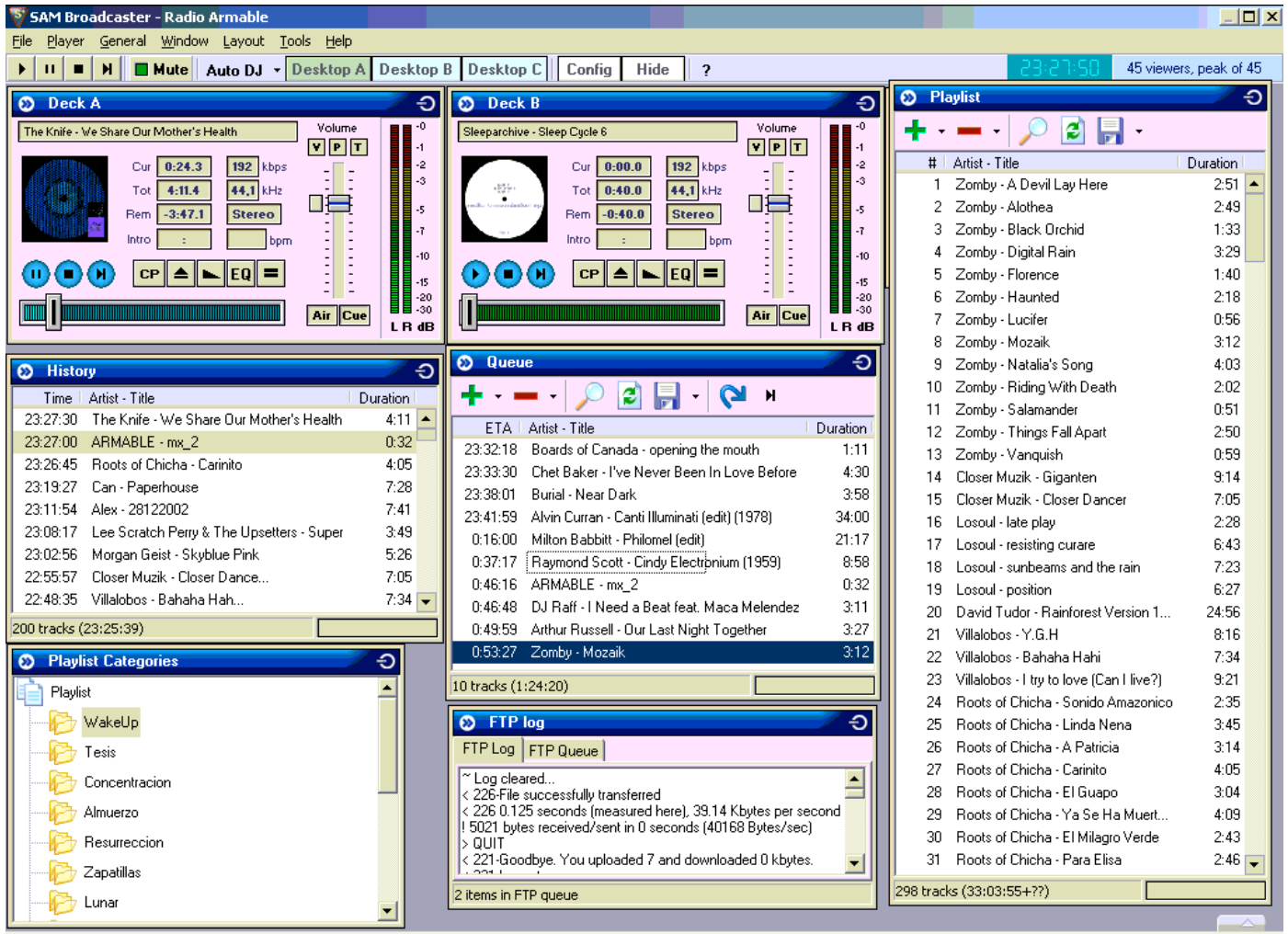
Playlist creation applications



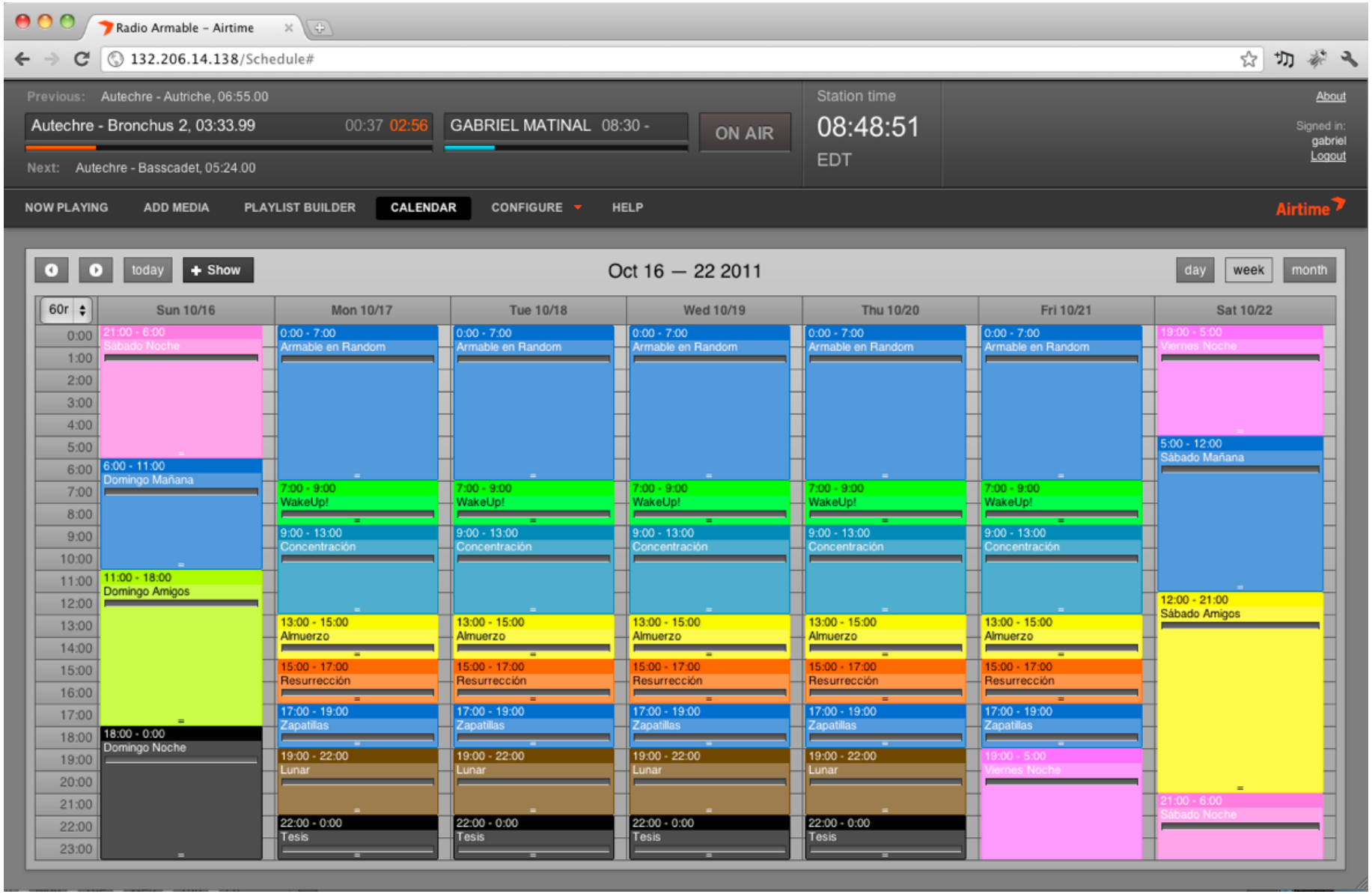
Playlist creation applications



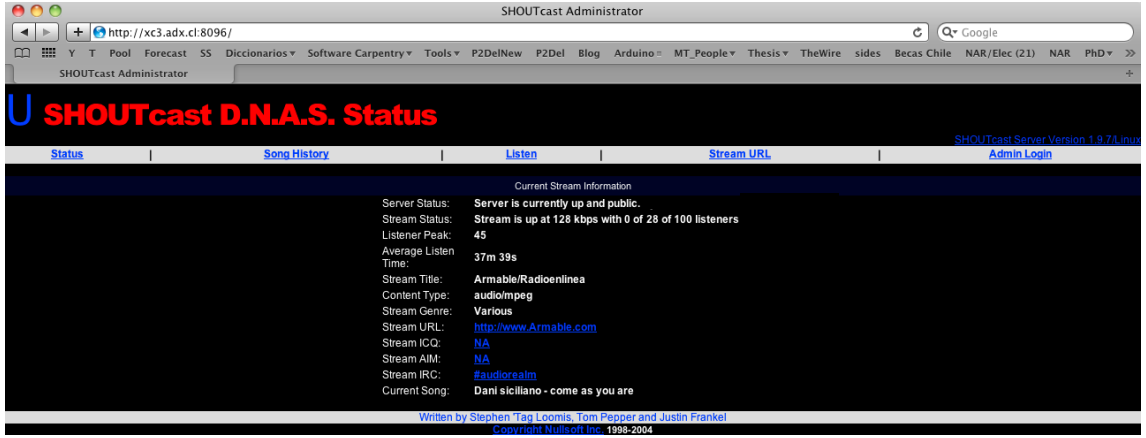
Automated radio broadcasting

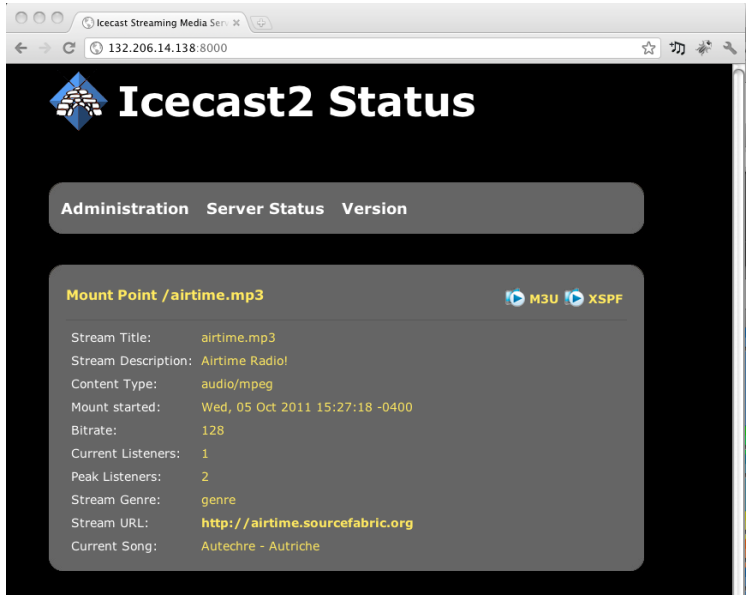


Social radio applications



Streaming Media Servers





Streaming Media Servers

Demo!

- Playlisting app: iTunes
- Audio routing (+mixing) app: Audio Hijack
- Streaming media server: Icecast

Intellectual property

Compiled from http://www.cipo.ic.gc.ca/

- Copyright Act: any original literary, dramatic, musical (musical compositions with or without words) or artistic work is automatically protected by copyright the moment it is created
- In the simplest terms, “copyright” means “the right to copy”: the right to reproduce a work, or a substantial part of it, in any form
- In the case of music or sound:
  - **a recording consisting of sounds**
  - **a performance of a musical work**
  - **an improvisation of a musical work**
- Copyright in Canada

Intellectual property

Compiled from http://www.cipo.ic.gc.ca/

- The **work's creator** is usually the **copyright owner**
  - Exceptions: an employer has copyright for works created by employees unless there is an agreement to the contrary
- When you own copyright on a work you can **control how it is used**
  - **Selling the right** to use the work, or **getting a permission** to use it
  - You can **limit its use** to protect the value of the copyrighted work.
- Even though copyright protection is automatic, **registration** gives you evidence of ownership
  - **Certificate of registration of copyright:** evidence that copyright exists and that the person registered is the owner of the copyright

Intellectual property

Compiled from http://www.cipo.ic.gc.ca/

- Application in Canada for a **registration of a copyright** in a
  - Work
  - Performer's performance or sound recording
- Fee
- Copyright generally exists for the **life of the author/s plus an amount of time after their death**
- This amount of **time varies per country**
  - Canada: 50 years after the death of the author
  - US: 70 years after the death of the author
- **After that, the work becomes part of the public domain** and anyone can use it.

Canada and Copyright

- Why there were just a few on-demand music services in Canada?
  - Canada is a party to the **Berne Convention of 1886**
    - aspects of modern copyright law were set
    - copyright exists the moment a work is “fixed"
    - countries recognize copyrights held by the citizens of all other signatory countries.
- Canada **signed but did not ratified** both the WIPO Copyright Treaty of 1996, and the WIPO Performances and Phonograms Treaty of 1996
- File sharing in Canada
- That changed four years ago!

Fair use

http://fairuse.stanford.edu/overview/fair-use/what-is-fair-use/

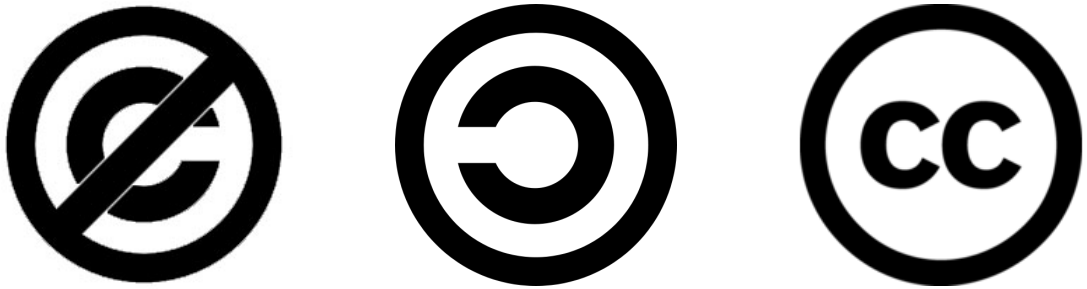
- Fair Use
  - copying of copyrighted material done for a **limited and "transformative" purpose**
  - based on the belief that the public is entitled to **freely use portions** of copyrighted materials for purposes of **commentary** and **criticism**
  - such uses can be done **without permission** from the copyright owner
  - if your use qualifies as fair use, then it would not be considered an illegal infringement
- The **Four Factors** are used to measure fair use**:**
  - the **purpose and character** of usage (the transformative factor)
  - the **nature** of copyrighted work
  - the **amount and substantialit**y of the portion taken
  - the **effect of the use** upon the potential market
- What If You Acknowledge the Source Material?
  - **Acknowledgment of the source** material (such as citing a photographer) may be a **consideration** in a fair use determination, but **it will not protect** against a claim of infringement.
- E.g.,
  - Text: limited copying for educational purposes
  - Audio/visual content: limited performance for educational purposed
  - Graphical content: limited display for educational purposes

Fair use

http://fairuse.stanford.edu/overview/fair-use/what-is-fair-use/

- Fair use music cases
  - A television film crew, covering an Italian festival in Manhattan, recorded a **band playing a portion of a copyrighted song** "Dove sta Zaza." The music was replayed during a news broadcast.
  - A woman was sued for copyright infringement for **downloading 30 songs using peer-to-peer file sharing software**. She argued that her activity was a **fair use because she was downloading the songs to determine if she wanted to buy them later**.
  - Campbell v. Acuff-Rose Music Inc. (AKA 2 Live Crew vs. Roy Orbison)
  - WAY
  - Danger Mouse's The Grey album
    - Jay-Z and Macca reaction

Copyright alternatives

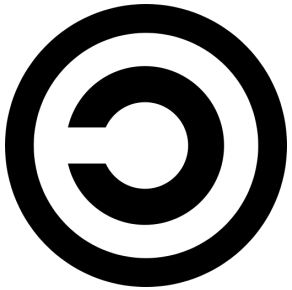


Public Domain

http://fairuse.stanford.edu/overview/fair-use/what-is-fair-use/

- "Creative materials that are not protected by intellectual property laws such as copyright, trademark, or patent laws."
- **Four common ways** that works arrive in the public domain
  - the copyright has **expired**
  - the owner failed to follow copyright **renewal rules**
  - the owner deliberately **places it in the public domain**
  - copyright law does not protect this type of work

Copyleft



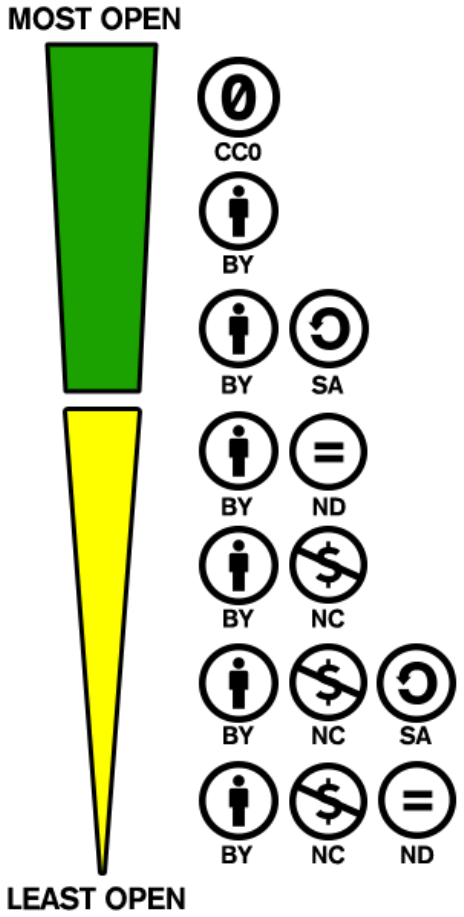
- Practice of using copyright law to offer **the right to distribute copies and modified versions of a work,** requiring that **the same rights be preserved in modified versions of the work**
- Copyleft can be characterized as a copyright licensing scheme in which an **author surrenders some, but not all rights** under copyright law
- **Instead of allowing a work to fall completely into the public domain** (where no ownership of copyright is claimed), **copyleft allows an author to impose some restrictions** on those who want to engage in activities that would more usually be reserved by the copyright holder
- Four **types of freedom**:
  - Freedom 0 – the freedom to use the work
  - Freedom 1 – the freedom to study the work
  - Freedom 2 – the freedom to copy and share the work with others
  - Freedom 3 – the freedom to modify the work, and the freedom to distribute modified and therefore derivative works
- Under copyleft, **derived works may be produced** provided they are released under the compatible copyleft scheme

Creative Commons

- "Reasonable, flexible copyright" at www.creativecommons.org
- Copyleft-like licenses for creative work
- Four main variables
  - Attribution
  - Non-commercial
  - No derivative works
  - Share alike

CC licenses

Creative Commons



CC0 Public Domain Dedication
Attribution
Attribution-ShareAlike
Attribution-NoDerivs
Attribution-NonCommercial
Attribution-NonCommercial-ShareAlike
Attribution-NonCommercial-NoDerivs

CC Music

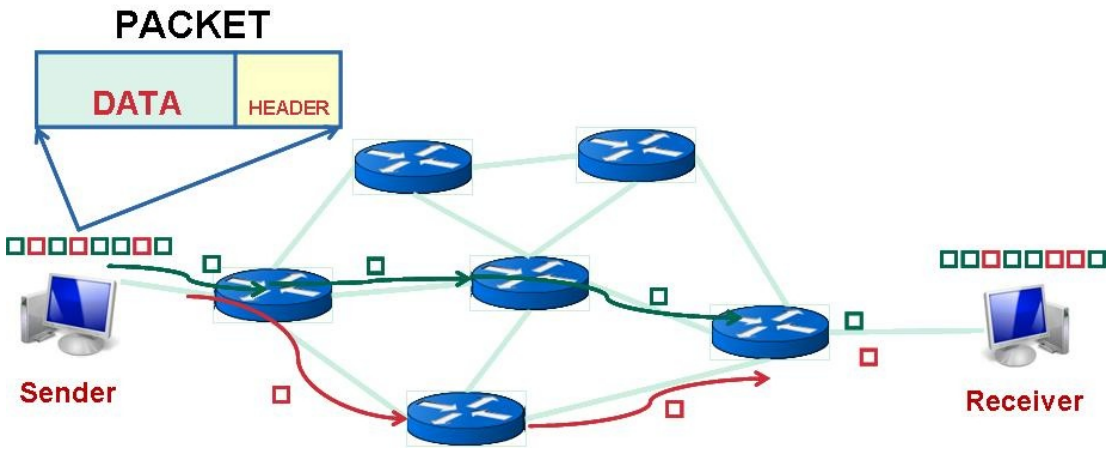
- A new public domain score and studio recording of Bach's Goldberg Variations - set the Goldbergs free!. The link!
- Chris Zabriskie, on CC
- Creative Common Music Communities
- Creative Commons Record Labels
- Freesound
- Kompoz
- iBeat
- Musopen
- InternetArchive
- IMSLP

Mid-term review

New music economy

- Three dimensions (tensions) between the old music economy (OME) and the new music economy (NME)
  - Connectivity vs. control
  - Service vs. product
  - Amateur vs. professional

Packet switching



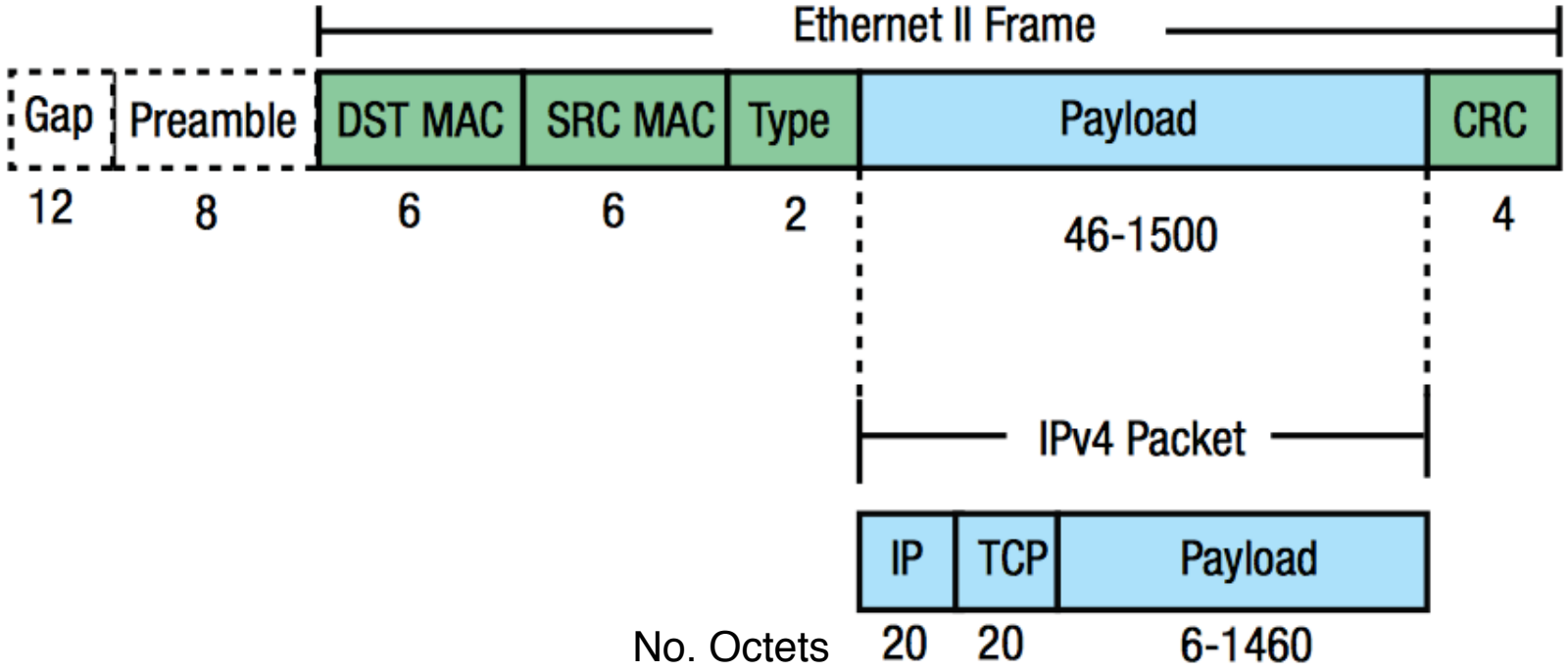
Taken from http://computernetworkingsimplified.com/physical-layer/overview-circuit-switching-packet-switching/

Internet technologies

- Ethernet
- TCP/IP
- OSI Model
- IP Addresses
- DNS
- Ports
- DHCP
- FTP
- SSH
- HTTP

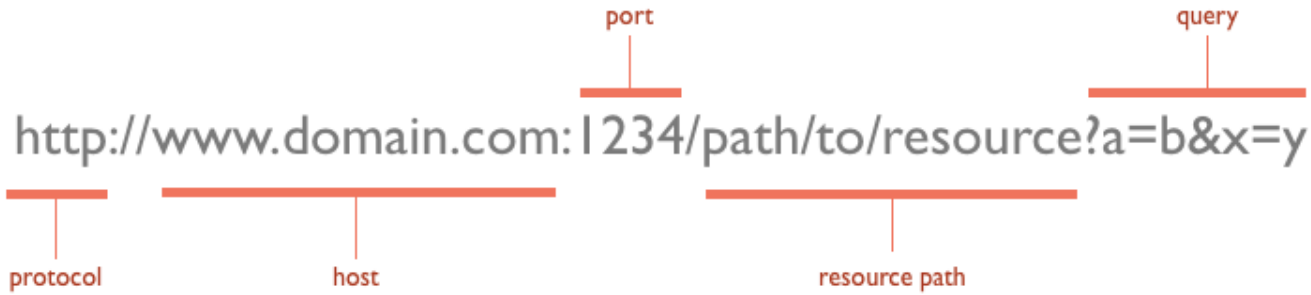
Complete Ethernet Packet

Taken from openmicrolab.com



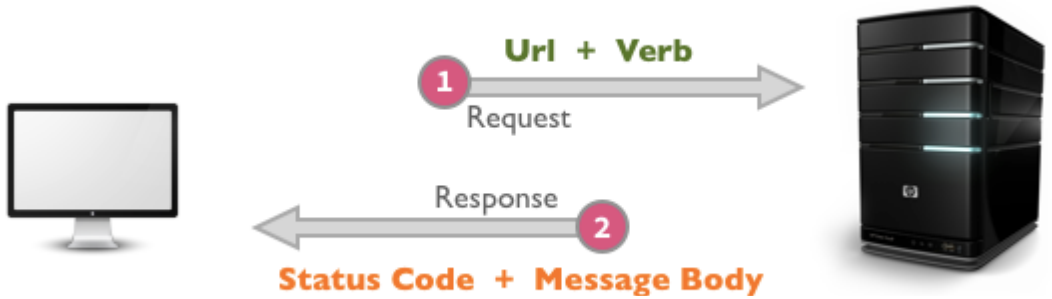
HTTP requests

- Request messages are at the heart of web communications using HTTP
- These messages are sent using URLs (Uniform Resource Locators)



HTTP

- “The first version of the protocol had only one method, namely GET, which would request a page from a server. The response from the server was always an HTML page.” (T. Berners-Lee)



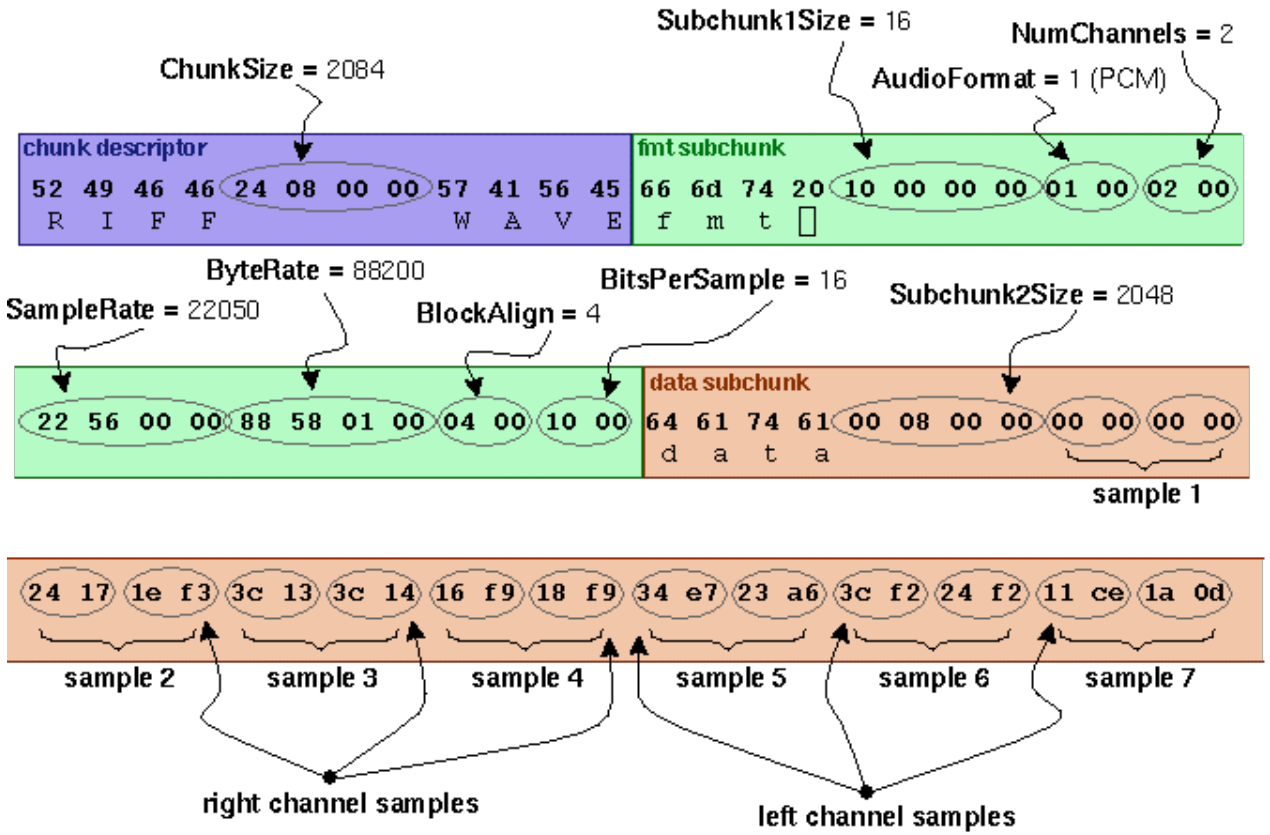
Taken from http://code.tutsplus.com/

- However, these days there are some other HTTP “verbs” that allow us to perform other actions on resources:
  - GET: fetch an existing resource
  - POST: create a new resource
  - PUT: update an existing resource
  - DELETE: delete an existing resource

Sound file formats

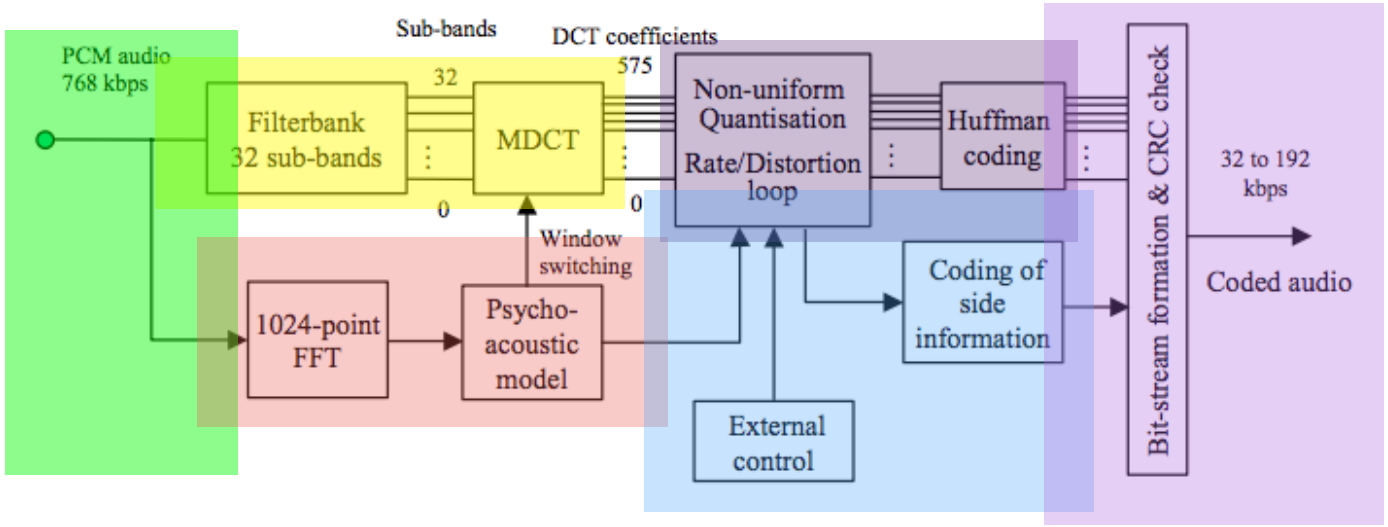
- Broadly speaking, sound content is delivered in two format categories:
  - As **structured audio**
    - Sounds are generated in a dynamic manner at runtime
    - MIDI, MODs (e.g., trackers)
  - As **recorded sound**:
    - often called *waveform* sound
    - audio data can be stored in
      - **uncompressed** formats
      - **compressed** formats
        - **Lossy** formats
        - **Lossless** formats

Uncompressed formats
The Canonical WAVE file format



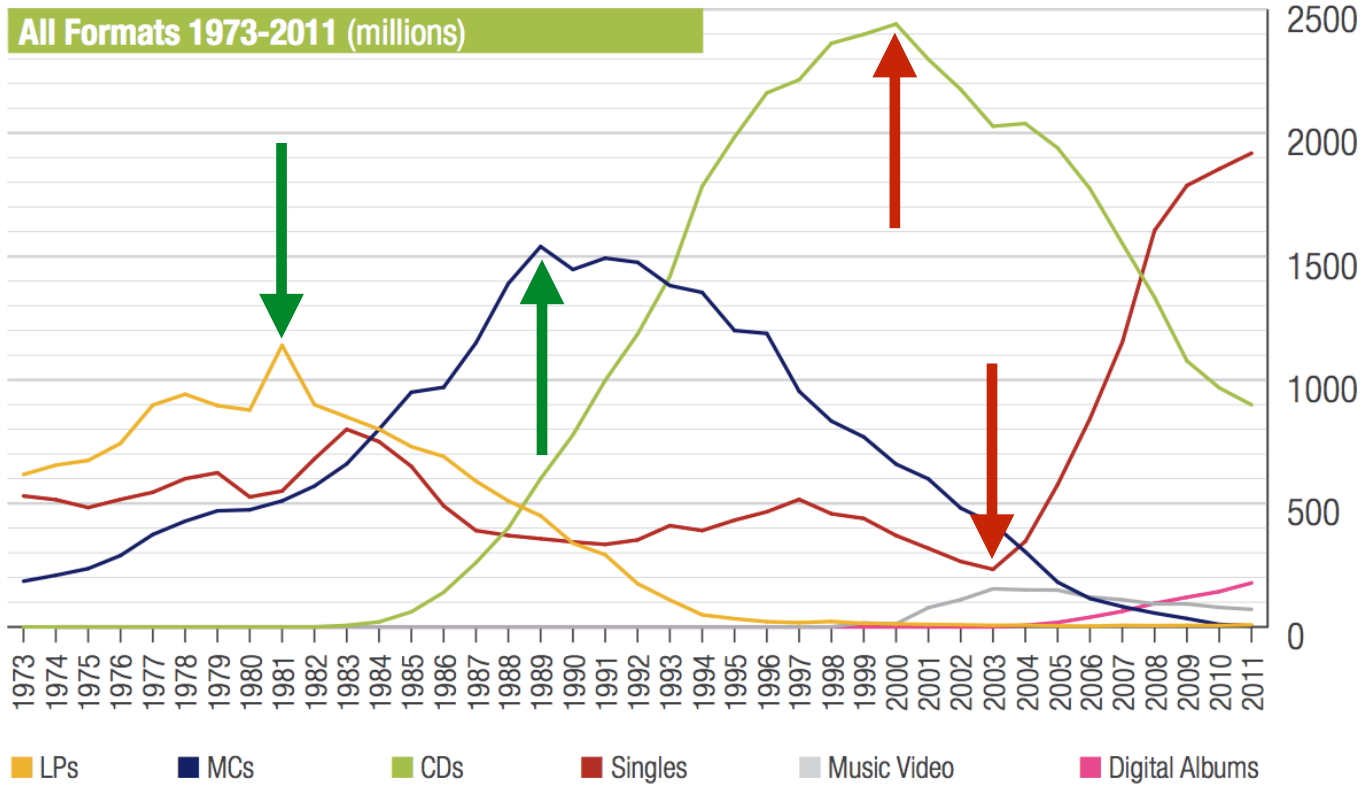
Taken from http://soundfile.sapp.org/doc/WaveFormat/

MPEG-1 Layer-3 encoder diagram



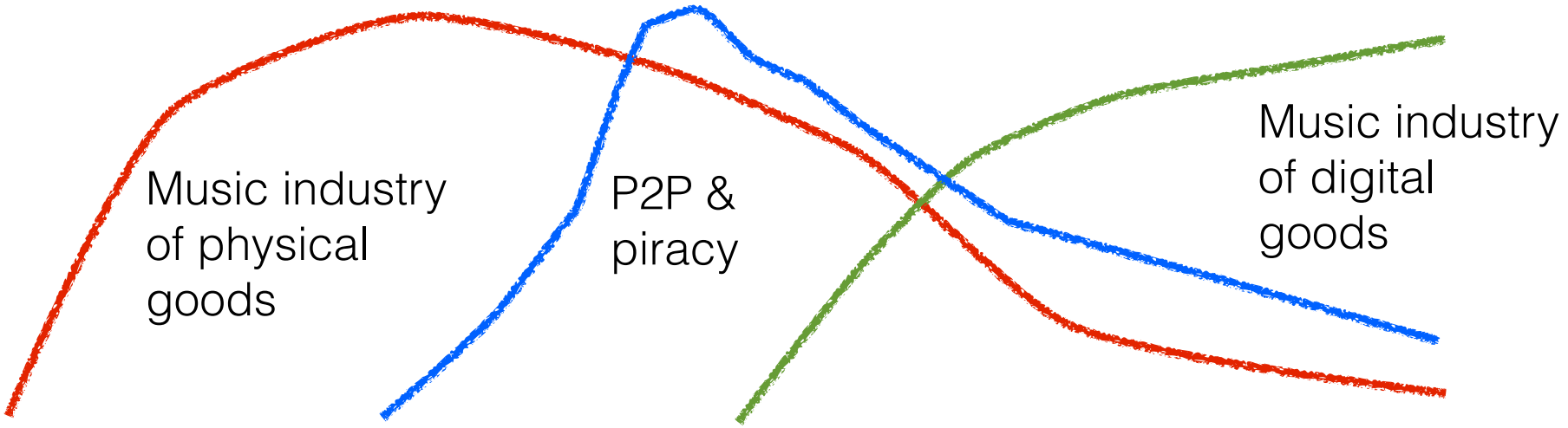
Taken from Brandeburg, K. 1999. MP3 and AAC explained. In Proceedings of the AES 17th International Conference on High Quality Audio Coding

Global music sales by format



Taken from Record Industry in Numbers 2011 (IFPI)

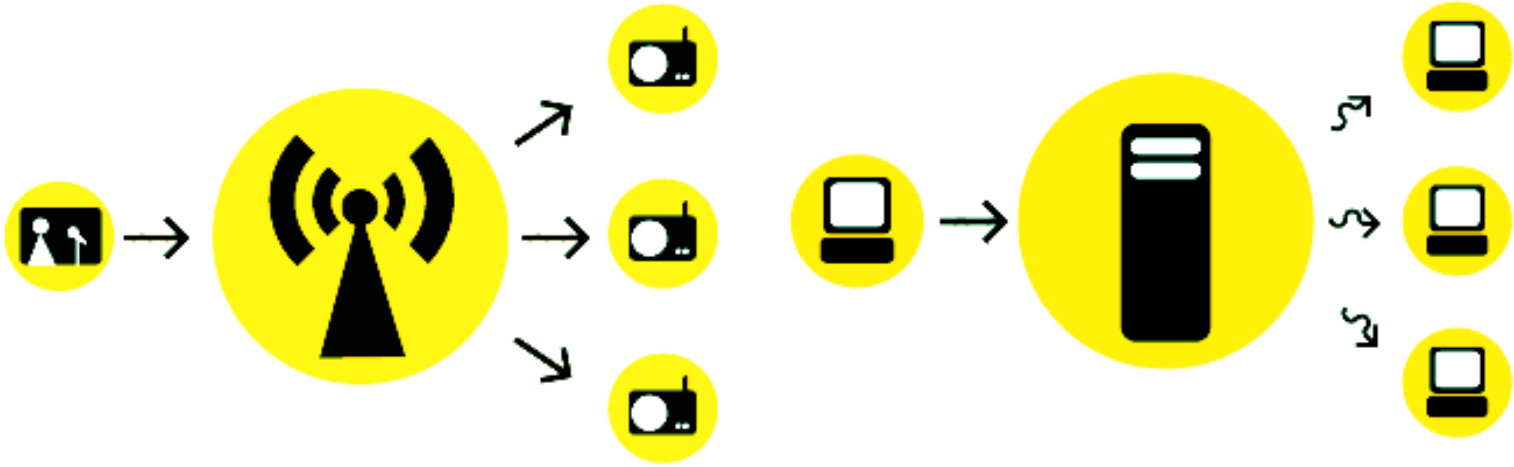
Music industry

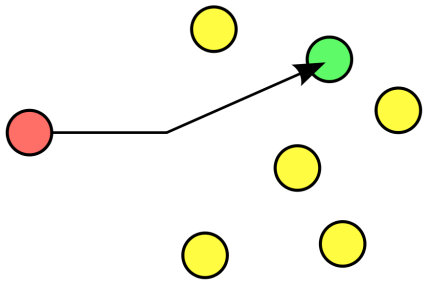


Digital music distribution models

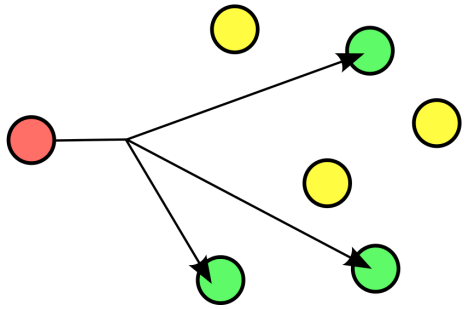
- Music stores (buy to own)
- Music tracks are delivered online
- Music locker (cloud based)
- Music streaming via subscription

Internet Radio broadcasting





Unicast: Sends IP packets to identified recipients on a network, with added bandwidth



Multicast: sends IP packets to a group of hosts on a network, with no added bandwidth, and not requiring prior knowledge of who or how many receivers there are.

Mid-term potential questions

Class 7
- Music recommendation systems
  - How does content-based recommendation work?
  - How does collaborative filtering recommendation work?
- Internet radio:
  - Differences between unicast and multicast streaming?

Class 6
- Music streaming services
  - Differences and similarities between music streaming services.

Class 5
- Music distribution
  - In the context of music sales, the year 2000 has been established as an inflection point, in which the ever-growing music industry stop its growth. Explain what happened year, and how the music industry changed.
  - How the music industry recovered from the inflection point established in the year 2000?
- Digital music distribution models
  - Provide the characteristics of the different current models of digital music distribution

Class 4
- Sound file formats
  - What is the main difference between "structured audio" and "recorded sound." Provide at least one name of a structured audio file format.
  - What is the main differences between lossy and lossless compression?
  - Provide an explanation about the perceptual model in which MPEG-3 Layer 1 is based

Class 3
- Internet technologies
  - Explain briefly how does Ethernet work
  - What is the Internet Protocol Suite (TCP/IP)
  - What is a port in the context of computer networking?
  - Can you explain the five constituent parts in this URL:
    - http://www.domain.com:1234/path/to/resource?a=b&x=y

Class 2
- Internet, WWW, and HTML
  - Explain the difference between the Internet and the web (WWW)
  - Explain what is HTML

Class 1
- The old and the new music industry
  - Explain the tensions between the new and the old music industry in terms of
    - connectivity vs. control
    - service vs. product
    - amateurism vs. professionalism