Chapter 2 - Literature Review

**Thesis title:** An analysis of the relationship between individuals’ perceptions of privacy and mobile phone location data - a grounded theory study.

**Andrea Gorra**, Leeds Metropolitan University, UK
Comments sent to a.gorra@leedsmet.ac.uk would be most appreciated.

Chapter 2  Literature review: Privacy, Mobile phone technology and Legal framework

When examining the subject area of communications data retention broadly three areas need to be taken into account: a) technological matters, b) social values such as privacy and civil liberties, and c) law enforcement issues (Whitley and Hosein, 2005; Raab and Bennett, 2006). Accordingly this chapter has been divided into three parts; the first part focuses on the role of geographical location within mobile phone technology, the second part discusses the relevant privacy related literature, and the final part of this chapter examines the legal framework relevant to mobile phone location data.

Grounded theory methodology has been used to guide collection and analysis of empirical data for this study (for more detail see Chapter 3 - Research Methodology). GT methodology is an inductive method, which means that typically only a very broad review of literature takes place before collecting the data.

However in addition to this initial review, the literature is visited again after the data collection has been completed. This study has followed this approach with the benefit that a second review of the literature could be guided by findings from all three empirical data collections - pilot study, interviews and survey. Particularly the development of the final GT categories, which was supported by written and visual memos, has prompted the review of further literature.

The literature discussed in this chapter is a result of this iterative process and some literature particularly relevant to the findings from study will be revisited in Chapter 5 and brought into relation to the empirical findings.
2.1 Part 1: Mobile phone location data

This section presents essential technical background to mobile cellular networks in order to provide an understanding of the role of location data. A brief history of mobile phone telephony is provided and followed by an explanation of the relevance of the cellular network architecture to mobile phone location data. Following this a technical overview of mobile communications network architecture is provided and finally location-based services and techniques of locating a mobile phone are discussed.

2.1.1 Background to mobile phone location data

The mobile phone is a key technology in an increasingly mobile and connected world. Growing technological convergence and ubiquitous networking leave behind a continuous and lasting trail revealing information about those involved in the communication. Picking up on this electronic trail makes it possible to identify the location of communication devices and individuals are indirectly locatable by carrying their mobile phone. In some cases, the location information supplied by a mobile phone can be very specific, depending in cell size and cell shape, which will be explained later in this section. Location-based services (LBS) for mobile phones (see section 2.1.3) are predicted to grow in the near future (see Lyon, 2005). A market research firm predicts that LBS revenues will reach €622m by 2010 (Moore, 2006).

The mobile phone as a location technology brings together the following three key features:

1) identification of the approximate location of the handset
2) on a continuous basis
3) in real-time

The combination of these features means that mobile phone users can potentially have their geographical location and movements traced at any time or all the time. In addition, a log of all data generated by a mobile phone is stored by the mobile phone service provider, and potentially shared, as discussed in the final part of this chapter (2.3 Part 3: Legal framework relevant to mobile phone location data). This makes the mobile phone unique in comparison to other location identifying technologies, such as CCTV or RFID, particularly as a mobile phone tends to be carried by one person on a regular or sometimes even continuous basis. CCTV for
example can track individuals in real time but not continuously, RFID tags can also reveal location information but tend not to do this in real-time or on a constant basis. WiFi networks can also be used to locate the position of the device connecting to the wireless network (Hosein and Escudero-Pascual, 2002). However, this is limited to the size and the geographical area covered by this network.

Location data can be perceived as both a threat and an opportunity. There are many opportunities that the attribute location can offer to businesses, governments and individuals. However, knowing a person's location at any given moment in time can also have implications for privacy, civil liberties and social justice (Clarke, 2000; Lyon, 2005).

2.1.1.1 History of mobile phone telephony

The development of wireless communication systems started in the 1930s with the use of 'Walkie-talkies' during the Second World War to enable foot soldiers to stay in contact with the headquarters (Elliott and Philips, 2004). In 1946, AT&T Bell introduced the first commercial radiotelephone service in the US, which allowed communication between mobile users in cars and the public fixed network. In the 1960s, Bell Systems launched the Improved Mobile Telephone Service (IMTS), which laid the basis for commercial-sector mobile communications. Developments in microprocessor technologies in the late 1970s and early 1980s enabled the introduction of the reliable wireless communications system, the so-called first generation.

First generation network - 1G

The first generation wireless technologies, also known as 1G, were relatively simple and used analogue signals. Mobile phone handsets based on 1G technology were mainly used by government agencies and the military before this technology came into general use in the business domain in the 1980s (Elliott and Philips, 2004). The systems in Europe and the USA had in common that they provided coverage of a very large area by using only one transmitter mast. The coverage area of a mast was fairly large, up to 150km, and required minimal infrastructure. In order to connect via large distances, the base station as well as the mobile phone had to transmit simultaneously at high power. This meant that the mobile phones were larger than today's handsets and used to be built into car boots. Moreover, due to the limited number of available frequency channels, only a small number of
subscribers could be connected to the mobile phone network (Walke et al., 2003). 1G systems were based on analogue signals which are radio transmissions sent in a wave-like form. The mobile device sends the waves to a base station where the signal is reconstructed as accurately as possible and relayed to its destination. Noticeable differences in quality occur due to errors recreating the signal wave. In addition, analogue signals are relatively easy to intercept, as they are transmitted in the clear (Deitel et al., 2002).

**Second generation networks - 2G**

In the late 1980s and early 1990s, the popularity of wireless communications grew and increased the demand for network capacity. Together with the disadvantages of analogue 1G systems, this led to the development of the second generation wireless system based on digital technology. Digital signals have different transmission properties than analogue signals and use binary coding using sequences of 0s and 1s to construct a signal's unique pattern. Digital signals use digital samplers and codecs to convert analogue voice data into digital data. Digital signals can be precisely duplicated by the receiving base station and send to its destination. This process results in a lower error rate than analogue transmission correction which results in clearer voice reception (Deitel et al., 2002). In addition, digital traffic is relatively simple to encrypt in order to prevent eavesdropping (Stallings, 2005).

GSM, the Global System for Mobile Communications, fundamentally differs from the 1G system because of its use of cellular network architecture, which will be explained in subsequent sections. GSM, also known as second generation network or 2G, was first developed in the 1980s through a pan-European initiative, involving the European Commission, telecommunications operators and equipment manufacturers. GSM is an open non-proprietary and interoperable digital standard for cellular mobile systems operating in the 900 and 1800 MHz band. In 1986, a number of different prototype systems put forward by companies and consortia from different European countries were trialled and led to the agreement of the main characteristics of the new system (Steele et al., 2001).

GSM is still in use to date by all European countries and has also been adopted in other continents, such as Africa and South America. There are over 540 million GSM subscribers in Europe, plus another 18 million Europeans using 3GSM networks, which are the 3G service delivered over the evolved GSM core network.
(GSM Europe, 2005). With GSM it was also made possible to send and receive limited amounts of data via the Short Messaging Service (SMS) and mobile internet browsing via the wireless Applications Protocol (WAP) (Elliot and Philips, 2004).

**Second and a half generation networks - 2.5G**

2.5G technologies represent a state of development between 2G and 3G and have overcome the limited data and primarily voice-centred services of the 2G networks. In the 1990s and early 2000s higher transmissions rates and always-on connectivity were enabled by General Packet Radio Services (GPRS). Data transmission speeds were now 10 times faster with 115kbits per second and based on packet-switching technology (International Telecommunication Union, 2003). Packet switching optimises the use of bandwidth available in a network and minimises the time it takes for data to travel across the network. The increased data transmission rates of 2.5G compared to earlier systems help to transfer data such as mobile internet content (Elliot and Philips, 2004)

**Third generation networks - 3G**

Third generation mobile telephony (3G) is the successor to the 2G and 2.5G systems. 3G improved previous systems by providing enhanced security and encryption features, improvements in screen displays and the ability to handle multimedia data, such as graphics and video streaming. 3G allows faster data exchange with data transmission rates up to 1920kbits per second, which enables the support of greater voice and data customers. Support can be provided for a wide variety of mobile equipment. 3G technologies were first introduced in Japan in 2001 and spread to Europe and the USA in 2002. UMTS (Universal Mobile Telecommunications System) is the third generation mobile phone technology mainly used in Europe and also in Japan. It uses the GSM infrastructure and UMTS/GSM dual-mode phones sold in Europe are able to make and receive calls on both networks. Elliott and Philips (2004) describe as aims of all 3G networks the following:

a) world-wide connectivity and roaming throughout Europe, Japan and North America
b) high data transmission rates and broad bandwidth, suitable for multimedia content
c) efficient spectrum utilisation

(Philips and Elliot, 2004)
2.1.1.2 The cellular concept

In terms of location data the most crucial concept was the introduction of the cellular network architecture with the change from the 1G to the 2G system. In 1972, a patent was registered by Bell Labs, the research subsidiary of the US telephone company AT&T that laid the foundations for today's second and third generation mobile phone systems (US Patent full-text and image database, 1972). The essence of this patent was to reduce the area covered by an antenna. In the 1G system, only one antenna served a very large area, whereas in the 2G this area was broken down into several cells.

Each area or cell is served by a single base station and is in the approximate centre of each cell. Due to the increased number of cells in the 2G system compared to 1G, many more masts were needed and this enormously increased the costs of infrastructure. Now it was possible to reuse the same frequency over relatively small distances, enabling coverage for a greater number of subscribers. This resulted on the one hand in the reduced size of antennas, and on the other hand in reduced distances between antennas. Another effect of this new system was that the long-distance transmissions between mast and handset were not necessary anymore, which resulted in decreasing sizes of mobile phone handsets and improved handset operating times. Adjacent cells cannot use the same frequency, as this could result in interference. Hence, it is necessary to only re-use frequencies in cells that are sufficiently distant to each other. The transmission power of a base station has to be limited to prevent interference with frequencies from other cells and to reduce health concerns (Stallings, 2005).

This concept of using an increased number of smaller cells required a mechanism to switch a user's connection from cell to cell. Hence, cells were arranged in a hexagonal pattern to enable all antennas to be the same distance apart, which is beneficial when a mobile phone user moves within a cell towards its boundary. This makes it easier to determine when to switch a user to an adjacent antenna and which antenna to chose. In comparison a square pattern, would only provide an equal distance to the centre of four cells (see Figure 2.1).
In practice, cells have irregular shapes and a precise hexagonal pattern is often not used due to topographical limitations, restrictions on positioning antennas and different amounts of traffic volumes as some cells have more mobile phone users than others. In addition, the coverage can overlap in looser shaped cells (Philips and Elliot, 2004). Different shaped cells have an impact on the accuracy of mobile phone location tracking based on the Cell-ID, as smaller cells provide a more accurate result than larger cells.

Figure 2.1: Contrasting hexagonal and square pattern (after Stallings, 2005)
2.1.1.3 Increasing the capacity of a mobile phone network

There was a considerable increase in the proportion of households with a mobile phone since 1998-99 from 27 percent to 78 percent in 2005 (National Statistics (2005). The increasing number of mobile users has caused the necessity to enhance the network capacity. According to the Mobile Operators Association (2006), base stations can only carry a maximum of around 120 calls at the same time.

A mobile phone network can be made suitable for an increased number of phone users in several ways:

- Add new channels (if possible)
- Frequency borrowing from adjacent cells
- Cell sectoring
  - divide cell into wedges, each with own channels (usually 3 or 6 sectors per cell)
  - each sector has its own subset of channels
  - the base station has to use directional antennas
- Cell splitting
  - generally cells are 6.5 to 13km in size
  - power level of antenna is reduced to keep signal within cell
  - since the cells are smaller, more frequent handovers (i.e. transfer of call from one Base transceiver to another) have to take place
- Micro cells
  - Antennas move lower down, e.g. from top of buildings to smaller buildings or lamp posts
  - Cell size is decreased and a reduction of power takes place
  - Micro cells are usually in place on the high street where great traffic volume can typically be found
  - 10 - 400 metres radius circle

(Stallings, 2005)

To summarise, due to an increasing number of mobile phone users, more mobile masts have been put up. This has resulted in smaller cells that can cope with the higher traffic volumes. In some cases, citizens have raised protests against new masts being put up in residential areas mainly due to concerns about health risks originating from the base stations (Mast Sanity, 2006). The use of smaller cells leads to a higher accuracy for mobile phone location identification based on Cell-ID.
Even though this is not the most accurate method, the cell-id or mast in use is a prerequisite for more accurate ways of determining a handset's location (as discussed in section 2.1.4, below).

2.1.1.4 Health concerns in mobile telephony

Health concerns are based on the radiation emitted by mobile phone handsets and base stations. The latter emits radiation continuously and much more powerfully. An independent expert group on mobile phones (IEGMP), also known as the Stuart Group, was established by the Minister for Public Health Tessa Jowell in 1999 to consider concerns about possible health effects from the use of mobile phones and bases stations (Independent Expert Group on Mobile Phones, 2000). The report concluded that no clear indication of short and medium term health hazards could be found. A precautionary approach to the use of mobile phone technologies was recommended until more scientific evidence was available. A more recent report by the board of National Radiological Protection Board states that “the main conclusions reached in the Stewart report in 2000 still apply today” (National Radiological Protection Board, paragraph 19, 2004).
2.1.2 Technical overview of mobile communications support

An essential part of a mobile phone network's tasks is to monitor the location of every registered mobile phone device, as mobile phone users are free to roam throughout the coverage area of a cellular network. Therefore the network must possess some way to track mobile phones so that it can successfully route incoming calls and text messages to them. Mobile phone location data is an inherent feature of mobile communication and for this reason some technical background to mobile communication is provided in this section. At first the interplay between mobile phone handset and mobile phone base stations is explained, followed by what happens when a call takes place. It should be noted that this section only touches on the general principles of mobile communications based on cellular wireless networks. Detailed technological developments are not described here as they are not of relevance for this study.

2.1.2.1 Overview of the cellular network architecture

The GSM network architecture consists of three parts which are essential to enable mobile communication: the mobile phone handset, the base station and network subsystem. Each of these elements of a mobile phone network is described in more detail in subsequent sections to explain how the position of a mobile phone handset can be identified. See Figure 2.2 for an overview of the key elements of a cellular network. A mobile phone network continuously generates a host of information which is stored in various databases. The most relevant databases for determining a mobile phone handset’s location are the visitor location register (VLR) and home location register (HLR) database.

Figure 2.2: GSM network architecture (after Walters and Kritzinger, 2000; Walke, 1999)
Mobile phone handset and SIM card
A mobile phone handset contains a radio transceiver, digital signal processors and a removable smart card, known as subscriber identity module (SIM). The SIM card can be transferred between handsets and contains the international mobile subscriber identity (IMSI), which is used to identify the subscriber to the system. The mobile phone handset or device is uniquely identified by the IMEI number (International Mobile Equipment Identity) and depending on the mobile phone contract personal details about the user are known by the service provider. Customers of monthly paid contracts need to register their personal and bank details with their service provider, whereas for a pre-paid contract registration is not always necessary (Stallings, 2005).

Base station subsystem: Base station (BS) and base station controller (BSC)
The base station subsystem (BSS) is responsible for handling traffic and communications between a mobile phone and the network subsystem. The BSS consists of two elements, one or more base station transceivers (BST) (commonly referred to as mobile phone masts) and the base station controller (BSC). Each base station transceiver includes a radio antenna which handles the radio-link protocols with the mobile phone and a link to one of the base station controllers and is assigned to a single cell. The base station controller manages allocation of radio channels, reservations of radio frequencies and handovers of mobile phone handsets from one cell to another within the BSS. The BSC can either be located with a BST or can control multiple BST units, hence serves multiple cells. A group of BSCs is connected to a mobile switching centre (MSC) via microwave links or telephone lines (Hubaux and Znaty, 2000; Steele et al., 2001).

Network subsystem: MSC, HLR, VLR, AuC, EIR
The network subsystem (NS) provides a link between the cellular network and fixed networks such as the analogue public switched telephone networks (PSTN) or digital integrated services digital network (ISDN). The NS controls handovers between cells in different base station subsystems, authenticates users and validates their accounts. It also provides functions for worldwide roaming of mobile users. The main part of the network subsystem (NS) is the mobile services switching centre (MSC), which is supported by the following four databases that it controls.
a) Home location register database (HLR)
The home location register database (HLR) contains an entry for every SIM card issued, including details such as telephone number, mobile equipment number, equipment type and subscription type. Typically, GSM networks have only one HLR (Walke, 2003). In addition, dynamic information about the mobile subscriber is stored, for instance the current location area (LA). A location area consists of one or a number of cells. As soon as mobile phone user leaves his current LA, this temporary data held in the HLR is immediately updated (Steele et al., 2001).

b) Visitor location register database (VLR)
The visitor location register (VLR) is a temporary database that maintains information about subscribers that are currently physically in the region covered by the mobile switching centre (MSC). Entries are added when visitors enter the VLR domain and deleted when visitors leave the VLR’s domain. The VLR stores information transmitted by the HLR, such as authentication data, telephone number, agreed services, allowing the MSC to make a connection. It temporarily stores a user’s last known location area and records whether or not the subscriber is active and other parameters associated with the subscriber. The VLR also contains information that enables the network to find a particular subscriber in the event of an incoming call.

c) Authentication centre database (AuC)
The authentication centre database (AuC) handles authentication and encryption keys for all subscribers in the home and visitor location registers. It stores data needed to authenticate a call and to encrypt both voice and data traffic.

d) Equipment identity register database (EIR)
The Equipment identity register database (EIR) lists stolen phones, stores subscriber and equipment numbers (IMEI) of phones that are to be banned from the network or to be monitored. It can block calls from stolen mobile stations and prevent network use by handsets that have not been approved (Walke, 2003).
2.1.2.2 Steps in making a mobile phone call

To demonstrate how the cellular architecture works, described below is what happens when a call takes place between two mobile users within an area controlled by the same mobile switching centre (MSC) (after Stallings, 2005; Walters and Kritzinger, 2000).

Step 1: A mobile unit is turned on - initialisation

The mobile phone scans and selects the strongest setup control channel. The mobile phone selects the base station antenna of the cell within which it will operate, which must not always be the geographically closest base station due to interference patterns or other electromagnetic phenomena. A so-called handshake takes place between mobile phone handset and the mobile switching centre through the base station, which is used to identify the mobile phone user and to register its location. As long as the mobile phone is switched on, this scanning procedure is repeated frequently and if the phone enters a new cell, a new base station is selected.

Step 2: Call origination - request for connection

A mobile phone initiates a phone call by sending the number of the called mobile phone on the pre-selected setup control channel. The mobile phone's receiver first checks that the setup channel is idle by scanning the base station's forward control channel, also known as reverse control channel. When an idle is detected, the mobile is able to transmit on the corresponding reverse control channel to the base station, which then in turn sends the request to the MSC.

Step 3: Paging

The MSC completes the connection to the called mobile phone by sending a paging message to particular base stations depending on the called mobile number. Which base stations are paged depends on the mobile phone network provider but is also dependant on the physical position of the mobile phone. The mobile phone network is aware of the current position of a phone because it is stored in the VLR. Hence, only the relevant base station and those base stations in surrounding cells are addressed. Each BS transmits the paging signal on its own assigned setup channel.

Step 4: Call accepted

The called mobile phone recognises its number on the setup control channel it monitors frequently (see step 1). The mobile phone then responds to the relevant base station, which in turn sends the response to the MSC. The MSC sets up a
connection between the calling and called mobile phone. At the same time, the MSC selects an available traffic channel within each base station's cell and notifies each base station, which in turn notifies its mobile phones. The two mobile phones tune to the channels assigned by the base station.

### 2.1.3 Mobile location-based services (LBS)

“Location-based services are information services accessible with mobile devices through the mobile network and utilizing the ability to make use of the location of the mobile device” (Virrantaus et al. 2001 in Steiniger et al., 2006). Location services are available to operators of all commonly used mobile phone networks in Europe: GSM (2G), GPRS (2.5) and UMTS (3G).

The value of knowing the location of a mobile phone handset has been acknowledged by private and public sector services. The geographic location of a mobile phone can either be used to enhance existing service applications or to create new ones. The following categories of location-based services (LBS) are commonly distinguished:

**Safety and security**
- Emergency services (Europe: E112, US: E911)
- Roadside assistance
- Child finder
- Asset tracking
- Employee safety
- Retention of traffic data, including location data in the EU

**Social**
- Games
- Friend finder
- Dating

**Navigation and information services**
- Traffic and weather reports
- Navigation information (Find the quickest route / Find my nearest …)
- Tourist services

**Tracking services and logistical telematics**
- Field Staff Management (Job scheduling)
- Fleet management
- Alerts on unauthorised locations

*Sources: Elliot and Phillips, 2004; m-location.com, n.d.; Qualcomm, 2003*
These LBS typically use either active or passive location requests. The former are initiated by the mobile phone user, for example, when a mobile phone user wants to find the nearest bank. Passive location requests are not initiated by the mobile phone user and primarily encompass security and safety related applications, such as location tracking for emergency services and tracking of work force for business users.

Location-based services need to meet the expectations of subscribers and of mobile phone operators in terms of implementation and cost requirements. Performance requirements of LBS include consistency, start time and accuracy. The requirements for accuracy vary depending on the application - the majority of applications require accuracy of 10-100 metres range. Implementation requirements consider impacts on handsets, e.g. drain on batteries, roaming between networks (2G to 3G) and network expansions. Costs to take into account by operators include handset costs, infrastructure, expansion maintenance and return on investment (Qualcomm, 2003).


2.1.3.1 Location-enhanced emergency call services

In many emergencies people do not know their exact locations, which is particularly important when making the emergency call from a mobile phone. By using location-identifying technologies, emergency response times and thus consequences of injuries can be reduced. Around 180 million emergency calls are made in the European Union every year of which 60 to 70 percent originate from mobile phones. Possibilities for emergency call management are being investigated by European initiatives (European Commission, 2003). In the US, business drivers for location-based services include government mandates such as Enhanced 911 (E911) that require the incorporation of location-determination capabilities in mobile phones. The US’s Federal Communication Commission (FCC) required US mobile phone service providers to provide public safety agencies with location information in the event of an emergency. Particular performance characteristics for location determination were specified (Centre for Democracy and Technology, 2006).

In Europe, calls made by mobile phone to ‘112’, Europe's universal emergency number, will be able to inform the emergency services about the caller’s location based on Cell-ID. Mobile phone service providers will share location information with the relevant emergency service. In Europe, opposed to the USA, no specific performance standards have been set for the member states.

Article 6 of the Directive on Universal Service and Users’ Rights relating to Electronic Communications Networks and Services (2002/22/EC of 7 March 2002) states that:

“Member States shall ensure that undertakings which operate public telephone networks make caller location information available to authorities handling emergencies, to the extent technically feasible, for all calls to the single European emergency call number 112” (European Commission, 2002). This makes it a legal requirement for mobile phone service providers to deliver location enhanced 112 services across Europe.

In the EU, similar to the US and Canada, consumer consent for use of location data for emergency services is implied. The European Union has left it to voluntary efforts by industry to exploit the commercial capabilities of location-based services, however, has made it mandatory for mobile phone operators to pass on the location of callers as it is technically possible for them (Gow, 2004).
2.1.4 Different methods for locating a mobile phone

The various techniques for identifying the location of a device can be divided into network centric, handset centric methods or a combination of both. Network centric techniques locate a device based on information supplied by the network or with help of a number of mobile phone base stations, no handset enhancements are necessary. Handset centric methods require an upgrade to the mobile device as the location is calculated by the mobile phone itself from signals received from base stations. Satellite based technologies, such as GPS (see 2.1.4.2) are an example for handset-centric positioning (Steiniger, 2006).

The following location technologies are relevant to mobile phone location data and will be described together with their particular performance and implementation characteristics. Table 2.1, below, list the different methods.

Table 2.1: Mobile phone location technologies and supported mobile communication standard

<table>
<thead>
<tr>
<th>Mobile phone location technology</th>
<th>Supported in mobile communication standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell of origin (COO)</td>
<td>GSM, GPRS, and UMTS</td>
</tr>
<tr>
<td>AOA (Angle of Arrival)</td>
<td>GSM, GPRS</td>
</tr>
<tr>
<td>TDOA (Time Difference of Arrivals)</td>
<td>GSM, GPRS</td>
</tr>
<tr>
<td>E-OTD (Enhanced Observed Time Difference)</td>
<td>GSM and GPRS</td>
</tr>
<tr>
<td>OTDOA (Observed Time Difference of Arrival)</td>
<td>UMTS</td>
</tr>
<tr>
<td>A-GPS (Wireless Assisted GPS)</td>
<td>GSM, GPRS and UMTS</td>
</tr>
</tbody>
</table>
### 2.1.4.1 Network-based mobile phone positioning

#### Cell of origin (COO)

The simplest technique to identify the location of a mobile handset is based on cell-id, also known as Cell of origin (COO), see Figure 2.3. Since a mobile phone can be situated anywhere within a cell, the accuracy of these methods depends on cell size, which can vary from a few hundred metres to several kilometres (for more detail see section 2.1.2.1 above).

This method of locating a mobile phone is the least accurate but also the most inexpensive one, as it does not require individual handsets or network infrastructure to be altered. The location of the cell used by the mobile phone can be identified within about three seconds (Lyon, 2005). Accuracy of the Cell-ID method can be improved by specifying the cell sector, as each base station typically has multiple antennas, each covering a sector of the cell (see Figure 2.4). For example, a base station with three antennas will produce a cell with three 120 degree sectors. By detecting the antenna with which the handset registers, the location can be narrowed down to a sector of the cell (D'Roza and Bilchev, 2003).
Angle of Arrival (AOA)
A technique known as Angle of Arrival (AOA) determines a user's location by measuring the angles from which a mobile phone's signals are received by two or more base stations. Because the mobile device is moving, this is not a very exact method (Steiniger et al., 2006). The base stations need AOA equipment to identify the direction of the phone’s signal. The AOA equipment compares the angle between the caller and various receiving base stations and uses triangulation to determine the caller's longitude and latitude. Limitations of the AOA technique include its reduced accuracy; for example when various forms of signal interference occur. This can be due to signals bouncing caused by high buildings, which results in a weaker signal or none at all. AOA is known to work best in less populated areas, as there is a lower likelihood of interference. Many companies combine AOA with TDOA according to Deitel et al. (2002). AOA is more accurate than COO and the handsets do not have to be modified. However, it can prove costly to modify and configure the base stations (Lyon, 2005).

Time Difference of Arrival (TDOA)
The Time Difference of Arrival (TDOA) technique measures the time it takes a mobile phone signal to reach the receiving tower and two additional towers. The signal's travel time allows determining the user's distance from each tower, which in turn allows calculating the user's position. By calculating the user's distance from the receiving tower and two adjacent towers, a (virtual) set of arcs is created, which intersection indicate the handset's location. Handsets do not need to be modified to utilise this location technique, as the calculation of the position is done by the network provider (Deitel et al., 2002).
**Enhanced Observed Time Difference (E-OTD)**

Enhanced Observed Time Difference (E-OTD), uses triangulation between at least three different base stations to provide more accurate location identification than Cell-ID. The distance between handset and base station is calculated based on the different times it takes a signal to arrive at the base stations once it leaves the handset. Accuracy can theoretically reach 30 metres but can lie in reality between 50 and 125 metres (Lyon, 2005).

E-OTD only works in GSM and GPRS networks, requires an upgrade to the mobile network infrastructure, and uploading of software to base stations to ensure compatibility. The base stations need to be fitted with location measurement units (LMUs) and, by measuring the signal from the mobile phone, the LMUs can triangulate the user’s position. This technique offers greater accuracy than cell-of-origin but at a slower speed of response, typically around five seconds (Prasad, n.d.). Differing from TDOA, the calculation of the position is done by the mobile phone handset and for this reason a handset update in the form of software modification is necessary (Steiniger et al., 2006). A similar technique known as OTDOA (Observed Time Difference Of Arrival) operates only in 3G networks (Deitel et al., 2002).
2.1.4.2 Satellite-based mobile phone positioning

Global Positioning System (GPS)
The Global Positioning System (GPS) uses a set of satellites and ground receivers to determine the location of a GPS-enabled device. A GPS's receiver location is calculated by comparing time signals from several satellites, of which each has to have a direct line of sight to the receiver. At least three satellites are necessary to determine the receiver's two-dimensional location (i.e. latitude and longitude). To achieve additional and more accurate information, for example altitude, four or more satellite signals are required.
The 24 satellites orbit the earth twice a day, transmitting radio signals from approximately 12,000 miles above the Earth. The satellite system, based on spy satellites utilised during the Cold War, was originally developed by the USA Department of Defence to help troops and missiles locate themselves on maps. In the 1980s, the US government made the system available for civilian use. Once a GPS receiver has been acquired, the location identifying service is free to use and accurate to an average of 15 metres. However, the drawbacks for GPS receivers that have been integrated into mobile phone handsets are that GPS receivers consume a considerable amount of battery power, are fairly expensive, and location positioning does not tend to work from inside buildings as a direct line of sight with satellites is needed. In addition, in urban areas the GPS signal can bounce of building walls and distort the result (Monmonier, 2002; Lyon, 2005).

Assisted GPS
Assisted GPS (A-GPS) links to the terrestrial-based system of cell sites to speed up the process of calculating a handset's position. A-GPS can be combined with Cell-ID, E-OTD or OTDOA. It requires an update to the handset and network infrastructure and shares the same drawbacks as all GPS receivers (Lyon, 2005).

Galileo - European Satellite Navigation System
Galileo is a European satellite navigation system currently in development. It aims to provide positioning services from 2011 and to offer at least the same performances as GPS (European Commission, 2006b). Galileo’s navigation infrastructure, which will consist of 30 satellites and relevant ground infrastructure, is funded by the European Union and the European Space Agency (European Commission, 2006a). Galileo is under civil control and aims to ensure that European economies are independent from other states’ systems, such as GPS. GPS is a
military system by the United States and is made available to civil users without any guarantee for continuity.

Figure 2.6 below shows a comparison of the accuracy of positioning methods used by mobile phones:

![Figure 2.6: Accuracy of positioning methods (after Steiniger et al., 2006)](image)

As can be seen from the graphic above (Figure 2.6), the most accurate ways of determining a mobile phone’s geographical location involve the satellite based system GPS. Location identification by cell-ID is the least accurate method, however this technique makes use of the default properties of mobile phones. In other words, every mobile phone can be located by Cell-ID without any modifications of either software or hardware.
2.1.5 Summary Part 1

As can be seen from this short discussion of technical background, the location of every mobile phone can be identified by triangulating the mobile phone signal, irrespective of mobile phone network generation and network provider. This first part of the chapter has explained this process by providing an overview of the cellular GSM network architecture and techniques used for locating a mobile phone. Location information as part of other communications data can be stored for future data-mining and analysis, as for example for commercial location-based services or for law enforcement purposes, as discussed in Part 3 of this chapter (section 2.3). The automatic generation of digitised personal data about every mobile user and its storage on a long-term basis, may also bear negative aspects that can impact in individuals’ privacy, as discussed in the next section.
2.2 Part 2: Privacy

The second part of this chapter focuses on the representation of privacy in the literature and starts off by introducing and classifying the wide range of privacy definitions. Following this, legal and technological matters relevant to privacy are reviewed together with issues of surveillance and social justice. In addition, this section addresses the role of privacy enhancing and invading technologies, and finally explores the relationship between location-based services and privacy in the literature.

2.2.1 Setting the scene

The difficulty of defining the concept of privacy is often used as an introduction to reviews of privacy literature (see for example Introna, 1997; Solove, 2002; Bennett, 2004). Alan Westin, a renowned US privacy scholar, comments on this subject matter that "no definition of privacy is possible because privacy issues are fundamentally matters of values, interests and power" (Westin, 1995, quoted in Gellman 1998, p. 194.). He suggests understanding privacy as an interest individuals have in leading a life free from interference by others. Despite this apparent slipperiness of the notion of privacy, numerous scholars have not ceased to provide their thoughts on this subject over the last several decades: such as regarding the impact of technologies on privacy (Cavoukian and Tapscott, 1995; Agre and Rotenberg, 1997; Bennett, 2004), surveillance and privacy (Clarke, 1988; Lyon, 1994; Zureik, 2005) or regarding communications interception in the information age (Diffie and Landau, 1998).

Fried (1968 in Singh and Hill, 2003) observes that privacy is of high value for people because it allows for transactions that result in trust, which otherwise without the reassurance of privacy would not be possible. Privacy is seen as an interest of the human personality and by some believed to be the most essential human right of the modern age (Privacy International, 2003b). Despite the apparent difficulties in finding a suitable characterisation of privacy, a right to privacy is enshrined in international treaties, conventions, and agreements. In 1948, the UN declared privacy to be a fundamental human right in Article 12 of the Universal Declaration of Human Rights (UDHR). The UDHR consists of 30 Articles which outline the United Nations General Assembly’s view on human rights guaranteed to all people. Yet, privacy was declared a fundamental right in the Universal Declaration of Human
Rights, differing from an *absolute* right, which is inviolable. Fundamental rights may under certain circumstances be removed (Gauntlett, 1999).

Shortly after the UDHR the Council of Europe released the European Convention on Human Rights, which states in its Article 8.1 that "Everyone has the right to respect for his private and family life, his home and his correspondence". It has been established under the jurisprudence of the European Court of Human Rights (ECtHR) (see for example Kopp v Switzerland (1999) 27 EHRR and Halford v United Kingdom (1997) 24 EHRR 523) that interception of post and telecommunications from business and private premises falls within the scope of Article 8.1 and therefore ensures private conversations to individuals (Davis, 2003). However, Article 8.2 ECHR allows interference with this right to respect for private and family if it is "necessary in a democratic society in the interests of national security, public safety or the economic well-being of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others" (ECHR, Article 8.2) (for more detail see section 2.2.3.1).

Meanwhile, some philosophers believe that privacy rights are unequivocal and unconditional (McWhirther and Bible, 1992; Velasquez, 1992 in Singh, 2003). Goodwin (1991 in Singh, 2003) interprets privacy as a two-dimensional concept requiring space and information, whereas Gavison (1980) identifies the following three elements of privacy: secrecy, anonymity and solitude. Fried (1968) defines privacy as "control over knowledge about oneself", which corresponds to Westin's (1967, quoted in Garfinkel and Gene, 2002, p. 205) definition "Privacy is the claim of individuals, groups or institutions to determine for themselves when, how, and to what extent information about them is communicated to others". Westin supplies the following views of privacy,

a) Privacy as limited access to self, that is the extent to which we are known to others and the extent to which others have physical access to us

b) Privacy as control over information, that is not simply *limiting* what others know about you, but *controlling* it. This assumes individual autonomy in that one can control information in a meaningful way.

Privacy is again suggested to comprise the two factors control and knowledge by Foxman and Kilcoyne (1993 in Singh, 2003), in other words to have control over what information is collected and the knowledge that a data collection takes place.
and including the purpose of collection. These definitions predominately relate to the
categories of information privacy as discussed in the next section. Theoretical
approaches to privacy predominantly differentiate between the following two
aspects; protection of personal territory and protection against governmental
intrusion. This relates to the traditionally recognised desire to balance the power
between government and private individuals, as discussed by Nissenbaum (1998).

2.2.2 Privacy categories

The range of scholarly thoughts on privacy above indicates that there seems is no
universally accepted definition of privacy. For these reasons, it seems the most
useful approach to come to terms with the concept of privacy by applying different
categories or spheres of privacy. All definitions of privacy supplied by the various
authors above can be assigned to one or more of the following privacy categories.

The following four spheres of privacy are commonly identified (see for example
Nissenbaum, 1998; Clarke, 1999; Gauntlett, 1999; Privacy International, 2003b):

- a) Information privacy
- b) Privacy of communications
- c) Bodily privacy
- d) Territorial privacy

In addition to these categories, Introna (1997) suggests the following three privacy
categories, which share the element of ‘personal information’ as a common
characteristic. These particularly relate to the empirical findings of this study, which
will be discussed in detail in subsequent chapters:

- e) privacy as no access to the person
- f) privacy as control over personal information
- g) privacy as freedom from judgement by others

These seven privacy categories as listed above are commonly referred to in the
literature and are detailed in the following; the two categories relating to information
(a+f) have been combined.

In the context of mobile phone location data and its long-term storage, the most
obvious threat to privacy is towards the right to respect for private life as contained
in Article 8 ECHR (see section 2.2.3.1). Information relating to an individual’s
geographical location and gathered on a continuous basis can be classified as
privacy invasive according to six out of the following seven privacy categories:
a) Information privacy, f) Privacy as control over personal information

The categories of 'information privacy' and 'control over personal information' are the most frequently referred to privacy categories in the literature and are particularly relevant regarding modern information and communication technologies (ICTs). Mason (1986) identifies two forces related to information that threaten individual privacy. For one part the growth of information technology with its ability to generate, store and analyse great amounts of personal information, and for the other part the increased value of information in decision making to policy makers. The increased deployment of computers in the 1960s and 1970s has enabled businesses to store vast amounts of data at relative low cost for almost unlimited periods of time and has facilitated the linkage and analysis of data stored in different locations. These technical developments have resulted in the need of rules to manage the collection and handling of personal data, particularly consumer information (Privacy International, 2003b). In addition, the increased use of technology had already in the 1970s resulted in growing concerns about the potential of information technologies to accumulate detailed collections of information about individuals. Therefore the data protection legislation has been developed, in order to prevent harm to individuals and negative consequences on society. At first by some individual countries and then at international level, for example through the European Union. The UK’s domestic version of data protection law was enacted in form of the Data Protection Act 1984. Today, the Information Commissioner has responsibility for promoting and enforcing the Data Protection Act 1998 and Freedom of Information Act 2000 (Information Commissioner, 2007).

The predominant aim of most data depositories or data warehouses is to analyse the accumulated consumer data and detect patterns in order to identify consumer choices and preferences. Profiles of consumers groups are generated, supposedly facilitating the establishment of long-term relationships between business and consumer. A well-known and often cited principle in customer relationship marketing proclaims, "it costs six times more to sell to a new customer than to an existing one" (Kalakota and Robinson, 2001, p.170). In addition to these commercial uses of personal data generated by information and communications technologies, the data is used for law enforcement by the police and intelligence services. For example, the profiling of suspicious groups of citizens as part of a pro-active risk management by the government are matters that will be addressed later in this chapter (see section 2.2.5.2). Mobile phone location data is an example of personal information
specific to a mobile phone user and hence falls into this information privacy category. However, the mobile phone user does not have control over the generation of mobile phone location information or who has access to the data and for what purposes, which will be discussed in further detail in subsequent sections.

b) Privacy of communications
The dimension of communications privacy deals with the claim of not being subjected to eavesdropping when communicating with others (Clarke, 1999a). This is also known as interception privacy. The European legal framework defines the right to communication privacy as fundamental and not as an absolute right, since court orders can be used to allow wiretapping (Gauntlett, 1999). However, it is widely agreed that wiretapping and electronic surveillance are highly intrusive forms of investigation that should be limited to exceptional circumstances (Privacy International, 2003b). Even though communications data does not encompass the content of communication, it nevertheless contains very personal information such as who has contacted whom, when, for how long and how often (see section 2.3.2).

c) Bodily privacy
The area of bodily privacy, also known as privacy of the person, is assigned to the matters of shielding one's physical self against procedures such as invasive body searches, genetic and drug testing. Person-identifying techniques based on physical and other characteristics, such as biometrics also fall into this category. However, mobile phone location data cannot be seen to be related to bodily privacy, as it contains information about a person's location and movements (see next privacy category).

d) Territorial privacy
This type of privacy deals with physical space and the protection of one's home or territory from direct intrusion. Territorial privacy also extends to the use of reasonable force to defend one's home against intruders (Gauntlett, 1999). The dimension of territorial privacy can also be categorised as media privacy, as it can also relate to behaviour in private and public places and visual surveillance, such as CCTV cameras. The standard of privacy protection regarding territorial and communication privacy was approved on international level in the UN Universal Declaration of Human Rights in 1948 (for more detail see section 2.2.3). The use of CCTV was regulated for the first time by the Data Protection Act 1998 to put the collection and storage of citizens' information to closer legal control (Moran, 2005).
Mobile phone location data encompasses information about the physical space that the mobile phone user accesses, which is then retained on a continuous and long-term basis under recent legislation (section 2.3.3.2).

e) Privacy as no access to the person

The privacy literature frequently refers to an article by Warren and Brandeis published in the Harvard Law Review from December 1890, in which the authors define personal privacy as the "right to be let alone" and "being free from intrusion" (Warren and Brandeis, 1890). However, this relatively early definition of privacy had not only related to the shielding of the home from intrusion but it was also a response to technology and market developments, i.e. the development of the camera and the emerging tabloid media. Warren and Brandeis' legal opinion tends to be presented as the initial one on privacy and still greatly influences today's privacy advocates. Mobile phone location data gives access to a person's location and communications habits and hence falls into this category of privacy.

g) Privacy as freedom from judgement by others

Johnson (1989 in Introna, 1997) describes privacy as a concept that varies from context to context and points out that "it is quite possible that no single example can be found of something which is considered private in every culture". Privacy is here seen in the sense of a private space of immunity from the judgement of others, based on preconceived ideas and norms. In the case of mobile phone location data this may result in an individual being judged by others depending on where she or he has been. From historical information it may be deduced how this phone user may behave in the future.

To conclude, mobile phone location data can be seen as related to the majority of - six out of seven - commonly referred to privacy definitions. As will be discussed in further detail in subsequent sections (2.2.5), the long-term retention of mobile phone location data, together with other communications data, can be interpreted as a form of surveillance. Hence, in order for citizens to accept and to consent to this form of surveillance, it is necessary to ensure that governments are accountable for its actions (Taylor, 2002). The following sections will focus on the legislative framework for privacy and surveillance in relation to communications technologies.
2.2.3 Legal dimensions of privacy

A common misunderstanding between legal rights, on the one hand, and moral or natural rights, on the other hand, is highlighted by Clarke (1999a) and Langford (2000). Privacy is seen as equivalent to liberty and this definition confuses the legal right of privacy with the condition of privacy. The notions of legal matters and privacy are not easily separable perhaps because the concept of privacy is often merely considered in the context of protection of privacy and therefore closely related to the legal framework. It could even be understood, that the subject of privacy is only recognised when an intrusion of privacy occurs.

Four major models for privacy protection based on various international standards are identified by Privacy International (2003b) and Beresford (2005). Firstly, comprehensive laws that govern the collection, use and dissemination of personal information by both the government and private sector. For example, Europe has a comprehensive regulatory model with a public official in charge for enforcing data protection legislation. In the UK, the Office of the Information Commissioner is an independent public body established to enforce compliance with data protection legislation, such as the Data Protection Act 1998. Bennett (1995) has shown that countries that have established a data protection agency provide better privacy protection for their citizens. Secondly, sectoral laws regulating certain areas of privacy protection, for example financial or medical privacy. Thirdly, various forms of self-regulation in which industry adopts particular codes for self-control and engages in self-policing. However, Beresford (2005) remarks that industry-wide consensus does not necessarily result in satisfactory consumer privacy. The latter two privacy protection laws can be for example found the in the United States. Fourthly, technological self-help for consumers through methods such as encryption, various digital payment methods and anonymous remailers. Depending on their application, these models can be complementary or contradictory. Privacy International (2003b) believes that the joint use all of the models together provides the most effective privacy protection for citizens.

The terms 'human right' and 'civil liberty' are often used in conjunction with the term privacy. Stone (2004) explains that the phrase human right is commonly drawn on in an international context, such as by the Universal Declaration of Human Rights and the European Convention on Human Rights. The term civil liberty tends to be rather
used on the municipal level. However, Stone also acknowledges that the distinction between both terms can be blurred.

2.2.3.1 **Protecting Human Rights**

At the end of the 19th century, the American lawyers Warren and Brandeis have established the basis for privacy protection constituted in legislation (see section 2.2.2). In Europe, the right to privacy is a highly developed area of law, and after the Second World War Human Rights have been laid out in legislation, also in an attempt of the democratic states to distinguish themselves from the Eastern Bloc regimes (Bowden, 2003). All Member States of the Council of Europe have ratified the European Convention for the Protection of Human Rights and Fundamental Freedom, also known as the European Convention on Human Rights (ECHR), which was adopted in 1950. All Member States of the European Union (currently 27 members) are members of the Council of Europe (a larger body of 45+ member states), and are thus bound by the ECHR.

The Council of Europe established the European Court of Human Rights (ECtHR), based in Strasbourg, to ensure observance of the European Convention on Human Rights. An important element of this Convention system is that an individual, company or other organisation can bring a case against their own government claiming a violation against the Convention. The UK had been a signatory of the ECHR since 1950 and thus has bound itself to observe the judgement of the ECtHR. Therefore, when the European Court in Strasbourg decides that the UK has violated a right under the ECHR, the UK needs to amend its law, policy or practice so as to give effect to the judgment, while it has some discretion in doing so. Hence, the British courts are not bound to follow the Strasbourg view, which makes it possible for a “distinct British approach to human rights” to develop (Davis, 2003).

The jurisprudence of the European Court of Human Rights on privacy is more complex than that on other Convention rights, argue Whitty, Murphy and Livingstone (2001). The ECtHR looks at underlying values instead of artificially distinguishing private and public categories. The Court has adopted a definition of ‘private life’ that is more general than merely considering a person's intimate sphere of personal autonomy. For example, the ECtHR has considered worthy of respect a person's sense of self and well-being but also relationships within the wider society. Relevant cases that have been decided by the Court in Strasbourg involve personal

In the United Kingdom, the Human Rights Act made the European Convention on Human Rights part of British law and with it the right to privacy for the first time in British history. The HRA was introduced in 2000 and is only applicable to public authorities and private organisations acting as public authority. Despite these limitations, the HRA also affects the British law for all organisations and citizens as it contains a set of principles that need to be brought into court in order to be interpreted. However, even prior to the Human Rights Act 1998 coming into effect, the ECHR had already had an impact on the law of the UK when British judges sought compatibility with the Convention in their judgement (Davis, 2003).

In addition, changes to the British law became necessary due to adverse rulings from the court in Strasbourg, for example, relating to telephone tapping and surveillance in the well-documented cases of Malone v United Kingdom (1984) and Halford v United Kingdom (1997), as mentioned above. In the former case, Malone claimed that he had experienced an infringement by a public authority to his right to privacy. He alleged that he had been under police surveillance, his telephone had been tapped and phone calls metered. 'Metering' is the equivalent of retaining traffic data from landline telephones and postal communication and includes data, such as the numbers dialled, time and duration of calls, as well as the address on a postal communication (Bailey, Harris and Ormerod, 2001). However, Malone’s legal challenge was unsuccessful as there was no enforceable right to privacy in English law. The public telephone system was not considered a confidential network and hence the interception of communications did not involve the violation of any of the applicant’s rights. Finally in 1985, the European Court of Human Rights ruled that police interception of individuals' communications was a violation of Article 8 of the European Convention on Human Rights.

Subsequently, Malone’s case led to the ratification of the Interception of Communications Act 1985 to comply with the Strasbourg judgement in Malone v UK (1984) but also with the privatisation of the telecommunications service. The Court
ruled that the legal regulation of the circumstances of telephone interceptions was not expressed with sufficient clarity in the UK as required under Article 8.2 of the ECHR (Whitty, Murphy, and Livingstone, 2001; Taylor, 2002).

The case of Halford vs United Kingdom (1997) concerned a private telephone conversation using an internal office telephone system, between a police officer and her lawyer regarding a sexual discrimination complaint against her police force. The interception of this conversation by the police had been declared admissible under English law as the Interception of Communications Act 1985 (IOCA 1985) did not apply to private telephone systems, and there was no other UK law that regulated these types of systems. However, Halford took her case to the European Court of Human Rights, which ruled that an intrusion on such internal telephone systems did not comply with the requirements under Article 8.2.

The Regulation of Investigatory Powers Act 2000 was introduced partly to comply with Halford v UK (1997) concerning private telecommunications systems and other new communication technologies. Part 1 of RIPA supersedes the IOCA 1985 and extends the definition of interception to include most forms of electronic communications including email (Taylor, 2002) (see also section 2.3.3).

2.2.3.2 European data protection law

European laws aim to set a common standard across all member countries and have to take into account different cultures and laws regarding privacy protection. Nevertheless it needs to be pointed out that the impact of EU Directives is different in each country, and depends on the different mentalities of the different nations (Hosein, 2003).

In 1997, the Telecommunications Privacy Directive (97/66/EC) was introduced by the European Union to specifically regulate telecommunications systems, such as mobile networks, telephone and digital television. The Directive guaranteed 'a right to privacy' regarding telecommunications traffic of individuals. Opposing to current practices, access to billing data for marketing purposes was restricted, and data related to phone calls had to be deleted after the phone call terminated (section 2.3.4). Two years earlier, in 1995, the Data Protection Directive (95/46/EU) was introduced by the European Union to protect personal information to ensure the free flow of personal data within the EU and to harmonise privacy laws amongst its
member countries. This Directive established that personal data can only be collected with consent of the individual. The data should be processed fairly and lawfully for limited purposes and only retained for a limited period of time. Individuals should have the right to access data collected about them and to delete or change incorrect data. Some data is categorised as sensitive data, such as data relating to political opinions, religious or philosophical beliefs, trade union membership, and therefore cannot be processed without the individual's explicit consent. In spite of this, an individual's agreement is not required in exceptional cases of public interest, such as medical or scientific research, for which alternative protection mechanisms have been established (Bainbridge, 2004; Moran, 2005).

The Data Protection Directive is, amongst other Directives, implemented in all European Union member states along with local laws. To enforce compliance with the EU Directive, a Data Protection Authority has been established in each country, to ensure independence of the executive governmental branch to avoid political influence. In July 1998, the British Parliament approved the Data Protection Act (DPA) to implement the European Union Data Protection Directive in March 2000. The British handling of data collection and dissemination is regulated by the Data Protection Act (DPA, 1998) and the Freedom of Information Act (FOI Act, 2000). Both the DPA 1998 and the FOI Act 2000 reflect the development to put the collection and storage of citizens' information to closer legal control (Moran, 2005). For example, for the first time the DPA 1998 regulated the use of CCTV, a technology that was increasingly being used for surveillance. The DPA contains eight Principles that ensure adequate handling of personal data, which is data related to specific living individuals. However, information that is necessary for safeguarding national security is fully exempt from the Data Protection Principles (Davis, 2003).

The FOI Act imposes a duty on public authorities, such as government departments and councils, either to publish or disclose on demand information and documents in their possession. However, many categories of information and documents are exempted from this duty, for example if the information could be accessed by other means or other legislation. The FOI Act is evidence of a move towards greater openness and accountability, which serves the idea of citizenship, makes abuse of power clearer to citizens and aims to enable voters to make informed choices (Davis, 2003). Compliance with the Data Protection Act and Freedom of Information Act is promoted and enforced by the Office of the Information Commissioner (ICO).
Chapter 2 - Literature Review

The ICO is the UK's independent public body set up to protect personal information and promote public access to official information an independent agency (Information Commissioner's Office, n.d.).

Davies (1997) criticises that data protection acts are not concerned with the full range of privacy protection issues. He warns that the narrow scope of the acts can lead to serious limitations, as these laws are only information laws and protect data before the people. In other words, data protection acts are merely concerned with the way personal data is collected, stored and accessed but do not question the action of collecting data in the first place. British data protection laws react to complaints of non-compliance rather than to proactively examine whether organisations comply with the law. For these reasons, the awareness of individuals regarding privacy invasions is seen as crucial, as claimed by the Parliamentary Office of Science and Technology (2002), Cady and McGregor (2002) and Accenture (2003b).

Further details about the British legislation relevant to mobile phone location data are discussed later in this chapter in sections 2.3.2 and 2.3.3.2).

2.2.4 Impact of technologies on the concept of privacy

Particularly considerations about the influence of technology on privacy are of relevance to this study. When Warren and Brandeis' published their definition of privacy at the end of the 19th century, protection of privacy might have sufficiently consisted in shielding one's home from intrusion and not being watched or overheard by others. Even though their definition of privacy as related to the "the right to be left alone" is still frequently referred to today, others increasingly focus on technological influences on the concept of privacy.

Already two decades ago, Mason (1986) claimed that we live in an information age, in an information society. Continuous advancements in information and communications technologies have resulted in faster computer processors, cheaper storage and growing use of networking technologies, such as the internet. One effect of this technological 'revolution' is the increased digitisation of data. It enables routine storage and analysis of personal data on great scale, such as for example when using credit cards, the internet or mobile phones. Consequently, the increased power of computers together with the advent of the internet have influenced numerous aspects of life and hence the concept of privacy.
A number of predominantly North American publications tend to merely relate privacy to electronic communications, and seem to be primarily concerned about the 'Death of privacy in the 21st century' caused by information and communication technologies (see for example Cavoukian and Tapscott, 1996; Gauntlet, 1999; Garfinkel, 2001; Cady and McGregor, 2002). However, their treatment of privacy provides a too narrow view of the subject matter. This highlights the importance of taking into account time and context of privacy publications, as well as country and regional specific regulatory approaches, in order to provide a sufficient understanding of the subject of privacy.

Before electronic means of data storage were widely used by businesses, data could often be found scattered over different places. The cumbersome and sometimes impossible retrieval of this data could act as a kind of privacy protection. When merging data from different sources over time privacy related issues arise especially when data that has been perceived as non-sensitive at time of disclosure is combined with other data (Nissenbaum, 1998). The analysis of consumers' purchasing habits is often accompanied by the creation of consumer profiles, which is another area frequently mentioned in the literature feared to invade individual's privacy (Cady, 2002; Graeff and Harmon, 2002). Consumer profiling evokes associations with racial or criminal profiling; taking into account that not only past actions can be analysed but also future consumer behaviour, which is often tried to be predicted and possibly influenced. This entails the danger shared with all types of profiling in that assumptions are made about a person which are not necessarily accurate or complete. Bowden (2003) warns that collections of consumer data often take place without the unambiguous consent of consumers, who may give their agreement at the very beginning of a business relation but then over time forget about it. This also raises issues regarding the ownership of personal data, particularly in sensitive areas such as health.
2.2.5 Surveillance society - Threats to privacy by digital data collections

The dictionary defines surveillance as “close observation, especially of a suspected spy or criminal” (The concise Oxford English dictionary, 2001). The French word surveillance means literally to "watch over, from sur- ‘over’ + veiller ‘watch’ from Latin vigilare keep watch, Origin: early 19th century." However, ‘to watch over’ can be interpreted in two ways, which represent the two faces of surveillance: on the one hand, protection or care and on the other hand, control. Lyon (2001) argues that the worrisome and unsocial aspects of surveillance pay for mobility, convenience, speed, security and safety in modern life.

Several authors claim that today’s Western society has turned into a surveillance society (see Lyon, 2001, Marx, 2002 and Stalder, 2002). However, these authors do not perceive surveillance in the context of totalitarian regimes or regarding a centralised Orwellian ‘Big Brother’ figure but rather as relating to the continuous gathering of information facilitated networked computer systems: “Everyday life is subject to monitoring, checking and scrutinising” (Lyon, 2001, p. 1). In Lyon’s view, all societies that depend on information and communications technologies for administrative and control processes can be described as surveillance societies. Indeed, the UK is the country with the most CCTV cameras in the Western world, according to the Human Rights organisation Liberty (n.d.). There are 4.2 million cameras in the UK, which means that there is one camera for every 14 British citizens, according to McCahill and Norris (2003 in Norris et al., 2004).

In its recent report commissioned for the UK Information Commissioner, the Surveillance Studies Network’s shares this broad definition of surveillance:

“Where we find purposeful, routine, systematic and focused attention paid to personal details, for the sake of control, entitlement, management, influence or protection, we are looking at surveillance.”

(Surveillance Studies Network, 2006)

The report warns that the routine tracking and information gathering mechanisms of the surveillance society are often not obvious to citizens. This makes it particularly important to incorporate checks and safeguards when collecting data to ensure accountability.

Nevertheless, some authors such as Bennett (2005), challenge the view of defining electronic data collections as a form of surveillance. Bennett stresses that a
distinction needs to be made between capture and storage of personal information on the one side, and on the other side, data analysis and manipulation of those whose data has been collected. He bases these claims on a case study in which he examines his own data tracks generated as an airline passenger when travelling within his own country Canada and across the border to the US.

In his view, the collection and storage of his passenger data by the airlines and airports does not constitute surveillance, as there is no further analysis of that data from which decisions are made. He criticises the use of the term surveillance in this context, as he perceives surveillance as a too broad definition. Bennett suggests using another concept or term to explain the practice of humans to monitor data in order to make decisions about individuals. He believes that identifying these practices as surveillance trivialises the real surveillance, for which he gives as an example the special attention to passengers with ‘risky’ surnames. In this context, it could be seen as more suitable to use the term dataveillance, which Clarke (1988) has introduced almost two decades ago to describe the systematic monitoring of people’s actions or communications through information technology. Nonetheless, whether the term ‘surveillance’ is used or not, the extensive collection of individuals data is seen with unease by privacy scholars.

2.2.5.1 Digital technologies changing the concept of surveillance: new versus traditional surveillance

The traditional definition of surveillance as close observation of a suspect individual with the naked eye does not necessarily apply any longer, due to the capabilities of modern technologies to routinely record data. Several authors such as Lyon (1994), Marx (2002) and Clarke (2005), agree that the concept of surveillance has changed over the last decades and have declared the emergence of “new surveillance”. Marx (2002) identifies 28 dimensions in which traditional and new surveillance differ. For instance, Marx describes the differences between ‘new’ and ‘traditional’ surveillance as concerning the collection of data. The generation of predominantly digital data in private life and commercial settings leads to fewer costs of data handling. Data collections tend to occur on a continuous and routine basis, which makes the gathering less visible and enables real-time availability of information, such as for example CCTV data, use of credit cards and mobile phones. Inexpensive storage space enables the long-term retention of data. In addition, faster computer processors make it easier to organise, store, retrieve and analyse data and help to
make predictions for the future. This can be relevant for commercial companies for marketing their products, as well as for governments in order to increase accuracy of planning as suggested by Stalder (2002). However, it seems that Marx solely bases his distinction of the differences regarding the two types of surveillance on the capabilities of modern ICTs, such as assistance with the handling of great amounts of data, or in other words on the continuing development of the technological age.

Marx (2002) places the emphasis of new surveillance on mass surveillance of groups consisting of anonymous individuals, rather than on the visible and noticeable monitoring of individuals. In addition, the increased storage and analysis capabilities enable a closer examination of certain subjects of surveillance, in real-time but also in retrospective. Lyon (1994) points out that it is not necessary to identify individuals in order to conduct profiling of behaviour and habits in relation to social demographics. He fears that this can lead to social sorting and demographic discrimination of groups of individuals. Davies shares Lyon's view and offers as an example the use of statistics by the British government, instead of suspicion, to justify the use of a range of new powers, such as DNA test and finger printing (Talk Simon Davies at Leeds Metropolitan University, 28th October 2005).

Different phases in the application of surveillance technologies are distinguished by Kim (2004). Firstly, physical surveillance, as described by Bentham (1791) and Foucault (1975), as for example by their ideas about the Panopticon; secondly the concept of electronic surveillance that places the emphasis on the use of technologies to expand the observing electronic gaze from prison or factory to the wider society with help of information technologies as discussed by Marx (1985) or Lyon (1994). Thirdly, data surveillance, also known as ‘dataveillance’ which is described as the systematic use of personal data systems in monitoring of the actions or communications of one or more persons by Clarke (1994b). As last phase of the social effects of surveillance technology, Kim (2004) introduces the notion of ‘hyper-surveillance’ and refers to the authors Bogard (1996) and Haggerty and Ericson (2000). However, the author fails to explicate the difference between the phases of dataveillance and hyper-surveillance, which seems to be minuscule.
Today’s immense data collections are performed by a number of different agents. This contradicts visions of surveillance related to Big Brother, as described in Orwell's novel 'Nineteen eighty-four', or in other words one single data collection agency. Different data collection agencies can encompass commercial companies logging consumers’ shopping or financial data, governmental agencies that record images of pedestrians crossing the streets, and interceptions of communication for keywords indicating terrorist activities by listening stations such as Menwith Hill in Yorkshire.

Some perceive a change in the privacy debate in that citizens used to be worried about either commercial companies or the government having access to their data. Even though neither of these concerns have vanished, there are now relationships springing up between commercial companies and the government (13th Annual Conference on Computers, Freedom & Privacy, Plenary Session #9: Data Retention in Europe and America, 2003). This collaboration between public and private sector actors particularly applies to the discussion about data retention. The government gains access to user’s mobile phone communications data via privatised, commercial mobile phone and internet service providers.

\subsection{Data shadows of individuals and implications of ‘dataveillance’ for society}

The dangers and benefits of surveillance are identified most concisely by Clarke (1998) and Lyon (2001). Lyon believes that society profits of surveillance in terms of efficiency, mobility, security and public order. Clarke predominantly identifies the advantages of physical security and the data matching practices of organisations to combat fraud. Both authors point out as potential dangers of dataveillance the unintended consequences and negative dimensions of surveillance.

The threats of mass dataveillance can be divided into dangers to individuals and to society. Clarke (1998) identifies as dangers of dataveillance for individuals that decisions based on incorrect data may lead to unfairness. This can particularly become a problem when an individual is unaware of the data collection and hence has no means to rectify wrong data. Consequently, this can result in blacklisting a person over a matter that is still under dispute or falsely related to that person. Metaphors such as 'data shadow', 'data trail' or 'digital personae' are frequently applied in the literature to depict the effects of dataveillance on individuals (see for example Clarke, 1994b; Lyon 2001; Stalder 2002; Bennett 2005). The ‘data shadow’ is made up of personal data and follows or precedes an individual. The danger here
is the potential judging and discrimination based on data that might not be accurate. Nissenbaum (1998) points out that even though those who are aware of collections of their data may be less easily targeted or manipulated. Nevertheless, the awareness of data collection can create a climate of self-consciousness and suspicion. Clarke (2000) further perceives as a danger of surveillance that it makes it possible for authorities to focus on undesirable classes of people before they commit an offence. For example, the 1984 Police and Criminal Evidence Act gives police the right to stop and search on grounds of suspicion. This situation results in an inversion of burden of proof from the accusing organisation to the individual who now needs to provide evidence of their innocence.

Collections of personal data may not only be interpreted as a threat to personal privacy but more importantly as a danger to social justice. Clarke (1998) warns that "dataveillance is by its very nature, intrusive and threatening". Dangers of mass dataveillance to society can once again be described as to create a climate of suspicion, particularly when individuals are aware of the data collection, as well as decreased respect for the law and law enforcers, deteriorating of society’s moral fibre and solidarity. The STOA report (European Commission's Science and Technology Options Assessment office) cautions of the surveillance capability of ICTs that may be used to track the activities of journalists, campaigners and human rights activists (STOA report, 1998). These technologies can result in a chilling effect on those who participate in democratic protest warns Privacy International (1999). Along the same lines Clarke (2005) indicates that privacy is important from a political viewpoint, enabling freedom of speech to think and act. In his view, the monitoring of individuals' communications and actions threatens democracy. Lyon (2003) argues that concerns about surveillance used to be expressed in terms of privacy and freedom. He warns that the digital divide is not just a matter of access to information but that information itself can be a means of creating divisions. Hence, personal data and communications data retention cannot merely regarded as an individual privacy concern but need to be framed as bearing an effect on society as a whole.

Already in his 1988 paper, Clarke identifies costs of collecting, storing and analysing data as a 'natural control' to limit dataveillance. However today almost 20 years on, this reason does not ring true anymore. As a result of most data being generated in digital format, the costs of automatic data collections and their analysis have decreased significantly. Clarke highlights the moral obligations of IT professionals,
marketers and governments in creating and using computer systems to handle individuals' data in an ethical way.

2.2.6 Privacy enhancing and invading technologies

The question whether technology itself can be seen as 'neutral', 'good' or 'bad' is frequently addressed by scholars (Clarke, 1999b; Cady, 2002). "There is nothing inherently good or bad about the increased power of technology" maintains Marx (2006, p. 38), and suggests that new technologies with privacy invasive potential must be subjected to empirical and ethical questions and not simply accepted as good because they are novel. In the same manner, Lyon (2002) argues that new information and communications technologies do not by themselves create surveillance but rather the application of technology and the policies accompanying its usage. May (2002) asserts that ICTs do not fundamentally change society and do not constitute an information 'revolution' that requires entirely new thinking. Even though the author admits that technology has social, political and economic effects and transforms some activities within society, he also argues that this does not alter society’s substance. In addition, he challenges the widely accepted view that the arrival of the ‘information society’ has fundamentally transformed society.

Some place technologies into categories according to their ability to invade or protect individual privacy (Clarke, 1999; Bowden, 2003). Data warehousing and data mining, as introduced in previous sections, are often perceived as predominantly privacy invasive technologies (PITs). To this category are also often counted authentication technologies such as biometrics and technologies, as they do not allow the anonymisation of data. Technologies that can be used to protect the privacy of users are known as privacy enhancing technologies (PETs). They protect the privacy of individuals and provide anonymity in transactions or provide a pseudonym. Examples of PETs are anonymous web browsers, remailers and encryption (Agre, 1997; Clarke, 1999, 2001). These PETs are an important part of the privacy model of technological self-help as introduced in section 2.2.3. Kim (2004) claims that PETs may help transparency as well as reconstruct interpersonal trust in the networked environment and can help the promotion and protection of privacy rights. However, the author also admits that PETs fail to acknowledge the drive (or motivation) for exploitation or dominance.
2.2.7 Mobile phone location data and privacy

It can be argued that location technologies pose new challenges for privacy policy and law. As explained above, location data has the potential to be used to create profiles, social sorting, discrimination and differential treatment, which often remain invisible to most members of the public (see Clarke, 2000; Lyon 2005; Surveillance Studies Network, 2006). Despite these claims, it needs to be clarified that a mobile phone's location may or may not correspond to the whereabouts of an individual presumed to be the user.

Mobile phones have always been 'location-aware', because location awareness is an inherent feature of mobile telephony as discussed in the previous part of this chapter (see for example section 2.1.2). However, in recent years the accuracy with which mobile phones can be located has improved, as well as the ability to store and process large amounts of data. It is now possible to preserve and access present and past movements of all mobile phone users over prolonged periods of time. As discussed in previous sections, British citizens are ensured a right to privacy under the European Convention on Human Rights, which as been brought into the domestic context by the Human Rights Act 1998. However, this right to privacy can be interfered with under a range of exclusive purposes listed in Article 8.2 ECHR (see previous section 2.2.1) and in accordance to the ECtHR's jurisprudence. Some have argued that the blanket retention of all citizens' communications data violates those rights to privacy and details will be provided in the subsequent section 2.3.3.2. Regarding the privacy categories commonly identified in the literature, it can be argued that location data falls into six out of the seven categories (see section 2.2.2). Only the category of bodily privacy, which relates to a person’s physical self, cannot be seen as directly affected by location data. Particularly in the area of information privacy, which is concerned with individuals' personal data, a distinction needs to be made between the use of data for law enforcement and for commercial purposes. The use of communications data for the former falls under state surveillance techniques and is regulated by legislation such as the Regulation of Investigatory Powers Act 2000 (section 2.3.3.2).

The use of communications data for commercial services is for example regulated by the European Union Privacy Directive (Directive 2002/58/EC) and Vodafone's Code of Practice that regulate the access to location data (Vodafone's Code of
Practice, 2005). The Directive explicitly covers unsolicited messages for marketing purposes and stresses the need for prior consent (Article 13). Article 9 of the Directive deals with the use of location data and requires that location data may only be processed when made anonymous or with the consent of the user. Vodafone’s Code of Practice (2005) distinguishes two types of location services, active and passive (see section 2.1.3), and stresses that particularly the latter require the informed consent of the mobile phone user. Gow (2004) stresses the role of consent as establishing the legal grounds regarding location data and its importance for understanding location privacy issues in the commercial context. He distinguishes between three moments of consent: collection, use and disclosure of personal data. In the same manner, White (2003) refers to three processes to identify privacy issues particular relevant to location techniques; firstly location identification, secondly data processing and thirdly value-added use. Spiekermann (2004) predominantly perceives unsolicited messages (‘mobile spam’) as privacy intrusions and points out that the mobile phone is a more personal device than for example a desktop PC.

To summarise, the authors highlight that the informed consent of mobile phone users is required to use mobile phone location data for commercial purposes. However, for the use of location data by governmental agencies, such as for crime and terrorism, the consent of mobile phone users is not explicitly obtained but acquired by the signature in the mobile phone contract. Further discussion of this matter will be provided in the next section.
2.2.8 Summary Part 2

As can be seen from the first part of this chapter, the continuous generation and long-term storage of digitised personal data is an inherent part of today’s networked information and communication technologies. This second part has discussed the relationship and impacts of these technologies on individual’s privacy. In addition to this, a range of privacy definitions were introduced and categorised according to several criteria. Current legislation relevant to mobile phone communications data and privacy were presented, while the use of location data for law enforcement and commercial use was differentiated. It has become clear that privacy is a concept that is hard to define and this is reflected in the jurisprudence of the European Court in Strasbourg, which decides on infringements of human rights, including privacy, on a case by case basis.

The following section offers an introduction to the legislative framework regarding the retention and access of communications data by law enforcement agencies, as relevant in the European Union with focus on the UK.
2.3 Part 3: Legal framework relevant to mobile phone location data

This final section of the chapter provides an insight into the legal framework relevant to mobile phone communications data. Firstly, the section provides some definitions for mobile phone communications data, of which location data is one element. The subsequent sections review the legislation and context up to and beyond the 9/11 terrorist attacks in the US, and as relevant to the retention of mobile phone communications data in England. The main two British legal instruments that is the Regulation of Investigatory Powers Act 2000 (RIP Act) and the Anti-Terrorism Crime and Security Act 2001 (ATCS Act) are discussed, as well as the European Data Retention Directive 2006/24/EC. Finally, the predominant arguments in the disputes and debates about the compulsory storage of all British’ and also European mobile phone users’ communications data are presented.

2.3.1 Definitions of mobile phone communications data

While Part 1 of this chapter (see section 2.1) has already provided some technical background to mobile phone location data, this section focuses on the legal definitions of communications data of which mobile phone location data is one element. The ATCS Act refers to the RIP Act for definitions of communications data to be retained by mobile phone and internet service providers. Communications data is defined under the RIP Act (section 21, 4 a, b, c) as follows:

Table 2.2: Definition of communications data (RIP Act, sect. 21, 4 a b c)

<table>
<thead>
<tr>
<th>(4) &quot;Communications data&quot; means any of the following-</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) any traffic data comprised in or attached to a communication (whether by the sender or otherwise) for the purposes of any postal service or telecommunication system by means of which it is being or may be transmitted;</td>
</tr>
<tr>
<td>(b) any information which includes none of the contents of a communication (apart from any information falling within paragraph (a)) and is about the use made by any person-</td>
</tr>
<tr>
<td>(i) of any postal service or telecommunications service; or</td>
</tr>
<tr>
<td>(ii) in connection with the provision to or use by any person of any telecommunications service, of any part of a telecommunication system;</td>
</tr>
<tr>
<td>(c) any information not falling within paragraph (a) or (b) that is held or obtained, in relation to persons to whom he provides the service, by a person providing a postal service or telecommunications service.</td>
</tr>
</tbody>
</table>
Chapter 2 - Literature Review

The 'Consultation paper on a code of practice for voluntary retention of communications data' (Home Office, 2003b, section 15) explains that communications data can be divided into three broad categories, such as:

- **Traffic data** identifies who the user contacted, at what time the contact was made, the location of the person contacted and the location of the user.
- **Service use information** are for example itemised telephone call records, i.e. the numbers called from a mobile phone, including times and duration.
- **Subscriber data** identifies the user of the service, providing their name, address, telephone number.

Appendix A of the consultation paper provides further details about data types retained and their retention periods, that is for example:

**Subscriber information** relating to the person encompass for example, name, date of birth, installation and billing address and payment methods.

**Telephony data** includes amongst other data,
- all numbers (or other identifiers e.g. name@bt) associated with call (e.g. physical/presentational/network assigned CLI, IMSI, IMEI, exchange/divert numbers) (for details about these see section 2.1.2.1 and glossary)
- Date and time of start of call, duration of call/date and time of end of call
- **Location data at start and/or end of call, in form of lat/long reference.**
- Cell site data from time cell ceases to be used.

**SMS and MMS data** include for example
- Calling and called number, IMEI
- Date and time of sending
- Location data when messages sent and received, in form of lat/long reference.

Subscriber information and telephony data are stored for 12 months, whereas SMS and MMS data is only stored for 6 months.

To summarise, the British legislation defines mobile phone location data as an element of traffic data, which is one of the three categories of communications data. The European Directive on Privacy and Electronic Communications (Directive 2002/58/EC, Article 2, c) defines location data as "any data processed in an electronic communications network, indicating the geographic position of the terminal equipment of a user of a publicly available electronic communications service". This Article applies to computers, PDAs and mobile as well as landline telephones. This Directive was more comprehensive than previous rules as it used the term 'electronic communications', instead of being restrictive to certain technologies.
The Directive 2002/58/EC was amended by Directive 2006/24/EC regarding the purposes for which the data can be retained by the service providers (see section 2.3.4). The 2002 Directive (section 14) further specifies that "Location data may refer to the latitude, longitude and altitude of the user's terminal equipment, to the direction of travel, to the level of accuracy of the location information, to the identification of the network cell in which the terminal equipment is located at a certain point in time and to the time the location information was recorded".

Hosein and Escudero-Pascual (2002) argue that new legislation concerning traffic data tends to be technology-neutral which does not take into account the differences in traffic data across different communications infrastructures and protocols. The authors point towards the differing definitions of traffic data from the Council of Europe (CoE) and the Group of eight industrialised countries (G8). Both definitions explain that traffic data "does not include" (G8, see Statewatch, 2001) or "does not refer to" (Council of Europe, 2001) content, however, Hosein and Escudero-Pascual (2002) point out that there are ambiguities particularly regarding internet data. For instance, click stream information of internet access logs closely relates to the content of everything that has been accessed and downloaded (Walker and Akdeniz, 2003). For example, within a communication, data identifying www.homeoffice.gov.uk would be traffic data, whereas data identifying www.homeoffice.gov.uk/kbsearch?qt=ripa+traffic=data would be content (Home Office, 2003a, page 7; Escudero-Pascual and Hosein, 2004).

As highlighted in the previous section 2.2.7, communications data is of relevance for commercial but also for law enforcement use. In this context, Green and Smith (2004) point out that different actors use differing interpretations of mobile phone location data, particularly regarding the relationship between mobile device and the individual associated with this device. On the one hand, location data is described as being connected to particular individual consumers and hence has economic value. It is for example used for mobile location-based services (see section 2.1.3), however only with the explicit consent of the mobile phone user (see section 2.2.7). On the other hand, industry sources and regulatory bodies claim that location data, as part of traffic data is anonymous and not connected to a specific individual. Green and Smith base this claim on one of their own case studies which highlighted the difference between billing data and traffic data. The former is personal, as it includes information such as the subscriber's name and address, and therefore needs to be regulated under the Data Protection Act. The latter is anonymised at
the end of the phone call but archived by the unique phone number and therefore still of commercial value to the telecommunications companies, according to Green (2006).

Nonetheless, it can be argued that even though an individual can hardly be identified solely by the mobile phone location data of his or her phone, the combination of location data with other data may constitute an infringement on a mobile phone user’s privacy.

Mobile phone location data discloses location at the beginning and end of the call and hence records changes in location. The following shows an example of outgoing mobile phone call records:

![Figure 2.7: Example of a telecommunications call data record (UK Presidency of European Union, 2005, p.8)](image)

### 2.3.2 Background to communications data retention

Communications technologies, such as the internet and mobile phones, are an inherent part of daily lives for the majority of European citizens. Every communication facilitated by such technologies generates a host of data, such as the source, destination, mobile phone location and partial web browsing logs. The data is stored for billing and legal purposes by mobile phone and internet service providers, which are, in the following, referred to as communications service providers (CSPs). This data about communication is also known as communications or transactional data. Since in western European societies citizens but also criminals increasingly use electronic communications and rely on technological infrastructure, the focus of legislators has turned to communications data in response to the September 11 US attacks. Attacks on the communications infrastructure are also included in the definition of 'terrorism' (Terrorism Act 2000, (2)) (Walker and Akdeniz, 2003).
Particularly, the process of adopting European Directives relating to communications data retention (2002/58/EC, 2006/24/EC) highlights the political climate change following the September 11 attacks (see section 2.3.4). Communications data is seen by security, intelligence and law enforcement agencies as a useful tool to prevent and investigate crimes related to national security. It is important to recognise that communications data does not include the content but instead only refers to information about a communication. The retained communications data can be analysed and helps investigators to identify suspects, relationships between them and to establish profiles. The Home Office, the ministry responsible for justice and internal affairs in the UK, points out that evidence from use of telecommunications can be the only physical evidence available to investigators (Home Office, 2003b, section 5). The rapidly changing technological environment and the requirements of law enforcement agencies to access the data while ensuring human rights of those affected by the legislation, poses challenges in legislating the area of communications data retention.

Historically, the content of communication has been perceived as more sensitive and personal than data about communication. Therefore greater authorisation and oversight mechanisms are still needed to gain access to the contents of communication (Escudero-Pascual and Hosein, 2004). However as Privacy International (2003b) points out, the sensitive nature of transactional data of new technologies makes it similar to content of communication. Data generated by an internet user’s transactions convey a great deal about this person, because a profile of interests, associates and social context can be constructed. Similarly, location information included in mobile phone communication data is more sensitive than the location information of traditional fixed telephony communication. Hence, the long-term retention of mobile phone communications data can be seen to affect a wide variety of human rights, including the right of freedom of speech and assembly, as well as the right to privacy. Considering the jurisprudence of the European Court of Human Rights in Strasbourg, the long-term retention of mobile phone communications data may be seen as a disproportionate interference in the private lives of citizens, and in turn the interference is not necessary in a democratic society (see section 2.2.3.1).
**Digitisation and privatisation**

With the invention of the microprocessor in the 1970s and with it the widespread use of personal computers, the processing power and storage capacity of computers has dramatically increased. The change from analogous to digital communications infrastructures makes it easier to store, gain access to and search transactional data. Difficulties of retrieving analogous transactional data had acted as a natural protection against unsolicited and excessive access and misuse of data. In addition, transactional data from analogous systems does not carry as much information as digital data carries today.

In 1981, the British Telecommunications Act transferred the responsibility for telecommunications away from the Post Office, creating a separate organisation, British Telecom (BT). At this time competition was introduced into the UK telecommunications industry. In 1984, the Telecommunications Act 1984 was passed and BT was privatised and thus lost its monopoly in running telecommunications systems. This was the first of a series of privatisations of state-owned utilities throughout the 1980s and into the 1990s (Dąbier et al., 2002). This privatisation process has resulted in a situation where law enforcement agencies needed to approach businesses and not state owned companies for transactional data. In addition, changes in business models such as "always on internet access" and flat rates for internet and mobile phone use, have made itemised billing largely unnecessary. Hence communications service providers only store detailed data about mobile phone and internet usage for billing purposes for a limited time. These changes, away from state owned telecommunications business to private companies, made it more difficult for governmental agencies to access the communications data, hence, the introduction of new laws became necessary (Walker and Akdeniz, 2003).

As new telecommunications technologies emerge, many countries adapt their existing surveillance laws and for example, the case law in Strasbourg has been used to update existing British legislation to take into account new technological developments (see section 2.3.2). However, often governments apply old legislative instruments to new technologies to address the interception of networked and mobile communications without analysing how the technology has changed the nature and sensitivity of information (Privacy International, 2003b; Who Knows Where You’ve been? Privacy Concerns Regarding The Use of Cellular Phones As Personal Locators, 2004). Changes in legislation were the most obvious and
immediate responses in Europe to the terrorist threats by the 2001 attacks in the US. Security and law enforcement agencies worldwide have encouraged their governments to adopt more comprehensive approaches to the retention and access of communications data (Blakeney, 2007). This may explain why much research regarding communications data retention has focused on these legislative changes (see for example Walker and Akdeniz, 2003; Escudero-Pascual and Hosein, 2004; Whitley and Hosein, 2005).

2.3.3 Legislative developments regarding communications data in the UK

In the United Kingdom, two pieces of legislation are particularly relevant regarding communications data: the Regulation of Investigatory Powers Act 2000 (RIP Act), which regulates the access to communications data, and the Anti-Terrorism Crime and Security Act 2001 (ATCS Act), which lays down rules for the retention of communication data. The European Directive on Data Retention (2006/24/EC) also has an influence on the data retention regime in the UK, as it lays down requirements for the retention of communications data for all European member states.

The value of interception of communications for law enforcement agencies and its role to fight terrorism of all kinds has long been recognised by the British government. Already the 1999 Home Office Consultation paper ‘Interception of Communications in the United Kingdom’ proposed and justifies changes to the interception regime by law enforcement agencies at the time (Home Office, 1999). In this consultation paper the Home Office admitted that legislation had not kept up with changes in the telecommunications and postal market. For example, the number of telecommunications companies offering fixed line services had grown from 2 to around 150, and there was the beginning of mass ownership of mobile phones, as well as citizens embracing communicating via the internet. Together with this increase in electronic communication, the threat of cyber crime emerged, which is defined as criminal acts using technology or action against technology. Hence interception was thought to play an important role in law enforcement (Akdeniz, Taylor and Walker, 2001). Consequently, these technological developments made a more comprehensive legislative approach more desirable.

The Consultation paper 1999 proposed to establish a clear, statutory framework for access to communications data and this has resulted in the introduction of the RIP Act. Part 1 of RIPA addresses those proposed amendments regarding interception
of communications, also in response to rulings from ECtHR in Strasbourg where the case of Halford v UK has been particularly relevant (see section 2.2.1). In this case, the UK legislation at the time in form of the Interception of Communications Act 1985 did not extend to non-public telephone networks and hence this form of interception could not be carried out in accordance with the law.

2.3.3.1 Regulation of Investigatory Powers Act 2000 (RIP Act)

The RIP Act 2000 is the latest step towards the development of an inclusive code of policing and surveillance, including the surveillance of internet data. As described in the previous section, the increased use of mobile phones and internet by citizens, as well as rulings from the European Court of Human Rights in Strasbourg made a more comprehensive approach necessary (Akdeniz, Taylor and Walker, 2001; Bailey, Harris and Ormerod, 2001). The RIP Act also sought to regulate surveillance of the types of communication not dealt with in the Police Act 1997, as for example pagers, mobile phones or emails. The Act extends the legislative powers in relation to post and telecommunications surveillance by replacing the Interception of Communications Act 1985, which it repealed. The Interception of Communications Act 1985 permitted law enforcement agencies to have access to communications, which later had to be limited under the Human Rights Act 1998 (Hosein and Whitley, 2002). Prior to the 1985 Act, the interception of communications was not regulated by statute but by Home Office administrative practice that was given implicit recognition by the Post Office Act 1969, s. 80. (Bailey, Harris and Ormerod, 2001).

The Regulation of Investigatory Powers Bill was introduced to the House of Commons on February 9th 2000 and was given the Royal Assent, which means that it became law, on July 28th 2000. The RIP Act aimed to ensure compliance with the ECHR and the Human Rights Act 1998, as Secretary of State for the Home Department, Mr. Jack Straw, highlights in the RIP Bill’s Second Reading in the House of Commons:

“This is an important Bill, and represents a significant step forward for the protection of human rights in this country. Human rights considerations have dominated its drafting. None of the law enforcement activities specified in the Bill is new. What is new is that, for the first time, the use of these techniques will be properly regulated by law and externally supervised. That will serve to ensure that law enforcement and other operations are consistent with the duties imposed on public authorities by the European Convention on Human Rights and by the Human Rights Act 1998.”

(HC Debate, 2000, Column 767)
The RIP Act enables secret services to monitor all online activity in the fight against cyber crime. It allows senior members of the civilian and military police, customs and members of the judiciary to demand that users hand over the plaintext of encrypted material or in certain circumstances decryption keys themselves. Part I of Chapter II of the Act deals with acquisition and disclosure of communications data. The provisions of Chapter II came into force on 5th January 2004 (Home Office, 2000).

The RIP Act was highly contested when it was introduced into Parliament, as the viewpoints of government, industry and privacy advocates differed widely. Issues of discussion were the development of legislative instruments in the face of continuously changing technological developments, meeting requirements of law enforcement agencies to access the data while at the same time maintaining the human rights of those affected by the legislation (Hosein and Whitley, 2002). After a public consultation in August 2001, controversies arose regarding surveillance concerns, as the RIP Act allowed the Home Secretary to designate a large number of public authorities to access communications data without a warrant. In June 2002, David Blunkett, the home secretary at the time, suggested expanding the RIP Act greatly by extending the number of public bodies to 1039 public authorities to monitor citizens' communications data. The draft secondary legislation under section 25 RIP Act proposed allowing access to communications data without a warrant to Government departments, such as the Department for Environment, Food and Rural Affairs, and the Department of Health, as well as Local authorities and other bodies, such the Environment Agency and the Scottish Drug Enforcement Agency (The Regulation of Investigatory Powers, Communications Data: Additional Public Authorities Order, 2002).

This resulted in a controversy about the expansion of power and those plans became known as "Snoopers' charter" (BBC News, 2002). The Surveillance Commissioner admitted in his annual report that he could not ensure meaningful oversight of so many bodies without assistance. Following protests of civil liberties groups, the media and opposition in Parliament and from the Information Commissioner, the Statutory Order was withdrawn on 18 June 2002 (Statewatch News, 2003; Reporters without borders, 2004).

In 2003, the Home Office initiated a second public consultation in form of two consultation papers, one on access to communications data (Home Office, 2003a) and one on the retention of communications data (Home Office, 2003b). These
provided some background and further justifications for the need for a range of public authorities to access communications data as a necessary and proportionate requirement. For example, the consultation paper 'Access to Communications Data - Respecting Privacy and Protecting the Public from Crime' (Home Office, 2003a) seeks consultation regarding permitting access to only parts of communications data, that is service use information and subscriber data, to certain public authorities.

After this consultation period, Parliament approved the implementation of Chapter II. In summer 2006 the Home Office published a further consultation paper relating to Chapter II, Part I of the RIP Act to invite comments regarding the revised draft code of practice (Home Office, 2006). In September 2003, the Home Secretary proposed a new amendment to the RIP Act. As before, this amendment gave many bodies including local authorities, the power to access communications data of citizens. The government appointed an Interception of Communications Commissioner to protect citizens from abuses of the RIP Act and ATCS Act (Reporters without borders, 2004). Sir Swinton Thomas was appointed as the Interception of Communications Commissioner from 11 April 2003 to 10 April 2006 under Section 57(1) of the Regulation of Investigatory Powers Act 2000. His role requires him to publish annual reports which review the interception processes (see for example Swinton, 2007). Hosein argues that instead of making the retention of communications data mandatory in UK, the government took the issue to Europe (Grossman, 2006).
The following table (Table 2.2) gives details about reasons for access and organisations that are entitled to gain access to communications data under the RIP Act:

Table 2.3: Access to communications data: reasons and organisations

**Communications data can be obtained for the following reasons:**

- in the interests of national security;
- for the purpose of preventing or detecting crime or of preventing disorder;
- in the interests of the economic well-being of the United Kingdom;
- in the interests of public safety;
- for the purpose of protecting public health;
- for the purpose of assessing or collecting any tax, duty, levy or other imposition, contribution or charge payable to a government department;
- for the purpose, in an emergency, of preventing death or injury or any damage to a person's physical or mental health, or of mitigating any injury or damage to a person's physical or mental health; or
- for any purpose (not falling within paragraphs (a) to (g)) which is specified for the purposes of this subsection by an order made by the Secretary of State.

(Home Office, 2000, Chapter II, section 22)

**The following organisations hold the powers to obtain communications data:**

- Any Police Force including PSNI, any Police Force of HM Forces and British Transport Police
- National Criminal Intelligence Service (NCIS)
- National Crime Squad (NCS)
- HM Customs & Excise
- The Inland Revenue
- The Security Service (MI5)
- The Secret Intelligence Service (MI6)
- The Government Communications Headquarters (GCHQ)
- Local Authorities such as FSA, DTI, the Emergency Services; government departments such as DEFRA, Dept of Health, Home Office Immigration Service; as well as County and District Councils (see SI 2003 No 3172).
- Department of Work and Pensions

(The Investigatory Powers Tribunal, 2005)
Communications data can be obtained by the authorisation of a relevant and named senior official of an organisation listed in Table 2.2, who must consider the request for communications data to be necessary and proportionate in order to authorise it. There is a much wider range of purposes, such as public safety, preventing serious harm or collection of tax (see Table 2.2, above), than those for which a warranted interception of a communication can be made in order to obtain the content of the communication (Davis, 2003). As explained in the previous section 2.3.2, the content of communication is considered as more intrusive than information about the communication.

There is a requirement for organisations to liaise with Communications Service Providers through a Single Point of Contact (SPoC) to ensure the smooth provision of service by CSPs. A SPoC is "an individual or group of individuals within a public authority who has been trained and accredited to facilitate lawful access to communications data" (Home Office, 2003a, section 9). SPoCs assess whether access in a particular case is reasonably practical for the CSP and also consider costs and resource implications for the CSP and the public authority. The Interception of Communications Commissioner oversees authorisations of access to communications data. CSPs may be compensated for the costs involved in retaining and giving access to data, with money that may be provided by Parliament (RIP Act, section 24). The Office of Surveillance Commissioners (2003) provides an oversight of the conduct of covert surveillance by public authorities according to Parts II and III of the RIP Act.

2.3.3.2 Anti-Terrorism Crime and Security Act 2001 (ATCS Act)

The Regulation of Investigatory Powers Act 2000 did not require communications service providers to retain data, and instead only regulated access to the data that may reside in CSPs networks. However some mobile phone service providers had admitted to store communications data for months and years (Millar and Kelso, 2001), which would not have been permitted by the Data Protection Act 1998 (section 2.2.3.2). The Anti-Terrorism Crime and Security Act 2001 provided a legal basis for the retention of data and specified periods of time during which communications providers would be required to retain communications data.

In August 2000, a proposal to store communications traffic data for up to 7 years was submitted to the Home Office by the National Criminal Intelligence Service (NCIS) on behalf of a number of UK law enforcement agencies, such as MI5, MI6, the Association of Chief Police Officers, and GCHQ (Ahmed, 2000). The reasons
given for these proposed changes were the growing need to fight cyber crime, international drug trafficking and the use of computers by paedophiles (Privacy International, 2002; Whitley and Hosein, 2005). The author of this report was the Deputy Director-general of the NCIS, which oversees criminal intelligence in the UK. The document proposed giving access to the retained communications data under the provisions of the RIP Act. Communications data was suggested to be retained for real-time access by communications service providers for 12 months. Once the data was 12 months old, it should be retained for a further six year period and deleted afterwards. Mobile phone location data was described as beneficial to investigations, as it could be used to locate the proximity of a mobile phone user to a crime scene, to trace associates of suspects and to locate places of significance. Oversight was proposed to be given by the designated chief officer (Gaspar, 2000).

This proposal was never formally adopted as government policy, even though the government’s policy did change in the months and years after. In addition, capturing traffic data of all citizens was not seen, at the time, as a reasonable and proportionate response to the growing use of new technologies by criminals. Britain’s Deputy Data Protection Commissioner pointed out that the retention of traffic data would constitute a conflict with the right to privacy as laid down by Article 8 of the Convention on Human Rights (Ward, 2001).

This situation changed with the introduction of ACTSA. The Anti-Terrorism Crime and Security Bill was introduced to Parliament two months after the September 11, 2001 attacks and received the Royal Assent already on December 14, 2001. The bill passed through the House of Commons with the government allowing it only 16 hours of debate, which was seen as an inappropriate length of time, considering the complexity of the Bill and the extensive number of 129 sections. While the House of Commons did not impose a single amendment on the Government, the House of Lords voiced more opposition to the Bill. A number of Lords who are also lawyers warned that the Human Rights Act could not provide a sufficient safeguard against the illiberal measures (Fenwick, 2002). The swift schedule of introducing and debating the Act has led a Parliamentary Committee to question whether the Act had been discussed and examined in sufficient detail, as its provisions result in major implications for civil liberties (Select Committee on Home Affairs, 2001). Others consider that the September 11 attacks served as a blanket reason for establishing measures that had been previously in preparation by government departments (Whitley and Hosein, 2005).
Tomkins (2002) criticises the Act as “the surely the most draconian legislation Parliament has passed in peacetime in over a century”. Other commentators refer to the Home Secretary’s “mantra that they are merely protecting democracy” (Fenwick, 2002, p. 724), and warn that “draconian anti-terrorist laws have a far greater impact on human rights then they ever will on crime” (Wadham, 1999).

Indeed, supporting this view is the fact that a voluntary code had been agreed between government and the communications industry (section 102.4). The ATCS Act requires communications service providers to voluntarily retain communications data "to safeguard national security or to prevent, detect or prosecute crimes related to national security" (ATCS Act, 2001, explanatory notes, Section 29). These voluntary requirements can be made mandatory if deemed necessary by the Secretary of State (ATCS Act, 2001, Section 104). Walker and Akdeniz (2003, p. 14) point out that this might indicate “how hesitant Parliament felt about the grant of these powers”.

At first a draft Code of Practice was established, which was consulted upon (section 103.1.a and b) and then brought before Parliament (section 103.4). The Act allowed the Secretary of State to make the Code of Practice compulsory (section 104). Whitley and Hosein (2005) point out that the Code did not introduce the process of data retention in the UK. A sunset clause was included within ATCS Act, which meant that if the government failed to introduce a Code of Practice within two years, the provision would not be implemented (Reed and Angel, 2007). The ATCS Act builds on the anti-terrorism measures enacted in the Terrorism Act 2000, such as the extension of police powers, and the implementation of criminal co-operation measures under the Third Pillar of the EU (Shah, 2005). The September 11 attacks have let to the introduction of additional powers addressed specifically to international terrorism, however many of these powers contained in the ATCS Act were to prove even more contentious than the provisions of the Terrorism Act 2000 which they complement (Bradley and Ewing, 2007).

The ATCS Act sets the maximum retention period for data held under the provisions to 12 months, whereby longer retention periods may be justified by the business practices of the communication service provider (Home Office, 2003c, section 16). The Act contains regulations dealing with terrorism and protecting national security, such as terrorist finances (part 1-3), racial hatred (part 5), bribery and corruption (part 12). Part 4 made it possible for the Home Secretary to detain non-British terrorist suspects indefinitely, which was deemed to be incompatible with the
European Convention on Human Rights and ceased to be in force. Relevant for data retention is Part 11, as it deals with the retention of communications data for national security purposes. The retention of communications data is a form of personal data processing, and hence it is subject to the Data Protection Act 1998. Oversight of the 1998 Act is by the Information Commissioner (Home Office, 2003c, Section 29).

The provisions made in Part 11 amend the RIP Act so that communications service providers must retain communications data and disclose it to secret intelligence and law enforcement agencies far more extensively than the RIP Act had initially envisaged. The Parliamentary committee for human rights, that is the Joint Committee on Human Rights expressed concern about the rule-making authority of the Secretary of State by this part of the Act (HL 51, HC 420, 2001-02). For example, the Secretary of State may add to the organisations that can obtain communications data and increase the purposes for which the data can be obtained. This is perceived as controversial since this enables the Secretary of State to greatly increase the ability of public authorities to acquire information about citizens and organisations (Davis, 2003). Finally, the Secretary of State had been made required to publish the code of practice in draft, consult with the Information Commissioner and not to bring the code into effect until it had been approved by both Houses of Parliament. The Information Commissioner also showed concern and referred to earlier discussions about communications data retention:

“The Commissioner has been aware for some time of pressure from the law enforcement community to require communications providers to retain details of communications data. This, it is claimed, would assist the detection of particular crimes and help with criminal intelligence gathering. Although such calls have been made it has not always been clear to what extent such retention is required beyond the period for which the communications providers would retain this for their own business reasons” (HL 51 and HC 420, 2001-02)

The ATCS Act was perceived as controversial for mainly two reasons. Firstly, some of the anti-terrorist powers are seen to be disproportionate to the actual threat facing the UK. It does for example not seem adequate to include Part 11, which deals with the retention of communications data, in an emergency Bill that had to be rushed through Parliament (Shah, 2005). Secondly, it has been argued that there is a lack of connection to the September 11 attacks, as the Act contains miscellaneous provisions that relate to criminal law and criminal justice matters, as well as it
increases the general powers of the police and other governmental agencies. One example of these new powers is ACTSA Part 11, which is about the new Code of Practice on retention of communications data (Fenwick, 2002; Davis, 2003). The Act also included an extension of already controversial police powers, as it amended the Police and Criminal Evidence Act 1984 and the Criminal Justice and public order Act 1994 (Shah, 2005).

Lord Waddington, a former Home Secretary, also argued that some of the legislative measures where not related to the September 11 attacks, as for example "The proposal to make incitement to religious hatred a criminal offence has been hanging around in the Home Office for a long time, at least since 1985 when it featured in a Law Commission report. So it has precious little to do with the events of September 11th (..)“. (HL Debate, 2001)

In its jurisprudence, the European Court of Human Rights found that retention of data by governments constitutes an interference with a citizen’s right to respect for private life (Article 8.1), whether or not the governments use the data against the individual. For example in the legal case of Rotaru v Romania (2000, App. no. 28341/95), the ECtHR found a violation with Article 8 took place, when the Romanian security services had stored information on Mr Rotaru’s past activities as a university student. Privacy International (2003a) has argued that the data retention as put in place by ATCS Act is more invasive than this case, even though it is only the traffic data that is stored and not the content.

Davis (2003) claims that proportionality as a human rights concept has the most considerable impact on the law of the United Kingdom. The doctrine of proportionality provides that any interference with a Convention right must be “proportionate to the legitimate aim pursued by that interference” (Whitty, Murphy, Livingstone, 2001). Three criteria need to be met to satisfy the proportionality test; firstly an immediate and very serious threat should be evident, secondly measures that are designed to meet the legislative objective must be connected to it, and thirdly the measures should not go further than needed to combat the threat. Fenwick (2002) argues that the “ATCS Act fails to meet the first and last of these tests”, and as discussed above, some claim that some of ATCS Act’s provisions are not connected to the September 11 attacks.
2.3.4 European Data Retention Directive (2006/24/EC)

The importance of the role that communications data can play in crime and terrorism investigations had been recognised by European enforcement agencies. However, the long-term retention of every citizen's communications data had been treated with hesitation before the September 11 attacks in the US. In response to the attacks, the cooperation on criminal affairs of a wide range of European law enforcement agencies on fighting terrorism had increased while previously the jurisdiction over police matters was mainly national and so were the rules for data protection (Bignami, 2007).

The Data Retention Directive 2006/24/EC of the European Parliament and of the Council was approved on 15th March 2006 after lengthy negotiations. The Directive was seen to affect the balance of power between the privacy of individuals and the right of the state to protect national security. It gives rise to a significant change in the basic principles of personal data protection (Davies and Trigg, 2006). According to Stefano Rodota (2006), then-chairman of the EU's Article 29 Data Protection Working Party highlights, any changes affecting data protection have an effect on the degree of democracy of European citizens, since the protection of personal data is an important element of freedom in society. The dynamics in the run up the final approval of the contested Data Retention Directive were complex and numerous stakeholders (such as the European Union institutions, civil liberties groups) were involved in the negotiations.

The European Data Protection Commissioners had been aware of the law enforcement agencies’ desire to for the long-term retention of communications data (Statewatch News, 2001b). Before the 11 September attacks in the US, the tone in the European Parliament had been largely critical of data retention proposals. The long-term retention of communications data had been largely dismissed as "improper invasion of the fundamental rights guaranteed to individuals by Article 8 of the ECHR", as for example by the Data Protection Commissioners in the EU at their spring conference in April 2000 (Munir, 2005). In July 2001 the European Parliament's Civil Liberties Committee approved the draft Directive on Privacy and Electronic Communications without data retention. A report by the radical MEP Marco Cappato argued that EU should restrict the powers of law enforcement agencies regarding the retention and access of communications data, and also rejected proposals to retain traffic data for up to seven years (McAuliffe, 2001).
However, shortly after the September 11 terrorist attacks, in October 2001, external pressures from the US were applied to European Union representatives. President George W. Bush sent a letter to the European Commission President Romano Prodi, which made a suggestion to retain data that would otherwise be destroyed under the European Data Protection Directives to counter terrorism: “Revise draft privacy Directives that call for mandatory destruction to permit the retention of critical data for a reasonable period” (Statewatch News, 2001). Despite these propositions to the European Union, the US does not require communications service providers to retain communications data of all of its customers but instead uses data preservation, which means that only data about certain suspect users is held. Nevertheless, in the US communications service providers may store traffic data for marketing purposes, whereas in Europe, the CSPs used to be required to erase this type of data as soon as it is no longer needed for billing purposes (Bignami, 2007).

Until a few weeks before the final vote on 30th May 2002, the majority of the Members of Parliament opposed any form of data retention. However, after pressure by the European Council and European Union governments, parliamentarians voted in favour of the Council's position on data retention (EPIC, 2007). The EU Directive 2002/58/EC on Privacy and Electronic communications was adopted on 25 June 2002, and states that member states may now pass laws permitting the retention of internet and mobile phone communications data. It leaves each EU member State free to adopt laws authorising data retention and it could be argued that the Directive encourages the retention of data, which can be seen as the first step towards the development of a more detailed data retention Directive. The Directive 2002/58/EC reverses the position of the 1997 Telecommunications Privacy Directive by explicitly allowing EU countries to require communications service providers to store the communications data of their customers:

“Member States may adopt legislative measures [...] when such restriction constitutes a necessary, appropriate and proportionate measure within a democratic society to safeguard national security (i.e. State security), defence, public security, and the prevention, investigation, detection and prosecution of criminal offences or of unauthorised use of the electronic communication system [...] Member States may, inter alia, adopt legislative measures providing for the retention of data for a limited period justified on the grounds laid down in this paragraph.”

(Directive 2002/58/EC, Article 15)
The 2002 Directive in its Article 6 requires CSPs to erase all traffic data that is no longer needed for providing services. This however had been changed in the 2006 Data Retention Directive, which allowed exceptions in its Article 15.1.

A 'Questionnaire on traffic data retention' was distributed by the Danish presidency of the EU to all Member State's governments in August 2002. The objective was to collate comments regarding the practice and experiences of traffic data retention in European member states in order to "facilitate the development of strategies and working agendas" (Questionnaire on traffic data retention, 2002). The responses showed that 10 out of 15 member states already had a legal obligation to store traffic data or were finalising a new legislation (Responses to the data retention questionnaire, 2002). A new questionnaire was issued two years later to obtain the practices of the new European member states (Questionnaire on traffic data retention 10767/04, 2004).

After the train bombings on March 11 2004 in Madrid, in which telecommunications played an important part, a proposal for the European-wide retention of communications data was debated at the EU spring summit end of March 2004 (Blakeney, 2007). Following this, in April 2004, the UK together with France, Ireland and Sweden submitted a joint proposal to the European Union for a framework decision on the retention of communications data for between one and three years or possibly longer. Interestingly, Spain, the country where the bombings had taken place, was not involved in this proposal. This data retention proposal suggested retaining communications data to "increase prevention, investigation, detection and prosecution of criminal offences, including terrorist acts" (Alvaro, 2005; Statewatch News, 2005a). This proposal was widely criticised. The Article 29 Data Protection Working Party in its opinion from November 2004 reiterated its previous view that the mandatory retention of communications data was not compatible with human rights and privacy laws and in conflict with the current data protection legislation. In addition, they raised concerns about the cost of the European-wide implementation (Article 29 Data Protection Working Party, 2004).

Finally, in June 2005, the controversial data retention proposal was partly deemed illegal by the Legal Services of both the European Council and Commission (Statewatch News, 2005b; EDRI, 2005a). This legal opinion was based on the fact that the proposed legislation about data retention was put forward as a framework decision. A framework decision is a legal instrument under the third pillar (criminal
law, policing), where the European Parliament only has a consultation right but no real influence. However, the proposed legislation should have been presented under the first pillar (economic, social and environmental policies) as it contained obligations addressed to civil parties, namely the service providers to retain and collect data. However, placing the legislation under the first pillar would have weakened the proposal because Directives cannot create substantive and enforceable rights. A Directive has distinct objectives but leaves the national authorities the choice of form and method within a set time frame (Storey and Turner, 2005; informal talk with a desk officer for this dossier of the European Union).

The Article 29 Working Party criticised the Data Retention Directive again in their opinion issued in October 2005. They argued that “Traffic data retention interferes with the inviolable, fundamental right to confidential communications” and that this should be limited to exceptional cases (Article 29 Data Protection Working Party, 2005). As mentioned above, the Working Party had continuously questioned the appropriateness of the mandatory data retention regime and issued 20 recommendations which, however, were not fully taken into account (Davies and Trigg, 2006).

The Data Retention Directive 2006/24/EC amended the Directive 2002/58/EC on Privacy and Electronic Communications, by adding a paragraph allowing data to be retained for purposes stated in the 2006 Directive, which are “investigation, detection and prosecution of serious crime, as defined by each Member State in its national law” (Directive 2006/24/EC, Article 11). The Data Retention Directive aims to harmonise obligations on communications service providers of the EU member states but does not aim to regulate the technologies for retaining data. Communications data shall be retained for a period between six months and two years (Article 6). Exceptions are possible, so that data can be retained for longer periods (Article 12). The Directive requires member states to introduce laws complying with this Directive by 15 September 2007, however all members states asked to postpone parts of the application of the Directive until 15 March 2009 (Article 15). For example, the UK has asked to postpone the retention of internet access, internet telephony and internet email.

There is the possibility for criminal law sanctions for intentional access or transfer of data that is not permissible under national law (Article 13). Access to data shall be
provided to the competent national authorities (Article 4). However as Davies and Trigg (2006) points out, a harmonisation of approaches between member states might be difficult, considering the lack of equivalent law enforcement authorities. Each member state should appoint a public authority to be responsible for monitoring the security of the stored data. This should be the same authorities as those referred to in Directive 95/46/EC (Article 28), which is in the UK Office of the Information Commissioner, an independent government authority and reports directly to Parliament. The Directives leave the issue of who is going to pay for the retention of communications data up to the individual member states. Critics of the 2006 Data Retention Directive point out that the Directive does not provide a clear definition of offences, nor procedures for data access. It does not specify security measures required to safeguard the data, which may result in an increased risk of breach of privacy in member states with weaker data protections. Finally, the Data Retention Directive 2006/24/EC does not provide a definition of "serious crime" and this may hence lead to different definition in the member states, the same is true for penalties that may be imposed upon failure to comply with the regulations resulting from the Directive (Data Protection Working Party, 2004; Vilasau, 2007).

2.3.5 Discussion about the rightfulness of communications data retention

All groups involved in the discussion acknowledge that the retention of some communications data is a useful tool for tackling cyber crime and terrorism, as it serves as evidence of mobile phone and internet communications. However, there are various arguments against the retention of communications data of all citizens for 12 months in the UK (up to 24 months under the European legislation) and access to this data by a range of governmental agencies in the UK.

2.3.5.1 Argument 1: Objection on the ground of human rights and civil liberties

The violation of human rights is the most frequently used argument against the retention of all European citizens’ communications data for an extended period of time (see also 2.3.3.2). The retention of communications data by service providers for longer periods than for business purposes can be seen as conflicting with the European Convention on Human Rights, specifically with Article 8 ‘Right to respect for private and family life’. In 1998 the Parliament approved the Human Rights Act to incorporate the European Convention on Human Rights into UK law, which established an enforceable right of privacy in British law for the first time in history.
Article 8.1 provides that "everyone has the right to respect for his private and family life, his home and his correspondence". However, this is subject to restrictions and there can be interference by a public authority "in the interests of national security, public safety or the economic well-being of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others" (ECHR, Article 8.2). As the data retention legislation is implemented to assure national security and support the fight against terrorism, European member states can implement the Data Retention Directive in accordance with Article 8 ECHR.

Since the Human Rights Act 1998, human rights standards have been routinely incorporated in legislation. Under section 19(1)a of the HRA, Ministers responsible for introducing government Bills to either house need to provide a written statement that in her or his view, the Bill is compatible with the Convention rights, or alternatively that it is not compatible and that the government nevertheless wishes the Parliament to enact the Bill into law. The Home Secretary made this compatibility statement with regard to the ATCS Act, however, the ATCS Act contains a number of provisions that affect human rights, which make the Act is not compatible with HRA. For example, the detention without trial is incompatible with the right to liberty of the person under ECHR Article 5(1), and this required a formal derogation from Article 5(1) under Article 15 of the ECHR (Tomkins, 2002, p. 7; HL38, HC 381 (2003-04).

Privacy International (2003a) argues that “blanket data retention would subject every citizen to the certainty of ongoing and unremitting interference in his or her private life”. The richness of mobile phone communications data which conveys who has been contacted when and from where for how long, makes it possible to compile a detailed profile of a person's interests, including communications patterns, network of acquaintances and social context. Data relating to mobile communications is more sensitive than those associated with traditional fixed telephony, as it conveys the location at the beginning and end of call, as well as details about text messages exchanged with other parties. Converging platforms and infrastructures, such as internet and mobile phones increase the sensitivity of traffic data, for example when mobile phone handsets are used to access the internet (Escudero-Pascual and Hosein, 2004). Communication technologies, such as mobile phones and email, serve as everyday means of communication in private and business life in Britain and the rest of Europe. They are a vital element of
modern life and the blanket storage of all communication transactions does not give citizens a choice to circumvent this type of mass surveillance. It could be argued that individuals can be discriminated from taking part in essential interactions with friends, family, businesses and employers when trying to avoid the use of modern means of communications.

2.3.5.2 Argument 2: Data preservation instead of retention

The question whether the retention of all citizens' communications data is proportionate is also at the heart of data retention debates (see for example Whitley and Hosein, 2005). Data preservation instead of data retention is seen as a more proportionate measure in responding to terrorism than a blanket data retention regime requiring the storage of all citizens' data (European Data Protection Commissioners in Home Office, 2003b). Data preservation is more targeted than data retention, as CSPs are requested only to retain data on a case-by-case basis, which means data about a particular user for a specified period of time. For example, the Data Protection Working Party had in their Opinion 05/2004 suggested the use of less invasive mechanisms, such as the so-called "quick freeze-procedure". Instead of a general storage of all data, this system would have allowed law enforcement authorities to request the service providers to store particular data and to obtain a court order to access this data (Article 29 Data Protection Working Party, 2005).

However, this approach has been rejected by the Home Office as data preservation "will never aid in the investigation of a person who is not currently suspected of involvement with a terrorist organisation" (Home Office, 2003b, p15). It is also debated whether surveillance of communications data can help to prevent terrorist attacks at all. Tony Blair is quoted to have said that "all the surveillance in the world could not have prevented the London bombings" (Davies, 2005). A Dutch study by the Erasmus University claims that data retention is unnecessary based on a review of 65 police investigations. 'In virtually all cases' the police could get access to all traffic data required based on existing account and billing information retained on average for three months (EDRI, 2005b).
2.3.5.3 Argument 3: No definite link between communications device and user

Communications data cannot be linked without any certainty to a personal identity, thus blanket data retention is ineffective in fighting crime. Even though communications data can serve as evidence for crimes planned with help of communications technologies, an undoubted link between the device that had been used for communication and an individual cannot be established with any certainty and further investigations, according to Walker and Akdeniz (2003). This becomes particularly apparent when a web-based email system such as Hotmail is used from a PC in an internet café, as Microsoft's Hotmail does not verify details of registered users and internet cafés do not require identification. The use of prepaid phones that do not need registration poses a challenge to law enforcement agencies, as they do not require the subscriber to be identified (Gow, 2005).

2.3.5.4 Argument 4: High cost of storage and retrieval of data

The telecommunications and internet industry have highlighted the high costs for storage and searches of the retained data, as legislation requires CSPs to store communications data for longer than needed for billing purposes. The internet Service Providers' Association (Ispa) expects costs of £26 million a year to set up sufficient data retention measures and £9 million in running costs to respond to law enforcement requests based on estimates from one large UK-based internet service provider (Leyden, 2005).

The requirement of communications service providers to store the customers' data for up to two years and the costs associated with it could also affect global competitiveness of the European industry compared to other western countries. Particularly as countries such as US and Canada have rejected the blanket retention of communications data. At the same time some fear that data retention may have an effect on consumer confidence, as citizens may avoid using electronic commerce in order to not have their legal transactions logged (Open Letter from civil society groups, 2005).
2.3.5.5 Argument 5: Potential use of data for other purposes

With the introduction of the ATCS Act, communications service providers are now required to keep customer records which they previously had to erase or anonymise. Under the Data Protection Act 1998 data could not be kept for any longer than necessary (fifth principle) and the data stored had to be adequate and relevant for purposes and not excessive (third principle). The DPA 1998 implements the European Union Directive 95/46/EC and governs the collection and process of personal data of individuals by the government and private sector.

As the current data retention legislation obliges CSPs to keep various types of customer data for longer than before, there are concerns that the data may not only be stored for the access by police and other public authorities but also for commercial purposes, such for consumer profiling, marketing purposes and CRM (customer relationship management). Green (2006) points out that the requirements of the state for long-term retention of communications data coincide in an unprecedented way with the extensive data-gathering practices of the telecommunications industry.

The retained communications data is also of interest to the music and recording industry to investigate music piracy and illegal file sharing in the European Union. The Creative and Media Business Alliance (CMBA), which represents large companies in the entertainment industry such as Sony BMG, Disney and EMI, has suggested to members of the European Parliament to be included in the latest European Directive on data retention (Sherriff, 2005; Thompson, 2005).

An additional concern regarding the use of data for law enforcement is that the European Data Retention Directive (2006/24/EC) does not define the term 'serious crime' used in its Article 1. In the same manner, the RIP Act allows access to communications data under a wide range of purposes, such as public safety or protecting public health (see section 2.3.3.1). Tony Bunyan, editor for Statewatch points out that the retained communications data maybe also be of interest to future uses of governments, as the data could be used for social and political control (Statewatch News, 2005c).
2.3.6 Synopsis of data retention discussion

The retention of some communications data to be accessed by law enforcement agencies is seen as beneficial and is supported by all stakeholders involved. The central elements for debates have been whether the blanket retention of all citizens' data for an extended period of time is justified and proportionate for the purpose of fighting terrorism. In the UK, there are discussions about whether some agencies should only have access to certain types of communications data, as traffic data is seen as more sensitive than user and service data. As a result of the 2003 consultation process, access by numerous agencies to communications data has been deemed justified by the Home Secretary.

A strategy of data preservation is favoured by many parties, as it would not infringe on the privacy of innocent communications users. In addition, this would also save money to the CSPs as less data would have to be stored and accessed in response to requests by eligible organisations under the RIP Act. Nevertheless, preservation is not seen as sufficient by the Home Office, as this would only enable access to data for a limited time before it would be deleted by communications service providers. The main issue seems to be whether retention of data is seen as proportionate and the responses by stakeholders differ widely.

2.3.7 Summary Part 3

The third and last part of this chapter has introduced the legal framework as relevant to the retention of mobile phone communications data in the UK. It has become apparent that a complex interplay is taking place between the national legislation in the United Kingdom and the European legislation, such as the European Convention on Human Rights. This section has detailed the legal instruments relevant to communications data retention: the Regulation of Investigatory Powers Act 2000, the Anti-Terrorism Crime and Security Act 2001 and the European Directive 2006/EC/24. Legislation on data retention has been widely debated within the United Kingdom and the European Union, from its initial introduction to the implementations of the RIP Act and the ATCS Act in the UK and the Data Retention Directive 2006/24/EC by the European Union. The section has painted a picture of data retention in legislative terms, while carefully ensuring to consider the viewpoints of the most relevant stakeholders in the debate the retention.
2.4 Summary of chapter

The multidisciplinary nature of the subject area of mobile phone location data and the long-term retention of communications data retention in general have been addressed with a threefold approach. This chapter has firstly focused on recent technological developments regarding mobile phone location data within the UK, followed by a discussion of privacy in the literature. The last part has provided essential background to the legislative landscape of the UK and the European Union.

Mobile phone location data is an inherent part of mobile communications, and advances in computer technologies have facilitated the long-term storage and analysis of this data. Communications data may be obtained and used by law enforcement for crime and terrorism investigations but also by commercial providers to offer services based on location. The widespread use of information and communications technologies for administrative and control purposes is seen as a trait of surveillance societies. The potential of inappropriate use of the retained communications data may result in infringements of the mobile phone user’s privacy, which makes checks and safeguards to ensure accountability necessary. Before the September 11 attacks there had been widespread criticism of the long-term retention of communications data. However, the attacks have been used in the UK and EU-wide to evoke a paradigm shift in storing and accessing this type of data. The retention of data was justified under Article 8.2 of the European Convention on Human Rights, but the wide range of purposes for which the data can be accessed under the Regulation of Investigatory Powers Act 2001 and the European Data Retention Directive 2006/24/EC, have been widely criticised.

The next chapter focuses on the methodology that has been used to collect empirical data in order to portray the views of ordinary citizens in the complex debate about technological influences on individual privacy.