1

INTRODUCTION

"Technology and society shape and reshape each other ceaselessly like tireless meshing gears in a perpetual motion machine."

—Regis McKenna *Real Time: Preparing for the Age of the Never Satisfied Customer*. Harvard Business School Press, 1997

Two of the major attributes of freedom are personal communication and personal mobility. Mobile communication devices have the potential of offering both. Furthermore, mobile devices that compute and communicate have the potential to enable transformative services, for a wide range of industries, that impact personal, professional, communal, cultural, and societal life situations. In order to realize this potential in its fullest extent a large number of new technologies, infrastructure services, business models, social processes, and individual and group behaviors are emerging. Therefore, what we see happening in the mobile services space today is just the beginning. Before we look at the question of where this is heading, let us first examine the technical, business, and social dynamics of the mobile services ecosystem.

1.1 INTRODUCTION

It has long been observed that technology shapes the metaphors we use to describe ourselves, our view of life, our motivation, and our needs and wants. The question

The Fabric of Mobile Services: Software Paradigms and Business Demands, by Shoshana Loeb, Benjamin Falchuk, and Euthimios Panagos

of whether *technology shapes our lives* or whether *we use technology* to more vigorously pursue our characteristic ways of life has been of concern to scholars, spiritual leaders, business people, moralists, and many others for as long as technology in its many forms has been part of our lives. If we focus on the last three centuries and the process of societal modernization, we witness the rise of the corporation, the increase of affluence, the breakdown of the extended family, and the growth of individualism. Regardless of whether these trends shape us or we shape them, they indicate a convergence towards self-absorption and personal freedom.[1] The mobile revolution can be viewed as the next manifestation of this convergence.

The mobile revolution has been enabled by a swirl of new disruptive phenomena, including "experience economy," "social media," "network effect," "long tail," the "big switch," "mashup," "ambient intelligence," "swarming," and "always on". The day-to-day infrastructure of the old order is transforming and giving rise to a new ecosystem and value landscape of innovative businesses. Engineers, software developers, network planners, service developers, content creators, media providers, industrial designers, and a host of others provide the building blocks at the foundation. The ecosystem is dynamic – some of the new businesses succeed in it and some fail.

Where do successful disruptive businesses come from? Some, especially in the venture capital and business communities, believe that you can take any group of smart people and lock them in a room, and they will come up with interesting mobile services ideas. This is only part of what it takes. Even if we assumed for a moment that investors and large ecosystem players had found a way to streamline the creation of new service concepts, the question "What will it take to build a sustainable business around these services?" still remains unanswered. In other words, and at a very high level:

1. Are people going to use the service even if offered at no cost, given the many hundreds of other new services that are becoming available all the time?
2. How does the service fit in the technology value landscape – what pre-existing enablers can support it or add value to it, and at what cost? Such enablers might include location servers, authentication systems, and information services linked into an information "mashup."
3. How is the company offering the service going to sustain itself and, furthermore, attain an expected set of financial goals?

In today's "long tail" of the "markets of one" [AND1], how do we assess the value of existing or new technologies? In strict economic terms, the value society puts upon a piece of technology or a service is determined by its functions. As we are finding out, and will discuss in later chapters in this book, mobile services are mostly an exception to the economic value rule, as functional capabilities are not the only thing that users pay attention to (e.g., [AND], [GOL]). With the advent of the "Experience Economy," where the technology or the services are commoditized and undifferentiated, users find

[1] See [STA], [FIS], and [CAR1] for a perspective.

value in the experience of using the technology and getting the service offer. As Pine and Gilmore put in their breakthrough book *The Experience Economy: Work Is Theater and Every Business a Stage* [PIN]: "Experiences have been always around, but consumers, businesses, and economists lump them into the service sector along with such uneventful activities as dry cleaning, auto repair, wholesale distribution and telephone access. When a person buys a service, he purchases a set of intangible activities carried out on his behalf. But when he buys an experience, he pays to spend time enjoying a series of memorable events that a company stages – as in a theatrical play – to engage him in a personal way" (see also Naisbitt seminal work [NAI]). This emphasis on paying for the intangible experience rather than the tangible goods and services is partly what leads to what economists call irrational behavior [ARI]. This branch of economics is now termed "behavioral economics," and it seems to be governing much of what is happening in the mobile services space and, hence, makes it hard to predict which technologies will survive and which will die soon after being introduced. There are lots of ways to get it wrong; but certain things we do know. For example, mobile services need to be easy to use, simple to operate, and quick to develop and extend. A delicate balance amongst business models, service features and functionality, and price points is required for achieving a wide user adoption and leading to a sustainable business. But in addition to the dynamics of the individual services, we need to consider the interaction of the service with its ecosystem.

In the book *Design Rules: The Power of Modularity* [BAL], the authors discuss the concept of a "value landscape" as it was first introduced by Alchian in his seminal paper on "Uncertainty, Evolution and Economic Theory" [ALC]. The economic system is viewed as a very large value landscape that rewards the creation of new technologies, products, services, new businesses, and, ultimately, new knowledge. Within this landscape, value is not only "the height" or "latitude" of a specific technology or service but also a force that pulls both technologists and companies "upward" in the landscape. This perception of value is what compels entrepreneurs and investors to create companies that seek to both improve existing technologies and create new ones. If the success of a product or a service were predictable, the "force" of market value would pull the ecosystem players unambiguously in one direction. The process of new product or service creation would then be a straightforward deductive exercise focused on implementation. But as we have witnessed in the high-tech market in general, and during the dot-com boom in particular, location of the "high peak" in the value landscape is uncertain, and when revealed, even surprising. Therefore, the steps that need taking and the time needed to reach a given peak are unpredictable as well. In some cases, even the "direction" of the nearest peak may be unknown. The fundamental uncertainty, combined with the inherent complexity of mobile technology, products and services (the high dimensionality of the space), makes certain approaches to planning and implementing a new offering in the mobile services space much more effective than others. Specifically, in the presence of uncertainty and complexity, modular design becomes very valuable. Modularity enables design options that can radically change the market value of a given technology. This is why, as we will see later in the book, the approach taken across the board in the mobile services marketplace is to combine enablers, features, and functionality to

create new offerings that leverage modularity and re-use at all levels. This includes, for example, the "mashup" paradigm in which a new application or service is created through the "gluing together" of existing information and services running on remote servers.

The evolution of mobile services and the value landscape they create entail the co-evolution of several industries that did not have much prior opportunity to converge. Today, we are witnessing the collision of the mobile device industry, the media and content industry (e.g., Web 2.0), the computing industry, the software applications industry, and the consumer electronics industry, as well as the automotive, healthcare, and many other vertical specific industries. With a collision of this magnitude, there is no telling what the result is going to be. For instance, are mobile phones with cameras on a convergence path with Wi-Fi enabled cameras? What about telephony-enabled laptops and smart phones with sufficient computing and processing power?

The purpose of this book is to present mobile services from the perspective of the unique "fabric" of its ecosystem and its essential enablers. It is the enablers that we focus on not necessarily the services themselves, although we mention services as well as needed. We look at a few representative elements of the fabric, grouped into a few key topics:

1. Simplicity and user experience;
2. The "always on" challenge;
3. Opportunistic mobile service;
4. Design patterns.

In the rest of this chapter, we first briefly discuss what we mean by next generation mobile services, discuss three of the tangible components of the mobile services value landscape, the mobile device, the mobile Web, and the communication network, and then look at a service example. We then discuss the "social rules of mobile services" and overview the contents of this book.

1.2 UBIQUITOUS COMPUTING AND NEXT GENERATION MOBILE SERVICES

The mobile device today is an integral part of our lives. Some studies indicate that people who leave home accidently without the device usually discover this fact within a few short hours, much quicker than people realize their wallet is missing! Bloggers musing, almost nostalgically, about the days without the mobile phone (or the Web) are now commonplace. They wonder: "how did we manage to do the shopping without the option to call home to say that the store is out of bananas and what should we get instead?" or "What did teens do without being able to share every moment with someone remote and mobile?"

As is the case with many concepts that finally hit the "mass market," what we are seeing emerging today in the mobile services market was envisioned long ago, as part

of the vision of ubiquitous computing. Ubiquitous computing [SCH] (also see recent reviews in [GRE] and [MOR]) portrays a world in which computing power is so distributed throughout the environment that it becomes an integral, invisible part of life. The highway, the car, the living room, and the garment become sites of information monitoring, processing, mediation, and transmission. Simplicity dominates the user experience: interactions feel natural and spontaneous. The technology "smoothes the edges of modern life," guides us, intervenes, and gently reminds us. In this paradigm, all the information that today sits in isolation on our phones, personal computers, and Web sites becomes accessible from just about everywhere, at any time, and is exploited to deliver value in a manner appropriate to our current location and context.

That is the vision, the end game, the objective that the mobile services industry is striving for explicitly and implicitly. To make this vision a reality, we need further innovation in several areas that touch the mobile devices, the supporting networks, the application infrastructures, and service enablers. This includes, but is not limited to, the following:

- Smart mobile devices, some with nontraditional form factors (e.g., wearable and organic devices);
- Wireless sensors embedded in our environment and artifacts;
- Seamless communication and computing infrastructure supporting interoperability between wired and wireless networks;
- Service-oriented architecture for flexible and dynamic service orchestration;
- Semantic Web for machine readable Web, and application enablers such as presence and location;
- Dynamic and massively distributed programmable device networks;
- User-centric functionality including artificial intelligence and multimodal interaction (e.g., speech recognition and synthesis);
- Awareness of the user context with all of its complexities and nuances;
- Ways to ensure privacy and personal security.

More than the technical capabilities alone, ubiquitous computing in its current incarnation, albeit deficient, creates an experience users value. "It is no longer about technology," users say, "it is about the experience of being always connected." In the economic value landscape, experience is an intangible parameter and cannot be measured directly using traditional economic tools. How do you put a price tag on "smoothing the edges of life?" and, more importantly, which technologies do you use to make this happen?

1.3 THE MOBILE WEB

The promise of the mobile Web is to provide rich and personalized user experience that goes beyond what has been available so far using fixed personal computers and mobile devices. However, so far, the mobile Web has largely fallen short of expectations. It

Table 1-1: Web 1.0 ⟶ Web 2.0

Web 1.0 Entity or Trend	Web 2.0 Equivalent
DoubleClick	Google AdSense
Britanica Online	Wikipedia
Personal Websites	Blogging
Screen scraping	Web services
Publishing	Participation
Content management systems	Wikis
Directories (taxonomy)	Tagging ("folksonomy")
Stickiness	Syndication

Source: Adapted from [ORE]

has been noted that the current mobile Web is an abbreviated version of Web 1.0, and a large fraction of mobile Web users are currently denied the features of Web 2.0.

How did Web 2.0[2] come about? The concept of Web 2.0 began with a conference brainstorming session between O'Reilly and MediaLive International around 2004. It was noted that far from having "crashed," the Web was more important than ever, with exciting new applications and sites popping up with surprising regularity. Some of the companies that had survived the collapse seemed to have some things in common. Could it be that the dot-com collapse marked some kind of turning point for the Web? The consensus was that it did, and so Web 2.0 was born. Table 1-1 highlights some of the key differences between prototypical Web 1.0 and Web 2.0 sites and concepts.

The early days of the dot-com revolution produced two seemingly conflicting activities. On one hand, it was clear that "Webifying" almost any business model could lead to better value for the end user and greater efficiencies for the enterprise. At the same time, there was a surge of creativity when it came to using the Web to transform businesses and lives. Some of the service concepts were slow to take off due to technology gaps and inadequate business models, or inappropriate supply chain management and processes. For example, social networking sites – while useful – were not nearly as compelling before there was a large amount of ancillary content – such as blog entries, home videos, and original music – that could be linked to and from users. A similar mindset can be found today around "mobilizing" services in the areas of telecommunication, media and entertainment, advertising, gaming, healthcare, and automotive telematics, as well as enterprise services such as work force management, sales force management, customer relationship management, and supply chain management. Recent service offerings by wireless carriers[3] as well as new browser technologies[4] are attempts to remedy the situation. Technical and usability problems still exist but are likely to be resolved in the

[2]Source: http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-Web-20.html

[3]AT&T and Verizon Wireless provide access to multiple social networking sites using Intercasting Corp's technology.

[4]Opera Mini, Skyfire, iPhone Safari.

Table 1-2: Mobile data users in the US (2008)

Wireless Operator	Wireless Technology	Wireless Subscribers	% Data Subscribers
AT&T	GSM	71.4M	28% (20.0M)
T-Mobile	GSM	30.8M	25% (07.7M)
Verizon Wireless	CDMA	67.2M	31% (20.8M)
Sprint	CDMA	52.8M	25% (13.2M)

future. What may not be resolved that soon is the introduction of new compelling mobile Web services that will attract and sustain mass-market users beyond what is being offered today. One of the issues with mobile Web applications is achieving a critical mass to both benefit from the social networking "network effect" and provide a solid business model for the service providers. One of the gating factors is how many people today actually pay for data services on their mobile phone, as data services are needed in order to access the mobile web. Table 1-2 shows the current number of mobile services users in the United States and what percentage of them have data services that are required in order to access the Web from the mobile device as well as to receive email. The percentage of these users is only 25%, which means that there are great opportunities for growth ahead. Is this sufficient for a critical mass?

All this being said, a 2008 Nielsen Mobile Company[5] report says that mobile Web users have reached a critical mass:

> "The US is the most tech savvy nation with nearly 40 million Americans – 16% of all US mobile users – using their handset to browse on the move. In the US, the number of people using the mobile Web has increased from 22.4 million in 2006 to more than 40 million today. However, roughly 95 million Americans were paying for mobile Web access but did not necessarily use the service as mobile internet is often included as part of a larger mobile media package. Users may be either unaware or disinterested in the internet access that is provided. This is not true of owners of certain handsets, such as the first generation iPhone. 82% of iPhone owners access the mobile internet, making them five times as likely to do so as the average mobile consumer. The second-generation iPhone comes with 3G, allowing faster access to the Web . . . However, the firm found that browsing habits differed between a PC and the small screen. PC internet users visit more than 100 domains per month, on average. By contrast, the average mobile internet user in the US visited 6.4 individual Websites per month. UK use was slightly less at 5.5 per month, whilst Italian users visit 8.2 per month on average. This is due to Italians using handsets that are more sophisticated. Growth like this means the mobile Web is now a viable option for big business."

The more industry, technology, and market reports one reads the foggier the picture of what exactly is going on in the mobile market today may get, and the picture changes daily. In Chapter 2, we go into the details of why the market appears so uncertain and what the forces are (some of them opposing forces) that shape its somewhat unstable dynamics. For the purposes of our discussion here, we will take a brief high-level look at some of these dynamics.

[5]http://www.nielsenmobile.com/

1.4 EMERGING MOBILE SERVICES ENVIRONMENTS – NETWORK VIEW

To offer a mobile service today, you have to be of one of two minds. Either you think "Web and Internet," or you think "telcos and cellular." If you think Web, you tend not to worry too much about dealings with telcos, as you believe that your services will be offered to the mobile device straight over the Web, completely bypassing any telco control. In reality, things may not always be as simple, at least not for now. In the following, we will take a quick look at the two approaches to offering services.

The telco view on offering a next generation services environment is based on the IP Multimedia Sub-system (IMS) [CAM]. IMS as defined by 3GPP was originally architected as a "horizontal" (i.e., cross-cutting, or spanning) service control layer for rapid deployment of IP multimedia services. However, given market conditions, which favor a modular service enablement environment rather than an all-encompassing and, perhaps, constraining services layer, IMS is being rolled out as a middleware services infrastructure supporting several services enablers and reusable service enablement functionality, including the following:

- Presence and location for mobile devices;
- User and device identity and preferences;
- Service plan information.

In principle, these enablers allow service creators to build differentiated, feature-rich IP voice and multimedia mobile services, such as:

- Presence and location enabled voice and integrated messaging services with mobility management and social networking interaction capabilities;
- Adaptive device-specific video and multimedia services;
- Trusted third-party e-commerce payment services with micro billing and bill back to users' subscriptions plans.

Figure 1-1 illustrates a layered view of IMS, with each layer representing a separate business and technical concern. The key value proposition of IMS is the decoupling of the services from the access network layers, thereby simplifying considerably the process of creation and introduction of new services. Web Services, a "plumbing" technology defined by W3C, are a big part of another related middleware called Parlay X Web Services, which expose network capabilities to developers as Web Services, thereby making it easier for operators and developers to leverage the network.

Table 1-3 illustrates some of the differences between the Web and the telco-oriented services methodologies. As these methodologies were developed with different technical and business goals in mind, it is not surprising that they represent opposing points on many dimensions. Telco architecture depends upon "event-oriented" network signaling protocols (e.g., SS7, SIP), making it ideal for the creation,

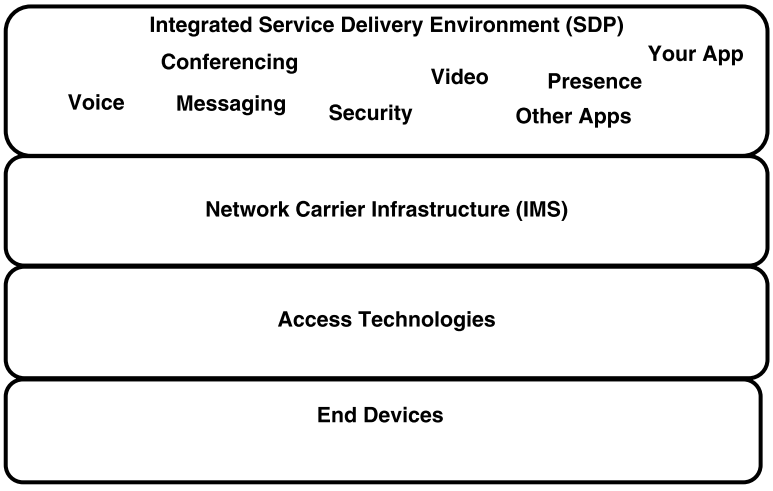

Figure 1-1: A simplified layered view of the converged mobile service ecosystem. What matters to service creators are the top (applications), which may use enablers from the second, and bottom (devices) layers.

control, and delivery of real-time network services such as voice/video-over-IP, push-to-talk, seamless voice mobility services, etc. A service-oriented architecture (SOA), on the other hand, is based upon XML and Web Services frameworks, which makes it ideally suited for longer latency and feature-rich "transaction-oriented" service creation and delivery. Examples of transaction-oriented services enabled by SOA could include: 1) ad-supported mobile multimedia streaming and download services, with revenue sharing among the content providers, operators, and handset makers, and 2) comparative on-line shopping services with micro payments and location-enabled social networking.

To get the best of both worlds, IMS operators are now utilizing SOA-based Services Delivery Platforms (SDP) implemented over the IMS layer (see Fig. 1-1)

Table 1-3: The differences between Web services and traditional telco services

	Web Services	Traditional Telco Service
Architecture	Transaction-based	Event-based
Latency	Medium to high	Low
Transaction time	Long	Short
Complexity	Low	High
Programming	Scripting	Coding
Enablers	E-commerce, user metadata, ad-insertion, gaming, media sharing, mapping, dynamic charging, LBS, personalization	Routing, screening, authentication, charging, messaging, presence, real-time communications

Source: Adapted from [MOD]

to help streamline "service creation" and, in turn, the introduction of new IP services. They are also using IMS standards for underlying service enabling and delivery aspects, such as charging, device presence, location, subscriber databases, and hierarchical levels of network and services security. The dynamic aspect of new service creation and delivery is especially important to the operator, given the highly competitive IP service provider environment, which necessitates rapid new service creation and deployment cycles. These cycles are increasingly hinged upon "access independent" approaches, short (12–18 month) payback periods, and low marginal costs of new service introduction. This may allow the network operators to leverage their SDP/Web services platforms to engage in profitable value chain partnerships with Web companies and content/e-commerce providers, involving revenue sharing and transactions fees. The partnerships may involve offering the value chain partners real-time "user information," such as user identity, security credentials, handset presence, service/feature characteristics, location, and demographics. They also may enable the introduction of targeted advertising and transaction-based business models for influencing users' demand and adoption of premium 3G services. To summarize, the IMS/SDP approach may provide wireless operators with the following business benefits:

- *Support Services Enablers* – It is reasonable to expect that there will be no single killer application but a wide variety of interesting IP-based services whose features can be aggregated to form even newer services. For example, combining a presence and location-based voice service with buddy lists and media may result in a new social networking application. This approach ensures that overall revenues are not dependent on the most popular services – which change often – but are distributed among a wide range of services.
- *Experiment with New Business Models* – The SDP/IMS platform is designed to enable operators to implement new services delivery business models, especially ad-supported and transactions-based models. The new business models hold promise for operators looking to 1) introduce feature-rich wireless services, e.g., presence and location-enabled voice and multimedia services, supported by targeted ad insertion based on the users' identities and demographics; and 2) generate transaction fees from ecommerce activities that are enabled via their services. Using an experimental approach towards new services definition and trialing is clearly a trend that will lead to rapidly discover the next generation of sticky and profitable services.
- *Create an "Open" Environment for Ecosystem Providers* – Most network operators have recognized that they alone cannot create and deliver *every* valuable service. An SDP/IMS infrastructure enables an "open" applications services creation and deployment environment that is conducive to incubating an open-ended ecosystem (value chain) comprising third-party content providers, service providers, applications developers, and device and systems vendors. Note that, in this case, the network operator is still in the picture, but as we will discuss in Chapter 2, they control less of the ecosystem.

1.5 THE MOBILE DEVICE ENVIRONMENT

Mobile devices today have progressed remarkably from mere voice communication machines (i.e., telephones) into state-of-the-art platforms that can send and receive e-mails, take mega-pixel digital pictures, access Websites, monitor biometric information, and so on. Furthermore, modern mobile devices allow service creation through such operating systems and platforms as Symbian, Linux, Java ME, Android, Widgets, and iPhone SDKs. It is expected that these platforms and their future versions will steadily gain momentum and support from a worldwide ecosystem of application developers, device vendors, and wireless network operators.

Steps toward weakening the hold that the phone companies have over the mobile device applications continue to be taken. In the year 2000, the Java profiles (e.g., MIDP) implemented on Connected Limited Device Configuration (CLDC) opened up the applications playing field to Java developers. More recently, the formation of the Open Handset Alliance Project (which includes mobile operators, handset manufacturers, software companies, and semiconductor companies, and as members)[6] is doing a similar thing. All the tools necessary for the development of applications for the platform are free and open, and the devices will treat (almost) all applications equally. This will create a disruption in the value landscape and Google, which is the driving force behind the effort, will benefit from the wider availability of advanced applications on mobile devices.

Due to the richness of technology, business context, and user context, mobile devices are viewed from a different perspective by different industry plays. Below are some examples of these perspectives:

1. The telecommunications industry traditionally views the mobile device as a small phone you can take with you.
2. The media industry views the mobile device as the *fourth screen* on which media content can be viewed. It comes after the movie theatre, the television, and the personal computer.
3. The computer industry views the mobile device as an extension to the personal computer.
4. The gaming industry views the mobile device as an extension to the gaming console.
5. The healthcare and well being industries view the mobile device as an extension to the doctor's office, the hospital, the nurse, or the counselor. It can support reminding services for medical compliance to a glucose monitor to a personal well-being coach.
6. The automotive telematics industry views the mobile device as an extension of the car and use it as a gateway between the car and a wide range of connected services.
7. Many other verticals, such as real estate, are coming up with unique applications that add new dimensions to their business.

[6]See: www.mobilehandsetalliance.com

Looking ahead to the next one or two years one can expect device features such as large high-resolution screens, powerful processors, larger storage space, considerable battery life, intuitive input stream, video phone, compass, USB port, accelerometers, and software-defined radio. These features will enable gestural input and innovative applications and many other features.

1.6 THE PERSONAL ASSISTANT EXAMPLE

One of the hardest applications to put together is one of the first ones that comes to mind when one thinks of an advanced futuristic technology – the *personal assistant*, the all-intelligent system that can know what you want even before you do. In this section, we devise a thought experiment in which we construct such an application, and then add more complex functionality to the basic service it provides. This thought-experiment is aimed at highlighting enablers and architectural options.

The personal assistant application is intended to find information for the user when it is most needed (e.g., a plumber to come for some fix-up, a restaurant on our way). In the simplest case, the system that implements this application is nothing but a straightforward directory lookup – looking up the service provider we need in a directory. The directory may be populated manually through an administrative interface and may be accessed through a lookup function. A user interface is required to input requests and return answers. The interface is then enhanced with some scheduling and calendaring functions, allowing the user to specify the time he needs the service and to schedule recurring services. It may also allow the user to specify the location of the service and other attributes such as cost.

As the next step up in sophistication, the system manages the user's profile, as well as richer descriptions of service provider capabilities. At this stage, the system is still doing simple matching between user profiles and the service providers' directory; but will need an additional interface to input and manage the profile. Next, the system will incorporate user feedback on past selections. The feedback requires an additional user interface to give the feedback and another to consult it. Next, the system hooks up to other information systems, allowing it to understand provider rankings.

Next, the system can apply more sophisticated selection criteria and use collaborative filtering techniques in order to guess, or predict, which services the user might like based on what other people with similar preferences like. This feature may also help increase the application's up-take in a community.

After introducing the notion of a community of end-users, the system now can incorporate the notion of a community of service providers and allow them to introduce their own profiles and preferences for matching in a more dynamic way than a static "yellow page-like" style. Sophisticated recommendation and notification algorithms will be needed to meet the requirements described above. Operationally, once a service request has been initiated, the system parses it and matches it with information from its database to identify qualified service provider candidates.

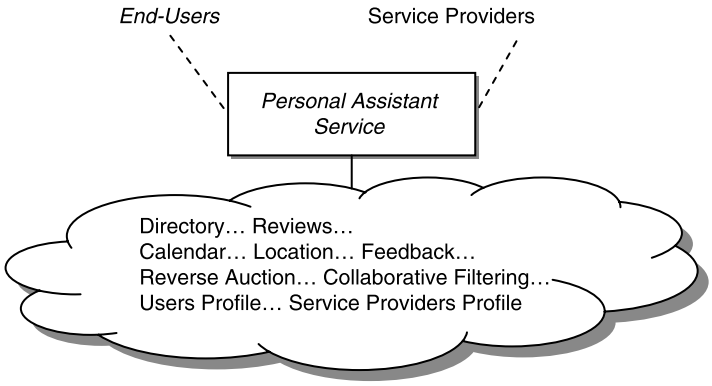

Figure 1-2: Personalized assistant example.

Recommendation algorithms will further filter and sort the results of the service request so that the list of results can be presented in a manner that is meaningful and attractive to use. After a selection is made, the system notifies the chosen service provider using one or more methods of communications (e.g., e-mail, Web-based, SMS, traditional telephony).

Figure 1-2 illustrates the system at a high level. Now things get a little more interesting as context comes into play. Imagine that the user carries a mobile device and allows the application to know his whereabouts and imagine further that the application knows personal information such as his calendar and to-do lists and that he is now "driving from Point A to Point B". This version of the application becomes complex and "context aware" [SCH].

An interesting next step would be to evolve the system from a Personal Assistant into a Market Maker (e.g., [NIS]). One way this can happen is to add an auctioning function to the infrastructure in which the service providers might bid on the right to satisfy a user's service request, perhaps through a reverse auction paradigm (as implemented by a "market maker"). An important thing to notice here is that advertisers can play the role of service providers. Here is how the different parties benefit from the service:

- Market makers may benefit by charging a fee to service providers to participate in the auction and, in some cases, increase their revenue due to larger profit margins;
- Local service providers benefit from participating in the program by gaining access to revenue generating opportunities that they may not be exposed to otherwise. Furthermore, service providers who win the auction process may increase revenues through follow-on services.

In the rest of the book, we will be revisiting in much more detail aspects of this example.

1.7 THE SOCIAL RULES OF MOBILE SERVICES

So far in this chapter, we have seen selective glimpses of the dynamic technology and business value landscape. The pace of change in the market will soon make obsolete much of what is being written today. Therefore, in this book we attempt to find some insights that we believe are universal about the technology, the business, and our lives; insights that may guide our understanding of service evolution. Our cognitive abilities are no different from what they were over thousands of years ago (with some believing that they have deteriorated due to technology [CAR]). Furthermore, if we check the recorded history we will find out that our complaints about life and the difficulties of life have not changed much. Therefore, as we indicated previously, the human part is *the* gating factor in our acceptance of anything new, including new mobile services. Consequently, perhaps there are some insights hidden in the swirl of "the mobile services market." Let's examine some of these insights. A good starting point is to look at Rachel Hinman's Adaptive Path Blog, [HIN], and examine some examples of how mobile devices are changing the social rules:

1. It is acceptable to talk to yourself in public places when using a Bluetooth headset or speaker feature on the mobile phone. This, of course, sometimes creates a backlash.[7]
2. Individuals show addictive "Crackberry behavior" that may be alarming to friends and family in spite of the enormous increase in personal and business productivity associated with the "always on" device.
3. One is always perceived to be available now that one has a mobile phone, and not responding sometimes creates negative connotations.

This last point is an example of technology creating a situation (e.g., being always connected) that then leads to a problem when one is not immediately responsive. There might be many technical solutions to solve this, but each may in turn create a fresh set of new problems (privacy-related and so on).

According to Hinman, social rules are based on our understanding of expectations – expectations about ourselves and about how we want people to engage with us. In the early days of any major disruptive technology, social misalignments and issues arise (see also [FIS], [STA]). It is clear that we associate different expectations to the various communication channels. When we call someone, we expect them to answer and if they do not answer, we leave a voicemail and expect them to get back to us. The rules for e-mail and text messaging are slightly different – often the social rules are personal and reflective of the relationship the sender and receiver have – or the social contract they share. When someone does not answer the phone, there is a set of likely explanations: the receiver is busy, the user is not within reach of their phone. A conflict seems to occur when expectations are not in alignment. For example, technology such as GPS positioning – which we will

[7]See Fred Wilson's "the no voice internet" http://www.avc.com/a_vc/2008/09/the-no-voice-in.html

cover later on in the book – or an Instant Messaging status message of "busy" adds a layer of complexity. Using these tools, a sender can know more about the receiver's state and adjust their expectations accordingly perhaps at the expense of the receiver's privacy. For a service to be widely adopted, the service designer and operator must be sensitive to all these social rules. For example, as we see later, even if the service operator may know the full context of the user, it may not want to reveal it all. The service must:

- Allow people to do what they need to do to retain social contracts and privacy;
- Support different levels of profiles and different levels of engagement to reflect the variety of expectations people have for various relationships.

Hinman and her workshop participants state that mobility implies changing contexts and changing interruptability. When designing a mobile service, assessing user interruptability may be key to user acceptance and it all comes down to the question: *Who is responsible for managing the user attention and intention?* Is it the user through explicit protective measures expressed via preference lists, calendar entries, to-do lists, profiles, or application configuration files, or is it the responsibility of the service designers or the service operators? As an example, when people first signed up to Twitter,[8] which allows them to report to their friends what they are doing at any moment, they chose to follow lots of friends, but they eventually turned off the service because they were inundated with too many "tweets." As Hinman points out, was the problem a flaw in the design of Twitter or was the problem in that people did not fundamentally understand the "rules" of social networks on mobile? This is one of the hardest questions facing the market today. Fundamentally, Hinman argues users should have no barrier to what they can do with a service, and that ultimately the burden for providing controls should be embedded in the system. However, as Hinman indicates it is often difficult to predict what will happen and design appropriately. Therefore, in the end one is not really offering services to users, but offering services to users who are embedded in a context. The issue is recognizing the different roles that people have and the communication channel they use – SMS, mobile Web, voice, applications. Depending on the context of your application, expectations are different. Measuring intangible user experience, like measuring any intangible (see for example methodologies discussed in [HUB]), is still a challenge.

The problem of capturing user context from various available information sources such as calendar entries, to-do lists, e-mail interactions, and phone calls is hard (e.g., [WAT]). Progress has been made on this problem, usually under the umbrella of intelligent digital personal assistance as we discussed in the previous section. Software agents, when implemented effectively, help users with everyday duties like managing timetables, organizing travel, and so on. Mobile intelligent agents may soon enable applications likely to create a lot of interest, both among operators and users, in fields as diverse as interactive personalized content, e-commerce, or gaming.

[8] www.twitter.com

1.8 THE FOCUS OF THIS BOOK

In this book, we attempt to blend several perspectives on the evolution of the mobile application technology, services, and, ultimately, markets. This is not an easy task as we are bound to occasionally tilt too much towards one perspective. It is our hope, however, that the temporary unevenness will smooth out through the flow of the book as a whole.

In Chapter 2, we first paint with large brush strokes the picture of the mobile services market as a highly complex system that is in a state of transition. This reflects at the macro level with several large industries such as telecommunication, computing, and entertainment on a collision/convergence path, each with its own momentum. At the same time, at the micro level, the level of the individual user, the system exhibits what economists call irrational behavior, where "user experience," which is an intangible, counts more than tangibles. We then proceed to discuss the notion of a value chain as an indication of a business model and show how the value chain warps as the market evolves, when players either expand their offerings into neighboring parts of the value chain or buy companies that dominate a market segment. This is followed by a discussion of the two worlds that the mobile services emerged out of from the network viewpoint – telecom and Web – and proceed to look at current innovative business models. We conclude with a specific example of personalized mobile service and its various business dilemmas.

Location-based services have several key technical components. In Chapter 3, we devote a significant discussion to the technical enablers that make such services possible. This includes an overview of the cellular, radio, and satellite communications systems with a particular focus on how their respective positioning technologies and accuracy plays a role in mobile services. We then examine the functional components that comprise today's location-based services (LBS) deployments. For example, mapping, geo-coding, content filtering, and service logic are critical pieces that operators need to deploy in order to run LBS successfully.

In Chapter 4, we shift perspective and turn to a discussion of the merits and principles of simplicity and user experience in mobile services. This includes a look at what the new expectations are for mobile user experience, how needs of mobile users differ across segments, and how mobile hardware and services are being designed with simplicity in mind (and how they may *not* be). We complete this discussion with real-world examples of mobile simplicity and user experience, focusing on some real products and services.

To meet the needs and expectations of mobile users, mobile services have "always-on" requirements. In Chapter 5, we give this challenge an in-depth look by first examining why it is so important for mobile services to be highly available and perceived as always on. We then look at the infrastructure and IT components and concerns that provide always-on capability and the important distinctions between scaling "up" versus "out."

In Chapter 6, we look at an emerging paradigm for mobile services that many refer to as "opportunism." Although this term often refers to link layer aspects of communications, there is a larger sense for the term in the mobile services realm. We will

explain the term, why it is an important new paradigm, and what it takes to support it successfully. Such support technologies include Semantic Web concepts [BER], artificial intelligence, agent-based systems, and others. We will provide detailed examples of opportunistic services that are currently in research or in real deployments; this will give the reader a clear picture of the scope and function of such services. For the benefit of planners, we outline what is still missing (e.g., in standards, in markets, and in technology) and make predictions on where opportunism is heading.

Designing and implementing mobile services takes special insight and skills. Today's current art is a mix of older techniques and software patterns with emerging paradigms and best practices. As mobile services continue to diverge from what services used to look like, newer design patterns will dominate. In Chapter 7, we will take an in-depth look at the old and new patterns that form today's mobile services view. These will include architectural patterns such as Model-View-Controller and client-server, as well as issues relating to User Interface structure and design (e.g., wizard, slideshow, and pagination techniques). Data distribution and resource management will be presented as key issues. Widely seen as a fundamental enabler of future mobile services, event-driven architectures will be also examined.

To wind up the book, Chapter 8 examines future trends, advanced research, and cool services. The players, the challenges, and the current state-of-the-art will be studied. In that vein, we survey some selected advanced new research in mobile services that will give readers a good idea of the sort of innovation that is around the corner.