



The Internet as Data Source

Music: Legal, illegal, digital and analogue

Commissioned by:

the Dutch Ministry of Economic Affairs

Project:

2007.017

Publication number:

Date:

Utrecht, 29 January 2007

Author:

Ir. Ing. Reg Brennenraedts

Inhoudsopgave

1	Market Profile	5
1.1	Product characteristics.....	5
1.2	Market Structure.....	5
1.3	Market trends and developments.....	6
2	Existing Indicators.....	10
2.1	Existing indicators on paper.....	10
2.1.1	<i>Central Bureau for Statistics (CBS)</i>	10
2.1.2	<i>OECD</i>	10
2.1.3	<i>Trade Associations</i>	11
2.2	Existing Internet as Data Source Indicators.....	15
2.2.1	<i>Market Research Bureaus</i>	15
2.2.2	<i>Scientific Sources</i>	17
3	Where to find information.....	18
4	Locating data sets.....	19
4.1	Site centric	19
4.2	User Centric	21
4.1	Network Centric.....	22
5	Experiments.....	24
5.1	Legal content on sound carriers: Bol.com	24
5.2	Illegal content without sound carrier: Limewire	26
6	Case conclusions in the light of IaD	28

1 Market Profile

Making and listening to music are as old as civilisation and yet, these days, there is no clear definition of these notions. For the purpose of our research, we apply the following explanation: *making a deliberate succession of sounds*. This definition does not cover the entire market. Music as product or service can be offered in various forms (live, analogue, digital, on-line, off-line etc). Here we will focus solely on the *unlimited, frequent and on call desire to listen to music*. In concrete terms this means the market in which pieces of music are sold, for example through record shops in the Netherlands such as “Freerecordshop” and websites like “Bol.com”. However we will also analyse the market in MP3s and related formats. In order to focus on the case of music, we will not include radio broadcasting in our analysis.

In this chapter we will first of all look at the characteristics of the product (section 1.1) then at the market structure (section 1.2). We will conclude by identifying the trends in this market (section 1.3).

1.1 Product characteristics

Currently the majority of music on offer is digital, such as CDs, DVDs, MP3s etc. Only a very small portion is still in analogue format, for example vinyl LPs and cassettes. A typical aspect of the digital music market is that it is not affected by its use. In other words, consumers can listen to a piece of music, as often they like, without it disappearing or its quality diminishing.¹ Alongside these generic characteristics, the product has qualities that are specific to this market, which we will discuss below.

The product appears on the market in two basic forms. Traditionally music was sold by selling sound carriers. Consumers did not just buy a piece of music but also whatever carrier had the music on it. Examples of this are the CD (along with the data on the CD), the LP (along with the data on the LP) etc. Since the arrival of broadband, it has become possible to distribute data on a very wide scale. Thanks to this, music can be distributed without it being bound to a physical carrier. Nowadays many consumers no longer buy a single but download to MP3.

Illegal copying of music was done on a relatively wide scale even in the days of cassettes. Digitalisation, however, has enabled copies to be made of music files very quickly and cost efficiently. Many consumers have seized this opportunity, and this now seems to have led to wide-scale piracy. Measures to control this such as Digital Rights Management, or dealing severely with offenders have not yet had much effect.

1.2 Market Structure

The Dutch music industry has always had an international flavour. You only have to glance at the hit parade to notice that most of the recording artists are no longer from the

¹ Literally this only applies to digital audio forms. The characteristic of analogue music is that every time it is played, the quality is marginally reduced (the carrier eventually wears out). It would be going too far to state that it is therefore a product for single use. As the market is dominated by the digital domain, we consider the case to be a market product with a non-singular use.

Netherlands. Looking at the large record labels, we see a high degree of monopoly in the market. Various consolidations have led to the market being dominated by four companies as shown in Figure 1: Warner, Sony BMG, UMG and EMI.

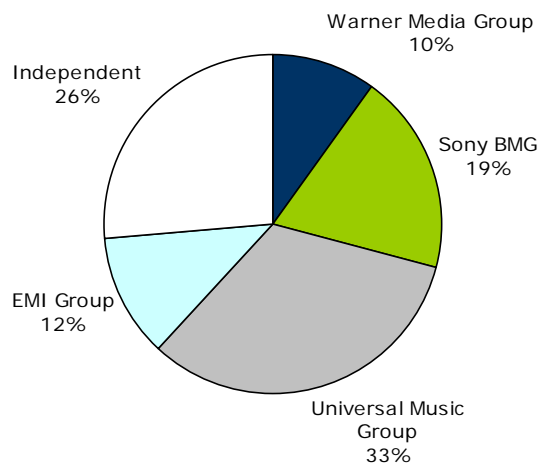


Figure 1: Distribution of the global market in music sales in 2005²

The level of regulation is limited, probably due to the fact that the market has become international and the main focus of the regulation is protecting copyrights. Initially, it is about preventing piracy but also concerns playing music in public for example in restaurants, bars, department stores and supermarkets. The trade associations that represent the industry are trying to protect these copyrights and an association has been formed in the Netherlands: Brein (which translates as 'Rights Protection for Entertainment Industry Netherlands').

1.3 Market trends and developments

The arrival of the Internet had a rather dramatic affect on the market for music and the industry itself. In the early Eighties, when music began to switch from analogue to digital, people did not realise the far-reaching consequences. At that time, the transfer from LPs and cassettes (analogue) to CDs (digital) was also prompted by the desire to improve the quality of sound and durability. Around 1990, when CD players in computers became common, this opened up the possibility to copy digital music easily. However, carrying this out on a large scale did not prove to be technically or economically worthwhile: the storage capacity (e.g. hard discs) was relatively expensive, the compression techniques (e.g. MP3) were not yet or only scarcely available, and the transfer of data was problematic; the amount, speed and quality of Internet connections were still very low.

The end of the 1990s saw the price of CD burners and CD writers fall to such a level that they became attractive for many households. In this way, people were able to 'burn' original CDs simply and cheaply at home. Once faster and cheaper Internet connections arrived on the scene, global distribution of digital music was possible. At the same time, the capacity of PCs had increased enormously. This all resulted in wide scale piracy. Initially the music industry reacted (obviously) very negatively to this illegal distribution of music through the Internet. Nevertheless, in recent years we have seen a change in its

² This diagram has been created using IFPI data

attitude towards the Internet's potential. The music industry is very much aware of the power of the Internet and now ensures that music can be distributed legally through this medium.

Above, we have outlined the choices available to today's consumers; whether they wish to buy an original CD or a copied CD, download music illegally or legally. Figure 2 illustrates these options in the form of a diagram. The cells contain examples of activities and the arrow indicates how the market is developing: over a period of time, new domains appear.

	legal	Illegal
Bound to physical carrier	<p>Sale of CDs, music DVDs, LPs in shops</p> <p>Sale of CDs, music DVDs, LPs on websites</p>	<p>Illegally 'burning' CDs, music DVDs, LPs</p>
Not bound to physical carrier	<p>Legal sale of music files through the Internet</p> <p>Distribution of music without copyright</p> <p>Unlimited distribution of music with artist's approval</p>	<p>Illegal distribution of music files through the Internet</p> <p>Illegal distribution of music files through protected networks</p> <p>Other forms of piracy</p>

Figure 2: Development and differentiation in the music market

It is interesting to note how changes in the music market have influenced the chain of values in the music industry. Figure 3 depicts the traditional chain of values in this sector. It shows that a rather substantial chain existed with a great number of players. Besides the artist was the recording studio, the producer of the physical carrier, the record company and finally the retail sector (obviously the 'legal' part of it).



Figure 3: Chain of values in the traditional music industry³

The table below shows a comparison of the new and old chains of values. The old chain of values assumed that the consumer bought a CD in a record shop. In the new chain, the

³ Based on OECD (2004) Digital Broadband Content: Music

consumer pays for downloading an MP3 file. The table clearly illustrates the declining influence of the large recording companies. The sections of the chain they used to influence have changed significantly. Moreover the barriers for entering the market have lowered considerably while the cost efficiency of the entire market seems to have greatly increased.

	Traditional	New
Creation of music	Creative process not greatly influenced by IT	Creative process not greatly influenced by IT
Recording	Large studios for recording and mixing	Traditional but also possible with PCs
Production	Warehouses where CDs are pressed and labels printed	Superfluous
Marketing	Record companies use contact with radio, TV, printed press etc.	Possibly traditional, but wide distribution possible through blogs, Youtube, websites etc.
Distribution	Complex network of national outlets	Distribution via the Internet
Wholesale	Distribution network to wholesalers	Wholesale via the Internet
Retail	CD arrives in the shop	Download from the Internet

Table 1: Shifts in the chain of values

Advances in technology have resulted in new business models for offering music. One example of this is paying to download a certain piece of digital music. Another example is Microsoft's Zune, the equivalent of Apple's iPod, in that a subscription model is applied. Customers pay a fixed monthly sum of about 20 Euros, which gives them unlimited access to a music database. Another variable is offering free music whereby the costs are covered by advertisements (the advertising model). A significant element of this is collating demographic details on the customer that are then used for marketing promotion. In exchange for watching or listening to an ad, people can download the music. Such business models all come from distribution chains in which consumers pay directly or indirectly for the music they download. The traditional record shop is one sector that has suffered tremendously from the changes in this market: its consumers get more and more of their music illegally and order their CDs more often through on-line webshops. In the UK for example there are only two large companies currently active in this retail market (Virgin and HMV) now that five companies (Andy's Records, MVC, Tower Records, Music Zone and Choices UK) have gone bankrupt in the past four years.⁴

Apart from the differences for record labels and traditional shops, the situation for artists has also changed. Many well known names have suffered negative consequences from the changes. As we will see later, the industry's turnover has dropped. Artists who work on a small scale have adapted accordingly to the new situation. It is now possible to make albums on an extremely low budget and distribute them over the whole world via the Internet. The artist earns his income in other ways such as merchandising and live performances. An interesting example of a band that became known through the Internet is the British band 'The Arctic Monkeys' – see Box 1.

⁴ See for example http://www.theregister.co.uk/2007/08/21/choicesuk_bellyup/

Arctic Monkeys - History

[...]

They began rehearsing at Yellow Arch Studios in Neepsend, and their first gig came on 2003-06-13 at The Grapes in Sheffield city-centre. After a few performances, they began to record demos and burn them onto CDs to give away at gigs. With a limited number of CDs available, fans began to rip the music back onto their computers and share it amongst themselves. The group did not mind, saying: *"We never made those demos to make money or anything. We were giving them away free anyway — that was a better way for people to hear them. And it made the gigs better, because people knew the words and came and sang along."* They themselves took no responsibility for their music, admitting that they did not even know how to get their songs onto the Internet. When asked about the popularity of the band's MySpace site in an interview with Prefix Magazine, the band pointed out that they did not even know what MySpace was, and that their fans had originally created the site. *"[When we went number one in England] we were on the news and radio about how MySpace has helped us. But that's just the perfect example of someone who doesn't know what the fuck they're talking about. We actually had no idea what it was."*

They began to grow in popularity across the north of England, receiving attention from BBC Radio and the British tabloid press. Mark Bull, a local amateur photographer, filmed the band's performances and made the music video to "Fake Tales of San Francisco", releasing it on his web-site, alongside the contents of Beneath the Boardwalk — a collection of the band's songs which he named after a local music venue.

In May 2005, Arctic Monkeys released their first EP, Five Minutes with Arctic Monkeys, featuring the songs "Fake Tales of San Francisco" and "From the Ritz to the Rubble". This release was limited to 1500 CDs and 2000 7" records, but was also available to download from the iTunes Music Store. Soon after, the band played at the Carling Stage of the Reading and Leeds Festivals, reserved for less known or unsigned bands. Their appearance was hyped by much of the music press and the band was received by an unusually large crowd for the billing they played. The critically acclaimed performance included spontaneous sing-along's of tracks that were only available as demos on the Internet.

[...]

Box 1: History of the British band The Arctic Monkeys⁵

Despite the success of some legal distribution of music through the Internet, piracy is probably very extensive. As duplication and distribution costs are next to nothing, the direct or on mutual terms distribution of music between consumers has really taken off.⁶ There are loads of ways to offer illegal files. It can be done through websites, one-click hosting applications like RapidShare, newsgroups and P2P applications. In recent years especially P2P applications such as Napster, KaZaa, BitTorrent, and Gnutella have attracted a great deal of attention. Besides being able to share information in many different ways, files come in a wide range of formats. These nearly always require a certain compression method⁷ while others enable you to protect a file against duplication, so called Digital Right Management.

⁵ Source: http://en.wikipedia.org/wiki/Arctic_Monkeys

⁶ At present it is forbidden for consumers in the Netherlands to offer illegal files but downloading is not. The question is how long this will continue. The justice department is currently investigating how to prohibit this. Source: http://www.nu.nl/news/1220075/50/Hirsch_Ballin_wil_downloaden_verbieden.html

⁷ Examples of popular compression formats are mp3, ac3, wma, Ogg Vorbis and aac.

2 Existing Indicators

In this chapter we will discuss indicators that give us insight into the music industry. Firstly section 2.1 describes indicators that are collated traditionally but in section 2.2 we will look at indicators that already use the Internet as data source.

2.1 Existing indicators on paper

Existing indicators of the music industry generally fall into two categories. The first is mostly aimed at the 'traditional market', looking at sales of traditional sound carriers. The second focuses on music that is distributed legally and illegally via the Internet. We will look at both categories, thereby using different sources: CBS, OECD and various trade organisations.

2.1.1 Central Bureau for Statistics (CBS)

The CBS holds data on developments in the retail trade including of course the outlets which sell sound carriers (SBI code 52453). This data enables us to analyse individual parts of the sector. The CBS also has details of the (possibly) illegal distribution of music. Statistics showing the activities of Internet users are published in 'De Digitale Economie 2006'. These figures are taken from a survey called POLS that investigated people's lifestyles. The survey showed that in the past three months, about 50% of Internet users have engaged in playing or downloading games and music as well as downloading images.⁸

Besides information on possible piracy, there are also details of the legal purchase of music. It appears that 25% of the Internet users interviewed ordered films or music on the Internet in the past three months.⁹ CBS also has details of the absolute and relative size of the market in digital music sales, figures they probably received from OECD.¹⁰ In 2005 this market was worth 5,000,000 Euros. Online sales only accounted for 2% of total music sales.

2.1.2 OECD

OECD has published various reports that give us insight into developments in the music market. Good examples of these are the various versions of Information Technology Outlook. They have also issued a document called "*Digital Broadband Content: Music*" which is specifically aimed at this market. In their report OECD used figures from the international trade organization IFPI to chart the scope of the retail sector. In 2003 the Dutch music market accounted for about \$500 million, meaning that Dutch people spend an average of \$31 a year. Looking at how this compares with the total global market: in 2003 this was estimated at \$32 billion. They also state that between 2000 and 2003,

⁸ Source: CBS (2006) De Digitale Economie

⁹ Source: CBS (2006) De Digitale Economie

¹⁰ They refer to the OECD – Information Technology Outlook 2006

worldwide sales of CDs dropped. The greatest drop was seen in the Netherlands, where album sales fell by 26.5% and single sales by 37%.¹¹

2.1.3 Trade Associations

NVPI

The NVPI is the trade association that represents the interests of the entertainment industry and acts on behalf of the producers of video films, interactive software and record companies. It has various market surveys carried out every year to track developments. These surveys are mostly concerning the traditional audio market. The graph below uses the NVPI's figures to illustrate the development in the Dutch audio market. It shows that the market has been halved over a 15 year period.

Development of the Dutch audio market (consumer turnover)

(€ million)

(million pieces)

turnover

volume

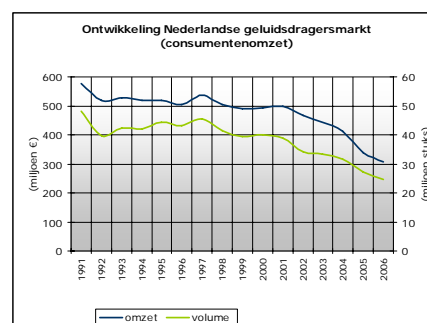


Figure 4: Developments in the Dutch audio market (consumers)¹²

Box 2 gives the NVPI's overview of the Dutch market

Mainly thanks to the sale of 10 million downloads, the number of music products sold has risen by 10.3% to 34.7 million. However, turnover fell by 7.5% to 316.9 million. Actual albums, singles and music DVDs saw their turnover drop by 7, 20 and 18% respectively. 2006 was a good year for Dutch music products. The national product share rose from 22% to 33% in December. Three of the five best-sold singles and five of the ten best-sold albums were of Dutch origin.

[....]

At 65%, the traditional entertainment business remains the major consumer market for audio products. Other outlets with a considerable share are the electronic shops (11%), Internet shops (11%) and department stores (10%). Last of all are the chemists and supermarkets (3%).

Box 2: The Dutch audio market in 2006¹³

Alongside wider developments, they also give a breakdown of the various audio products. Table 2 contains the figures over the period 1991 to 2006.

¹¹ Source: OECD (2004) Digital broadband content: Music

¹² Compiled from NVPI figures, Source:

<http://www.nvpi.nl/assets/nvpi/NVPI%20MARKTINFORMATIE%20AUDIO%202006.pdf>

¹³ Source: <http://www.nvpi.nl/assets/nvpi/NVPI%20MARKTINFORMATIE%20AUDIO%202006.pdf>

Year	LP	Music-cassette	DVDA/SACD	CD	Total
1991	9	17	-	529	555
1992	5	15	-	477	497
1993	2	14	-	479	495
1994	2	10	-	474	486
1995	2	9	-	471	492
1996	1	7	-	454	462
1997	1	4	-	492	495
1998	2	4	-	469	475
1999	1	3	-	458	462
2000	1	1	-	466	468
2001	1	1	-	461	463
2002	2	1	0.7	403	406
2003	2	0.3	5.6	335	343
2004	2	0.08	2.9	302	307
2005	2	-	2.0	262	266
2006	0.2	-	1.1	246.8	248

Table 1: Turnover sales of various audio products (in million Euros)

NVPI's figures can also show us how the price of a CD has developed, as illustrated in the diagram below.

Development of the average retail price of CD albums in Euros

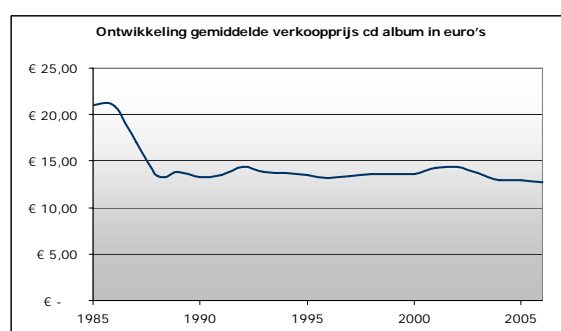


Figure 5: Development of the average retail price of a CD (in Euros)

IFPI

Besides the Dutch trade association, there is also an international organisation for the sector namely the IFPI. Just like the NVPI, they collect data and supply information on piracy. Box 3 provides a factsheet on piracy in the Netherlands.

Unauthorised file-sharing

- * Broadband penetration in Dutch households is 32%.
- * 30% of consumers download music from the Internet. That represents about 4.8 million people.

Awareness

- * 360,000 warning messages have been sent to Dutch file-sharers in 2004 and 2005.
- * Extensive media coverage since late 2003 about the illegality of file sharing.

State of the local industry

- * The Dutch market for singles has fallen by 50% in units and 54% in turnover between 2000 and 2004.
- * Sales of CD albums have decreased by 35% in 5 years, from 466 million Euros in 2000 to 302 million Euros in 2004.
- * The average price for a CD album reached an all time low in 2004 at 12.90 Euros.
- * Since 1998 the compilation market, traditionally one of the strongest strands of the Dutch recording industry, has more than halved in turnover (55%).
- * The under 20s old buy fewer CDs than they used to; in 2004 their share dropped to 19%.

Legitimate Services

- * There are several legal download alternatives in the Netherlands: Musicstream.nl, download.nl, freerecordshop.nl, itunes.nl, sonyconnect.nl, and radio538.nl
- * The Dutch Mega charts began their own download chart in February 2005.
- * Digital music sales estimated to double to around US\$2 billion in 2006
- * Single track downloads estimated up 89% at 795 million
- * Available tracks double to four million, via 500 online services in over 40 countries worldwide
- * Portable music players help drive digital music consumption
- * New revenue streams and business models emerge
- * Lawsuits impact illegal file-sharing, but "gatekeeper" ISPs must act to curb digital piracy

Box 3: Internet piracy in the Netherlands¹⁴

The IFPI has further information on the scale of the (legal) market for music that is not bound to a physical carrier. The table below illustrates that the Netherlands is one of the countries with a considerable market in this category. It also shows that most of the distribution in the Netherlands is through the Internet; distribution through mobile networks, for examples ringtones, is fairly limited here.

¹⁴ Source: http://www.ifpi.org/content/section_news/20050412m.html

	Total digital market			Digital Sales by Channel	
	US\$	Local currency		Online	Mobile
USA	636	USD	636	68%	32%
Japan	278	JPY	30,587	9%	91%
UK	69	GBP	38	62%	38%
Germany	39	EUR	31	66%	34%
France	28	EUR	22	47%	53%
Italy	16	EUR	13	31%	69%
Canada	15	CAD	18	71%	29%
South Korea	12	KWR	12,045	42%	58%
Australia	7	AUD	10	41%	59%
Netherlands	5	EUR	4	82%	18%

Table 3: The top 10 digital markets in 2005 (amounts in millions of dollars)¹⁵

In its 2007 report, IFPI gives an overview of the usage of P2P applications in Europe. The following diagram shows that relative use has declined slightly. However, this is more than amply compensated by the substantial growth in general Internet usage.

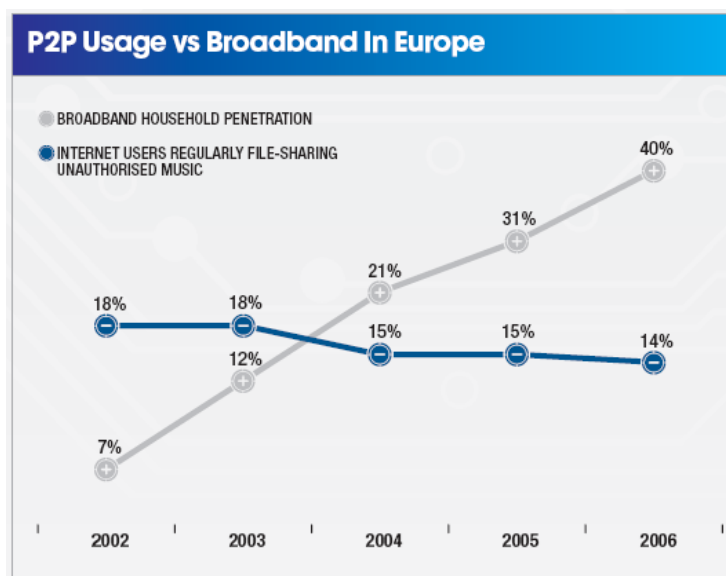


Figure 6: Usage of P2P and broadband in Europe. Source: IFPI (2007) Digital Music report

¹⁵ Compiled using figures from: <http://www.ifpi.org/content/library/worldsales2005.pdf>

2.2 Existing Internet as Data Source Indicators

2.2.1 Market Research Bureaus

Various market research bureaus are involved in analysing P2P traffic. Examples of these are: Big Champagne, NPD MusicWatch, Nielsen Soundscan, Ipoque and CacheLogic. It is not absolutely clear how they conduct their measurements but it is assumed that a number of them monitor the dynamics of some P2P programmes using a kind of spider. In fact this method of measurement is site-centric.¹⁶ Based on this method, they collect data for example on the number of times a piece of music is downloaded. They also publish figures on the total scale of the P2P market.

The following chart published in an OECD document with data sourced (very probably) from the US company BigChampagne, illustrates the use of a number of P2P networks. It shows that the use of the P2P programme FastTrack is declining while the other P2P applications are growing. It is also obvious that the quality of this data depends entirely on the number and type of other networks included in the research. Considering the fact that the total number of users of FastTrack is low, the expectation is that this does not represent P2P usage. And what is more, in the Netherlands (as far as we know), many other protocols are used besides FastTrack.

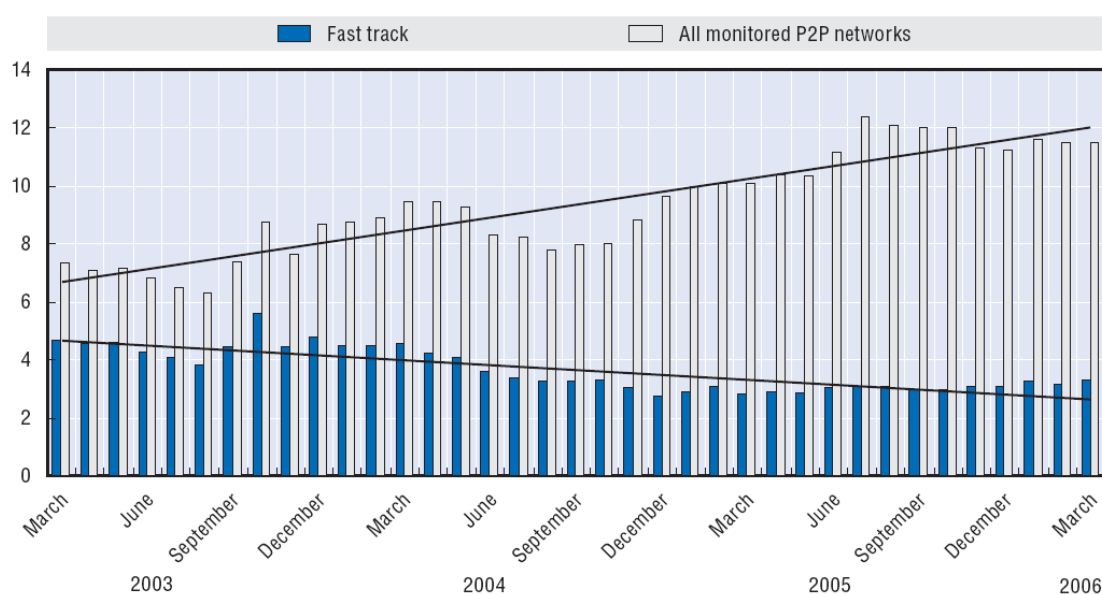


Figure 7: Global usage of P2P-networks (FastTrack and other monitored networks), measured as numbers of simultaneous users (in million Euros)¹⁷

Based on data supplied by BigChampagne, OECD has split P2P usage into national levels.¹⁸ As far as absolute numbers are concerned, we see that the USA and Germany have a very

¹⁶ In this case, measurements are made by looking at the host content. This can be done because the users' computers function as server. It would therefore be justifiable to say that this is a user centric method of measurement.

¹⁷ Source: OECD (2006) Information Technology Outlook 2006.

See: <http://www.oecd.org/dataoecd/27/59/37487604.pdf>

¹⁸ Source: OECD(2004) Digital Broadband content: Music.

large share. But if we look at relative numbers of users of P2P, Canada and the USA are at the top of the list. The Netherlands ranks round about tenth on both lists.¹⁹

	Percentage of all users		P2P users as a percentage of total population, Sept.-Oct. 2003
United States	55.4	Canada	1.2
Germany	10.2	United States	0.9
Canada	8.0	France	0.6
France	7.8	Germany	0.6
United Kingdom	5.4	Luxembourg	0.4
Italy	1.7	United Kingdom	0.4
Spain	1.1	Sweden	0.4
Netherlands	1	Belgium	0.4
Australia	0.91	Switzerland	0.4
Belgium	0.8	Austria	0.3
Sweden	0.7	Netherlands	0.3
Japan	0.7	Norway	0.3
Switzerland	0.6	Australia	0.2
Austria	0.5	Finland	0.2
Mexico	0.3	Denmark	0.2
Norway	0.3	New Zealand	0.2
Korea	0.2	Italy	0.1
Portugal	0.2	Spain	0.1
Poland	0.2	Iceland	0.1
Finland	0.2	Portugal	0.1
Denmark	0.2	Ireland	0.1
New Zealand	0.1	Japan	0.1
Ireland	0.1	Hungary	0.02
Hungary	0.1	Poland	0.02
Greece	0.1	Greece	0.02
Luxembourg	0.04	Korea	0.02
Czech Republic	0.04	Czech Republic	0.01
Turkey	0.03	Mexico	0.01
Slovak Rep.	0.01	Slovak Rep.	0.01
Iceland	0.01	Turkey	0.00
OECD countries	96.9	OECD average	0.24

Figure 8: Use of P2P as percentage of the total P2P usage (left) and number of users per head of the population (right)

OECD's Information Technology Outlook 2004, again based on BigChampagne data, gives a more in-depth overview of P2P usage. The following diagram shows that in most countries, music files have the largest share (that is the set of files the users refer to on their PC as 'downloadable').

See <http://www.oecd.org/dataoecd/13/2/34995041.pdf>

¹⁹ It is very noticeable that Korea (and to a lesser degree Japan) score so low. Especially because South Korea has the highest number of fast broadband connections in the world. We suspect that this research did not measure the most popular Korean P2P application: Soribada.

Figure 5.6. File breakdown for OECD countries based on peak simultaneous KaZaA users, September-October 2003

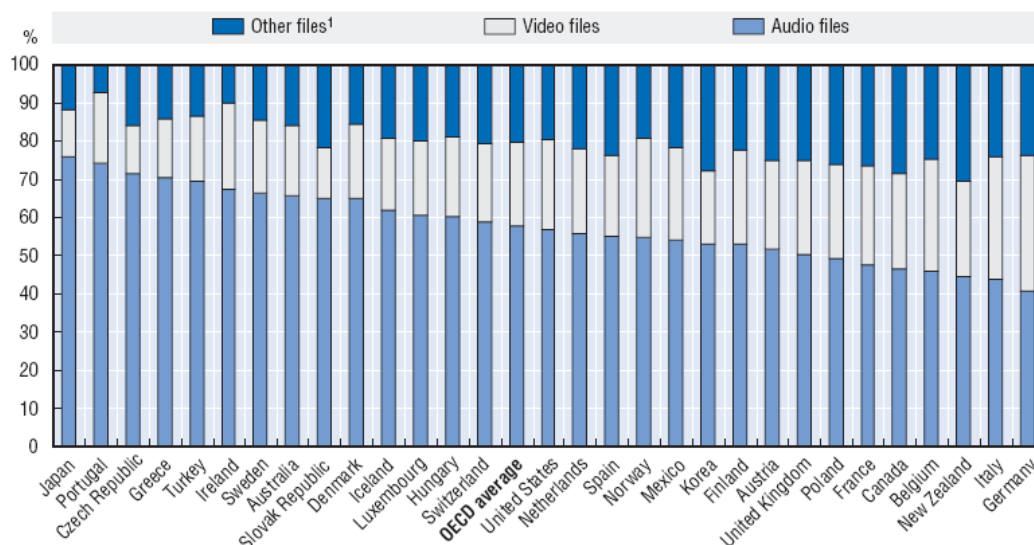


Figure 9: Specification of files shared via P2P

The German company Ipoque made an analysis of Internet traffic using network centric measuring. Their results highlighted that P2P accounted for between 50 and 90% of the traffic. The most used P2P applications were BitTorrent and eDonkey.²⁰ The company Ellacoya also used network centric measurements (such as deep packet inspection) on Internet traffic. They showed that P2P accounted for 37% of Internet traffic in the USA.²¹

2.2.2 Scientific Sources

Only one publication has been found on the usage of P2P applications. It stated that there is a relatively large number of Dutch users of Kazaa.

Country level domain	Occurrence	Country level domain	Occurrence
nl	1026 (25.84%)	be	115 (2.89%)
com	866 (21.81%)	ch	107 (2.69%)
net	527 (13.27%)	at	106 (2.67%)
se	281 (7.07%)	fi	89 (2.24%)
fr	148 (3.72%)	de	82 (2.06%)

Figure 10: Top level domains of about 4000 Kazaa users.²²

²⁰ Source: <http://arstechnica.com/news.ars/post/20070903-p2p-responsible-for-as-much-as-90-percent-of-all-net-traffic.html>

²¹ Source: <http://arstechnica.com/news.ars/post/20070619-the-youtube-effect-http-traffic-now-eclipses-p2p.html>

²² Lukas Hämmerle (2004) *P2P Population Tracking and Traffic Characterization of Current P2P File-sharing Systems*. <ftp://www.tik.ee.ethz.ch/pub/students/2004-So/MA-2004-04.pdf>

3 Where to find information

In order to form a good overview of this sector, we will use the same generic division as in other case studies. On the horizontal axis are the steps the consumer undergoes to obtain a purchase: information, ordering, payment and receipt of the product. The vertical axis contains the categories B2B, B2C and C2C. The diagram below illustrates the great extent of compression, however the nature of this market ensures that there is not a single point where *all* the data converge.

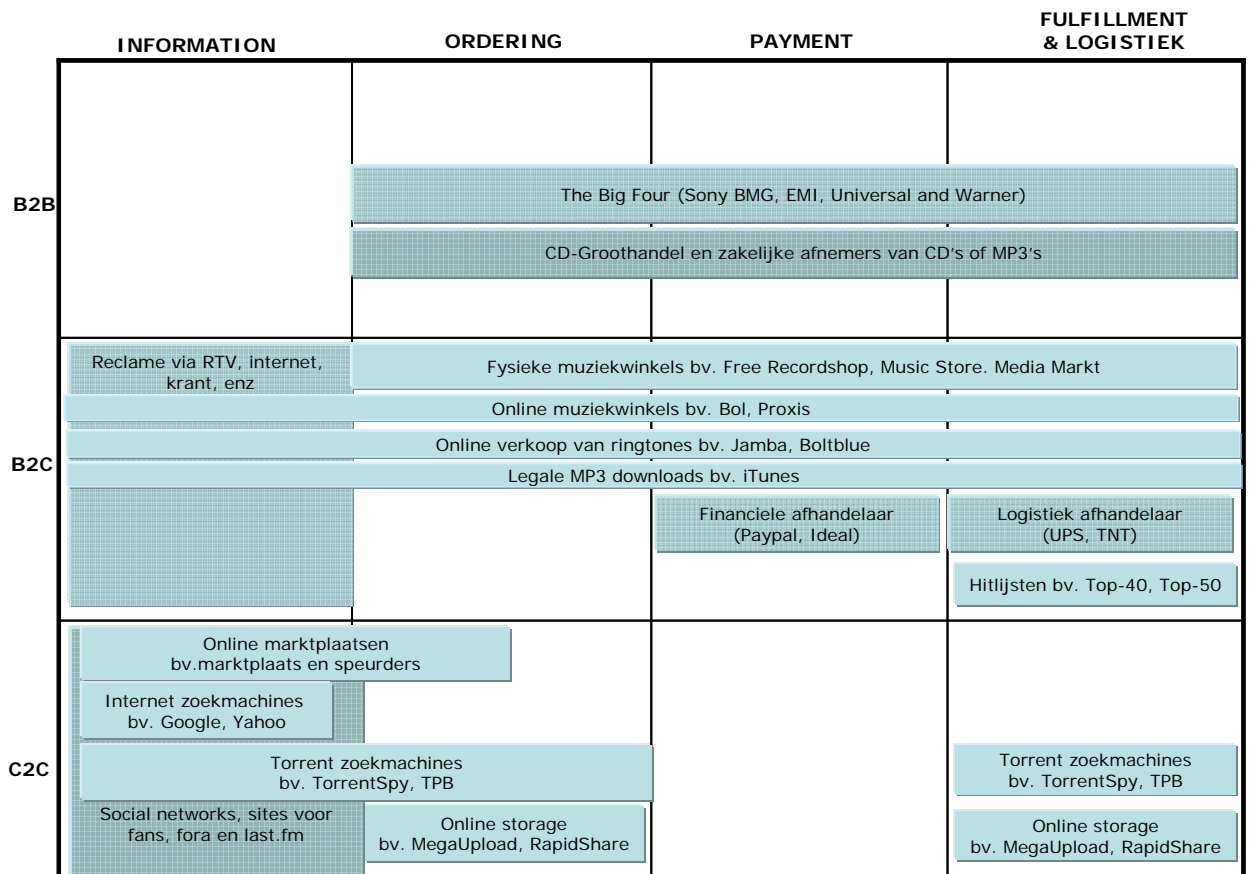


Figure 11: Compression points of information in the music market

4 Locating data sets

In order to locate data sets, we use the matrix that we designed at the start of this case study. We distinguish between legal/illegal and bound/not bound to a physical carrier. As we are focussing on locating data sets, we also further distinguish between site centric and network centric measurements. By combining these, we can form a framework of 12 cubes (see Figure 12), which we will use to locate data sets.

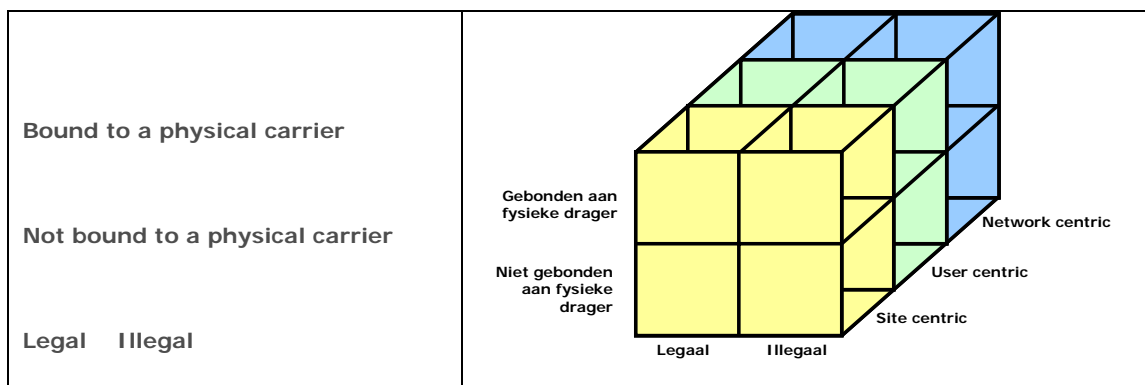


Figure 12: Framework for locating data sets

4.1 Site centric

If we use site centric measurement, it is possible to fill all the cells in the cube with interesting data sets. If it concerns the legal content, which is dependent on a physical carrier, it makes more sense to use webshop sites. One particular website springs to mind, that of the Netherlands' biggest sellers of CDs: Bol.com. The data set on the 'back' of this website contains interesting information on the size and content of what is on offer. By taking a longitudinal measurement, trends can be distinguished.

One way to chart the consumer-to-consumer trade in sound carriers is by utilising a website of consumer-to-consumer market places such as Marktplaats.nl and Speurders.nl. It is important to realise that we are talking about both legal as well as illegal content. But trying to distinguish between these categories from an advertisement on a website is no easy matter. Here too we can analyse the quantity and quality of products. Consider for example the variety of CDs, DVDs and vinyls on offer as well as average prices and the relative scope of a category.

If we try to locate data sets in the illegal circuit that are not bound to a physical carrier, we come up against an obvious problem. Nowadays, you can be prosecuted by copyright organisations and the police for offering data on illegal music files. This happened recently with the website Oink (see box 4).²³ Initiatives like Napster and Kazaa have been dealt with severely and they now no longer exist or are hardly active at all.

²³ As far as we know, Oink did not offer content but links to illegal content, which could be downloaded via BitTorrent.

Police take 'pre-release-site' off the air

AMSTERDAM – The British and Dutch Police have taken Oink off the air. The site specialised in the distribution of albums that had not yet been released. A 24 year-old man has been arrested.

This was announced by the IFPI, the international organization representing the record industry. Oink was the main site for downloading music before it was officially released. The music was usually leaked demos or trial recordings.

The site, which had roughly 180,000 users, was only accessible by invitation. New members were only welcome if they could offer music themselves. This year, 60 albums were put on the Internet via Oink before their release.

Servers. In Middlesbrough, England, the police arrested a 24 year-old site employee. He is suspected of piracy among other things. The Dutch police have seized several servers in Amsterdam.

Brein Tim Kuik, director of the anti-piracy association Brein, was pleased to see the authorities doing the right thing by clamping down on this type of site. He said, "The cumulative damages caused by the site are enormous. The Netherlands is known as a safe haven for illegal sites due to a number of hosting providers who bend the rules. That has to change."

(c) NU.nl/Wieland van Dijk

Box 4: An example of a site that has been taken off the air²⁴

Despite the fact that it is difficult to measure illegal activities, there are still a few websites offering illegal content or links to illegal content. Examples are the so-called Torrent Trackers. These are servers that enable clients to search each other's content and connect to each other. You can usually access them on the Internet and view the content they 'offer'. It should therefore be possible to use this data set by logging onto the site. Of course we must bear in mind that this is only one type of P2P programme. A well-known example of a Torrent Tracker is ThePirateBay.org.

Besides obtaining music through P2P applications, it is also simple to download illegal and legal music via the web, using standard search engines such as Google. Users only have to enter certain search details that let them obtain files very easily.²⁵ This is making use of what are called open directories. Applying this method, we could use the whole 'world wide web' as data source. It would then be especially interesting to measure how long a (Dutch) user is allowed to offer illegal content in this way. It goes without saying that this simple method would obviously be recognised by the anti-piracy movement.

If we turn to the legal section and music which is not bound to a physical carrier, it is interesting to look at sites that offer music legally (for payment). The most familiar example is probably iTunes, but there are some Dutch suppliers of these products such as PlanetMusicStream, Surf2Music and the FreeRecordShop. By using the database 'behind' these websites, you can gather various interesting indicators such as prices, selection, quantities etc.

Apart from the above-mentioned ways of offering music, one model is often overlooked. This concerns people who make their own music and offer it free of charge. The main platform for this is MySpace. Although it is usually not the idea, it is relatively simple to download from MySpace. But even though the music can be downloaded, this site is still within the boundaries we set. Currently hundreds of Dutch artists offer their music on

²⁴ Source: http://www.nu.nl/news/1284615/54/Politie_haalt_'prerelease-site'_uit_de_lucht.html

²⁵ The usual way to do this on Google is to enter a number of conditions the site has to comply with then the name of the performing artist or piece of music.

MySpace.²⁶ It could be worthwhile to look at the patterns of what is on offer and the profiles of the different users. These figures could give us insight into what free music is being offered.

As a considerable amount of sources have been mentioned in this section, we will supply you with an overview. Once again we have applied the matrix, which we defined earlier and the overview is in Figure 13.

Site Centric

		legal	illegal
Bound to physical carrier	to	<ul style="list-style-type: none"> ▪ Webstores for CDs (such as Bol.com) ▪ Trade in legal CDs via Internet (such as Marktplaats & Speurders) 	<ul style="list-style-type: none"> ▪ Trade in illegal CDs via Internet (such as Marktplaats & Speurders)
	Not bound to physical carrier	<ul style="list-style-type: none"> ▪ Webstores for MP3s (such as iTunes, Planet) ▪ Music offered on MySpace 	<ul style="list-style-type: none"> ▪ Torrent trackers (such as The Pirate Bay) ▪ Illegal offers via websites (open dir)

Figure 13 An overview of the data sources which can be accessed by site centric measurement

4.2 User Centric

Regarding **user centric** methods, we do not see how they can be used to measure music, which is bound to a physical carrier. However, we do know they can be used to measure music that is not bound to a physical carrier. We think that by inserting a kind of speaker (via hardware or software) it will be possible to monitor users' behaviour. In this way we could see how (often) they use legal or illegal methods of downloading music. But if we start generalising the data, we are likely to encounter problems when selective non-responses occur. Especially users who download a great deal will be hard to reach because they surely do not like someone 'listening in'.

Another method of user centric measuring could be developed with P2P applications. Some of these already let you view other users' hard drives.²⁷ This would enable us to have details of what is on offer. Figure 14 illustrates these possibilities in a matrix.

²⁶ Examples of these are: zijn <http://www.myspace.com/prettyparacetamol> and <http://profile.myspace.com/index.cfm?fuseaction=user.viewprofile&friendid=158911432>

²⁷ In fact you can only view that part of the hard drive a user has shared with others. This is also referred to as the share.

User Centric

	legal	illegal
Bound to physical carrier	<ul style="list-style-type: none">▪ No sources identified	<ul style="list-style-type: none">▪ No sources identified
Not bound to physical carrier	<ul style="list-style-type: none">▪ Monitoring via 'listening in' box at user	<ul style="list-style-type: none">▪ Monitoring via P2P applications▪ Monitoring via 'listening in' box at user

Figure 14 Overview of data sources that can be accessed by user centric measuring

4.1 Network Centric

Network centric measuring is an interesting way to chart (especially illegal) activities. We can see how deep packet inspection (DPI) enables us to look at the amount of P2P traffic in a data stream. Deep packet inspection means that a computer analyses the characteristics of a particular stream. Thus we can get a picture of the relative amount of P2P in a certain stream at a certain time and possibly obtain even more interesting data from this stream. There are indications that details can also be obtained about downloaded music files. At the moment it is not clear how 'deep' we can look into the stream.

When we actually get down to carrying out these measurements, the choice of concentration point is crucial. It has to be a stream that is typically representative of the (Dutch) population. Company and university networks for example would not be suitable. ISPs could be an interesting alternative. The questions we can pose while carrying out this exercise include:

- What is P2P's share of the total traffic?
- How will P2P develop in time?

If we are able to look more deeply into the stream, we can answer the following questions:

- What is the share of the individual P2P applications?
- What share do music files have in the stream?
- Between which countries do P2P streams flow?

Network centric measurement of P2P traffic can be problematic. First of all it is difficult to generalise the data from one single stream for the whole of the country. If we opt for an ISP data stream, we have to leave out company and academic networks. This is disconcerting because there is ample evidence that university networks are being abused for P2P applications.

Another consideration for network centric measuring is that absolute measurements of the size of the stream are not necessarily significant. As packages via the Internet can reach their goal via diverse routes, it is not always clear what the size of a particular stream signifies. For example this can be higher due to a malfunction in another network.

A further problem is the fact that P2P technology is used for completely different (legal) appliances (for example television, see www.joost.com). The total amount of P2P does not tell us which part is music or the ratio legal : illegal. However, deep packet inspection should be able to solve this problem.

The last problem to confront us could be the considerable calculations required for this process. Certainly if deep packet inspection is to be carried out in a large stream, a great deal of calculation has to be done. If the capacity for calculation is too low, the network becomes slower especially at peak times. Obviously this is not acceptable for ISPs or any other network controller.

Alongside DPI, in our opinion there are no other interesting data sets available for monitoring P2P traffic using network centric measurement. The matrix is also fairly straightforward, see Figure 15.

Network Centric

	legal	illegal
Bound to physical carrier	<ul style="list-style-type: none"> No sources identified 	<ul style="list-style-type: none"> No sources identified
Not bound to physical carrier	<ul style="list-style-type: none"> Monitoring P2P via DPI 	<ul style="list-style-type: none"> Monitoring P2P via DPI

Figure 15: Overview of data sources that can be accessed using network centric measurement

5 Experiments

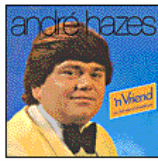
5.1 Legal content on sound carriers: Bol.com

In this experiment we are trying to trace the development of the market in legal content on sound carriers. In this case we will use the website Bol.com and its database. We can obtain the Bol.com database in two ways: by using spiders or by requesting parties to give us a copy of their database. By regularly examining these parties' entire music selection, we can form an interesting analysis. The next page illustrates the extent of information Bol.com provides about its product. Based on these details, we can consider the following:

- What is the average price of a CD?
- What is the average delivery time for a CD?
- How many CDs are available (long tail)?
- How many CDs are available in a certain category?
- How many people have commented on this CD?
- Are there certain networks for certain tastes in music?

The following sheet is an example of a CD for sale on Bol.com.

Een Vriend



Andre Hazes

Adviesprijs: €11,49
bol.com prijs: € 7,99
U bespaart: 30 %

5-10 werkdagen
CD | Emi | 28.01.2001
EAN: 0724353144425

Beoordeling van klanten - [Lees de beoordelingen](#)

[Vertel wat je van deze cd vindt en maak kans op € 100,- aan cadeaubonnen.](#)

Tracklist:

Klik, voor zover aanwezig, op de nummers met een muzieknoot om naar een muziekfragment te luisteren.

1. ['n Vriend \(4:57\)](#)
2. [Koningin Van De Zigeuners \(3:55\)](#)
3. [Waarom \(4:00\)](#)
4. ['n Ander \(3:17\)](#)
5. [Spijt \(4:43\)](#)
6. [Het Is Koud Zonder Jou \(4:06\)](#)
7. [Wat Is Dan Liefde \(3:36\)](#)
8. [Droom Maar M'n Jongen \(3:33\)](#)
9. [Alleen \(3:00\)](#)
10. [Vertel Me De Waarheid \(2:54\)](#)

Liefhebbers van deze cd bestelden ook:



[Gewoon Andre](#)
Andre Hazes
€ 7,99



[Classic Bachman-Turner Overdrive: The Universal Masters Collection](#)
Bachman Turner Overdrive
€ 7,99



[Hollands Corry & Corry](#)
Corry & Corry
€ 7,99

[Meer verwante cd's](#)

Liefhebbers van deze artiest bestelden ook:

- [Robbie Williams](#)
 - [Frans Bauer](#)
 - [Jan Smit](#)
 - [Blof](#)
 - [Guus Meeuwis](#)
 - [Anouk](#)
- [Meer verwante artiesten](#)

Extra informatie:

Dit product komt voor in:

[Nederlands](#) / [Pop](#)
[Pop](#) / [Diversen](#)

Extra informatie:

10 Track(s) | 16 Bits | JewelCase
Uitvoerend(e): [Andre Hazes](#)

Beoordeling van klanten:

[Vertel wat je van deze cd vindt en maak kans op € 100,- aan cadeaubonnen.](#)

9 juni 2005 Best
Een Vriend

Echt een geweldige cd, mooie nummers, vooral het nummer :Waarom vind ik een supernummer(is mijn favoriete). qegroet Niels

Figure 16: Details available about a CD on Bol.com

If we conduct measurements in this way, we can expect to encounter a number of obstacles. First of all Bol.com might not be too keen on our approach. They may consider their database as strategic information that should not become public. If so, they will not

make their database available voluntarily. One solution would be to access the site via a spider. But then you run the risk of a battle between spider builders and webmasters.

Another obstacle we have encountered is (again) to do with the generalisation of data. This happens mainly because there are so many suppliers of CDs and because the market is international. In concrete terms, this means that you do not know whether the collected data refer to the Dutch situation, the Flemish situation, or another situation. Moreover it is very difficult to judge Bol.com's share in a particular market.

5.2 Illegal content without sound carrier: Limewire

In our second experiment we try to map the development of illegal content without a sound carrier. We use the P2P programme Limewire as this makes it easy to see which files other parties share within the Gnutella network.^{28 29} For this we looked at 26 Dutch IP addresses. These 26 people shared all sorts of music files from Marco Borsato to André Hazes. They shared an average of 897 files with a standard deviation of 1278. A few of the users had very high numbers of files (1000 to 6000), while most of them shared about 300. The shared files showed there was a distinct preference for music, about 90% being audio files (MP3, wma, ogg). Only 5% were video files (wmv, avi, mpeg, divx) and the remaining 5% were other file formats (exe, xls, word, zip etc). When we look at the size of the files, we see a different ratio because video files are usually twice as large as music files. These 26 IP addresses were (according to the users) connected to the Internet via ADSL/cable (7x), T1 (3x) and T3 (17x).³⁰

These figures show us which music and media files Dutch people share. However, the data are limited by a number of factors. Firstly, the user can indicate which folders he is willing to share and, in this way, keep certain collections for himself. Secondly, the Limewire system only enables us to look at PCs in homes (most other P2P systems do not have this option). Thirdly, only users with a good connection will use these P2P networks. Fourthly, it is difficult to generalise this data.

Examples of data we can collect via Limewire include:

- How many files do users share via Limewire?
- What type of files do they share?
- How quickly does an average Limewire user's collection grow?
- From which geographical locations do users share their files?
- Who are these users' providers?

²⁸ See for example: Schollmeier & Kunzman (2003) *GnuViz – Mapping the Gnutella Network to its Geographical Locations*. <http://home.as-netz.de/gak/geggi/gnuviz.pdf>

²⁹ See for example: LT Nguyen, WG Lee, D Jia, O Frieder (2007) A Tool for Information Retrieval Research in Peer-to-Peer File Sharing Systems. http://www.cs.uiuc.edu/homes/hanj/refs/icde07/content/data/205_D3-8-176-Nguyen.pdf

³⁰ T3 connections are mostly used by companies and universities and deliver speeds of 45 megabit/second. T1 connections are usually purchased by smaller institutes and companies and deliver speeds of 1.5 megabit/second. ASL and cable connections are of course used at home and offer speeds of 256 kilobit to 8 megabit/second.

The disadvantages of this type of measurement are to do with the fact that there is only a limited number of users on the Gnutella network and there is no correct estimate of the total numbers of (Dutch) Gnutella users.

6 Case conclusions in the light of IaD

Interesting conclusions can be drawn from this case. First of all the government's involvement in this market is very small, which is in sharp contrast with other markets such as housing or pork. We cannot find any concentration points that have been suggested by the government. In the housing market, this could be the land register and in the pork industry it could be the Ministry of Agriculture. As no concentration points have been dictated from higher up, it is difficult to make really exact measurements in this market. Besides, there is a high level of internationalisation and it is apparent that this impedes the measurement of the Netherlands' market share. Concerning market concentration, we see that in the top layer (record companies) the concentration is high. Unfortunately these are international players. Other sections of the chain of values are more national, but here the concentration is lower.

Apart from the nature of the market, the characteristics of the product also play a role. The music sector has become a market in which digitalisation is the norm. Thanks to this, it is very easy and cheap to produce and copy the product. It seems a bit absurd that measuring developments in this market is so problematic. The digital nature of the product is that very quality that enables such a huge scope of data streams, legal or illegal. It is just not feasible to measure *all* the streams.

Despite the above-mentioned problems, the music market does seem to be interesting from the point of view of the Internet as data source. Especially site-centric measurement is a relatively simple way of revealing data sources. User centric measurements appear to be complicated, and our conclusion is that network centric measuring is definitely very complex. Nevertheless, every time we carry out measurements, we are able to extract a wealth of information.

On the other hand, we should also realise that the problems we have to deal with are the same ones everyone else is facing. A few commercial market research companies claim to have good data but in practice these prove otherwise. Due to the fact that we do not know which methods they are using and the great differences in the methods applied by various players, the reliability is dubious.

It is true to say that there is no reliable information to be found on many segments of this market, and even finding reliable information generally is very difficult. It is therefore advisable to focus on a very limited number of data collection methods. Although this may be a complicated process, the results could prove to be really worthwhile. We also recommend gradually shifting the focus away from the Dutch situation as this market has now become very international. Only looking at the Netherlands would be doing this market an injustice.



Contact:

Dialogic
Hooghiemstraplein 33-36
3514 AX Utrecht, Netherlands

Tel. +31 30 215 05 80
Fax +31 30 215 05 95
www.dialogic.nl

