Personal Health Records Success; Why Google Health failed and what does that mean for Microsoft HealthVault?

Ton Spil
Industrial Engineering & Business Information Systems
University of Twente
a.a.m.spil@utwente.nl

Rich Klein
College of Business
Florida International University
kleinrich@yahoo.com

Abstract
Five years of experimenting with Personal Health Records has not yielded the results that big companies like Google and Microsoft expected. Whereas Google pulled the plug on its product offering, Microsoft struggles to reach sufficient critical mass.

This study adopts a user perspective (51 interviews) in conjunction with grounded theory, to offer explanations why Google Health failed and predictions relative to Microsoft’s ability to reach a tipping point with respect to product/service viability.

Noteworthy, vendors ignore relevance, or perceived usefulness when designing PHRs. Moreover, low trust and high risks do not bode well for long-term success, with the widely used information systems success models often neglecting the latter two critical dimensions.

1. Introduction
Electronic personal health record (PHR). Personal health record, or abbreviated to PHRs, offer users a variety of advantages aimed at patient empowerment. Employing these systems, users control their own information, creating a more balanced and complete view, compared to existing provider maintained health records [1]. Further, PHRs offer extra features and functionality such as making online appointments, supplemental information specific to illnesses, information about health care providers, and options for self-care possibilities, among others [2].

Sunyaev [3] presents a framework for evaluating PHRs based on functionality, subsequently adopting the model to evaluate both Google Health as Microsoft HealthVault. The posited framework, however, do not lend itself to evaluating a service based purely on end-user functionality. Accordingly, the current study takes an end-user perspective employing a validated qualitative approach to match diverse quantitative measures.

The first product studied is Microsoft HealthVault, which started in October of 2007 as a platform to store and maintain healthcare and fitness information. With its launch Craig Mundie of Microsoft notes, “we wanted to see what Microsoft could do to anticipate the changes in healthcare and introduce technology that makes it more graceful to move from the old model to a data driven model”. The second product studied is Google Health, which offered users the opportunity to manage their own health information. Introduced in 2008, and retired with at the start of 2012, Google Health failed to capture widespread adoption and achieved only limited use [4].

Derived from behavioral models, the widely accepted Technology Acceptance Model (TAM) [5] evaluates the success of information system adoption through the use of usefulness, ease of use, and attitude. The subsequent UTAUT model [6] was introduced to integrate efforts aimed at expanding the initial TAM model, although recent studies critique this model as well [7][8]. Both user adoption models do not fully encompass all factors associated with user adoption, particularly in the context of Internet-based systems, where important domain specific factors; including trust [9][10], service quality [11], and risk [12]; remain unaddressed. Here, many initiatives have tried to extend the TAM model [13] [14][15] to cover e-
commerce specific success factors. For instance, an expansion of the seminal Delone and McLean Model of IS success [16] includes e-commerce specific measures. Despite sharing constructs and like propositions, however, no single model fully addresses all success factors of user adoption of e-commerce.

Employing a grounded literature search approach, we explore factors associated with user adoption of standalone PHR applications explaining these in greater detail through interviews with former Google Health and potential Microsoft HealthVault users. The next section provides background on personal health records. We then provide an overview of our research design and methodology, explaining success factors. We subsequently review of our results. Last, we discuss the study findings and use our results to evaluate Google Health’s failure and Microsoft HealthVault’s potential for success.

2. Personal Health Records

A personal health record is a health record maintained and controlled by the individual patient [17][18]. Information recorded in a PHR often includes medications, allergies, medical history, and so on. Noteworthy, differences exist between a PHR and an electronic health record (EHR) or electronic medical record (EMR). While one or more healthcare providers hold the latter two, an often cited definition for a personal health record, provided by the Markle Foundation¹ notes that a “personal health record (PHR) is an Internet-based set of tools that allows people to access and coordinate their lifelong health information and make appropriate parts of it available to those who need it” [19]. Expanding upon this definition, Tang et al. [18] contend that the form of a PHR can range from a standalone application (with no connection to other systems) to a “tethered” PHR, which offers patients a connection to their healthcare provider’s EHR. In the middle of the spectrum one can find hybrid PHR systems that can connect to other systems, acquiring and transmitting data. Figure 4 depicts the range of complexity in PHR systems [18].

¹ The Markle Foundation is a charitable organization situated in the United States that works to “realize the full potential of information and information technology to address critical public needs, particularly in the areas of health and national security” (Markle, n.d.).

Unfortunately, within the existing literature, few clinical trials and systematic reviews concerning the effects of providing patients with the option to use and maintain personal health records, particularly standalone applications, exist (a query of Pubmed resulted in only a few relevant publications). Tenforde, Jain, and Hickner [20] noted limited evidence of the value of PHRs, identifying only three randomized trials in their review. Furthermore, these trials plagued by “study limitations that obscure a clear interpretation of their results”, reported inconsistent results [20]. Another recent publication reports on a pilot study comparing Google Health and Microsoft HealthVault, when tethered to an advanced PHR maintained by the United States military [21]. The study context, namely, a highly tethered system within the U.S. military, undermines our ability to gain a greater understanding of the PHR adoption challenges in other contexts or as applicable to the general population. Unfortunately, as noted by Collins et al. [22] highly tethered PHR applications are limited to healthcare institutions with “sufficient financial, intellectual, and human capital resources” to support large scale initiatives fostering adoption within such environments. Accordingly, given that such systems are not limited to a specific user population, exploring the adoption challenges to standalone PHR adoption constitutes important pursuit.

Researchers citing the benefits of PHRs, often return to an early work by Tang et al. [18], who only name a number of hypothetical, or proposed, benefits. These include greater patient access to health information and data, with patients subsequently using information to “improve their health and manage their diseases” [18]. Moreover, given patients’ ability to collaboratively track individual health issues with their providers, lower communication barriers between patients and providers emerge. Another noted benefit lies in PHRs’ facilitation of “ongoing connection between patient and physician”. Providers note that with continuous and ongoing, as opposed to episodic, interactions, the time to address emerging health issues shortens. Finally, consistent with patient-accessible
EMRs/EHRs, Tang et al. contend that greater patient engagement in their individual health situation arises in the presence of increased medication adherence regimens. Further evidence of this phenomenon is seen in a recent randomized trial of 267 patients from 11 practices. This effort reports greater concordance with medication regimens and a reduction in potential discrepancies for PHR users [23].

Another potential benefit not named by Tang et al. [18], but by others (e.g. [1]), focuses on patient empowerment. Closely aligned with the notion of empowerment, Collins et al. [22] note increased patient engagement based upon their telephone survey of 17 organizations with tethered PHR applications. Consistent with the engagement theme, Wagner et al. [24] conducted a randomized trial of 453 hypertension patients in PHR and no PHR groups finding no significant impact between groups; however, self-identified active and frequent PHR users saw a significant reduction in blood pressure. Wagner et al.‘s finding highlights the need to conduct work outside the tethered system context used for their study [24]. Finding specific to active system users underscores our need to understand individual patient adoption issues, beyond the tethered PHR context.

In addition to potential benefits espoused by authors, a number of concerns emerge within the existing literature. One concern commonly voiced with respect to EMR/EHR and PHR systems focuses security and/or privacy [25][19]. Another important issue, relevant to standalone and tethered contexts, involves the accuracy of PHR data. Here, Witry et al. [19] conducted a qualitative study using focus groups consisting of physicians, with the aim of identifying participant perceptions about PHRs. Physicians voiced a number of concerns with regard to the accuracy of the records. Specifically, patients might input information that had not been verified by a professional and therefore might be incorrect. Patients might also record inappropriate information absent the ability to know what is significant [19]. Subsequently, providers often find themselves burdened with large quantities of unprocessed data [25]. Physicians further noted concerns that patients might even omit or alter their information to avoid possible consequences from their insurance carrier, as one example [19]. The accuracy concern contributed to physicians expressing a desire to control data input by requiring providers validate information prior to updating records [19]. Alternatively, in Liu et al.’s [25] study, physicians indicated limited concerns with respect to patients, family members, or caregivers deliberately altering records, as providers believe that “most patients just want to be cured”.

3. Method

Research method introduction

We employ an interview model-based research method called PRIMA [26] (previously termed USE IT [27]). The qualitative research method is chosen to afford a more detailed understanding of success measures, by complementing a literature study with an interview protocol aimed at unraveling potential underlying end-user motivations. Noteworthy, few qualitative research initiatives in the area of e-health appear within the existing literature focusing on end user patient adoptors.

Selection of interviewees and analysis considerations

Drawing on the UTAUT [6] model we include gender, age, and experience as moderators influencing the determinants of behavioral intention and actual use behavior. Our research involves consumer services making use voluntary. As prescribed by the research method, our interviews should represent homogenous groups [47], which we classify using the gender, age, and prior experience of interviewees. Previous research shows experience positively influences adoption, with users adopting one service expressing a greater likelihood of adopting another [28][29] with perceptions evolving over time [48].

Several studies have tried to extend existing models like TAM [13][14][15] as well as the Delone and McLean model [37], while other work has integrated different models [12]. Despite sharing constructs and like propositions, no single model fully addresses all success factors of user adoption within the e-health arena. Therefore, rather than draw upon a single model, we extract success factors identified across the literature specific to different models and theoretical perspectives, independently evaluating these factors using interview data. Success factors found by the extensive literature study serve as input for our research.

We completed a total of 51 interviews among potential users of Google Health (27) and Microsoft HealthVault (24). These interviews represented different homogenous groups [47]. First individual outcomes were extracted by scanning interview transcripts for success factors identified in the existing literature. For each success factor we formulated hypotheses...
checking each against all interviews conducted (scored as negative, not mentioned, positive)[51][52]. This process permitted us to triangulate data using different interviewers and different interviewees across a variety of socio-demographic criteria thus improving validity [51]. All interviewers employed the same instructions, 1.5 hours training and standard interview framework.

Interview model and analysis constructs

The PRIMA model [27] consists of five areas of analysis, namely, (1) Process, (2) Relevance, (3) Information needs, (4) Means and people, in addition to (5) Attitude.

Process consists of a description of all activities users perform completing certain tasks. Rogers [28] states that an innovation has to be compatible with existing values, experiences, and practices. Therefore, unraveling current processes emerges as a good indicator of compatibility. Rogers [28] defines compatibility as “the degree to which an innovation is perceived as consistent with the existing values, past experiences, and needs of potential adopters”. Subsequently, the degree to which an innovation is compatible can “either speed up or retard its rate of adoption” [28][29]. Karahanna et al. [30] validate three distinct aspects of compatibility, namely, compatibility with prior experience, compatibility with existing work practices, and compatibility with values.

Relevance seeks to answer the question “what is the value for the user of the e-health service”. While a subjective measure, users understand both what value means to them and when value exists. Accordingly, relevance primarily covers our definition of perceived usefulness. Consistent with our literature study, the usefulness of the e-health service closely aligns with the usability of the service [31], resulting in information relevant to both the usefulness and usability success factors. Research further posits that higher usability increases both perceived usefulness [32] and intention to use [33]. That said studies show weak, or no support, for a direct effect on intention to use [34][32][35]. Venkatesh [36] defines perceived usefulness as “the extent to which a person believes that using the system will enhance his or her job performance”. In other words, the system must deliver some appreciable value. Distinct from perceived usefulness [37], usefulness is often not objectively measurable, but rather a subjective perception of an individual user. Noteworthy, perceived usefulness consistently predicts purchase intentions across a large variety of research contexts [33][36][38].

Information needs describes which information users like to receive from the system and should align with the information the service delivers and captures. Information quality influences both perceived usefulness [34][39] and perceived usability, both while mediated by trust [40]. Researchers often measure information quality in terms of accuracy, timeliness, completeness, relevance, and consistency [16]. Egger [41] gives some guidelines for informational content. These guidelines encompass product and service information, information about the company, and information limiting user risks.

Means and people aspect examines the resources available to users given the assumption that hardware and support enable effective use of e-health services. In the cases of both Google and Microsoft, customer support constitutes the only direct contact with the end user, alongside other indirect contacts through user access of Website resources. System quality measures system design and construction aspects; however, measures generally include usability, availability, reliability, adaptability, and response time [16]. Service quality constitutes a key concern for service providers as reducing customer defections by only 5% bares the potential to boost profits by as much as 85% to 100% [42]. Further, superior service quality increases positive behavioral intentions while decreasing undesirable intensions [43], as evident in stimulating customer retention and improved loyalty versus preventing bad word-of-mouth communications. Given the impersonal nature of e-commerce, service quality is especially important to such transactions [43][44]. By using an Internet-based service, users incur different risks. Lee [12] identifies different perceived risks from the user perspective, specifically, performance, social, time, financial, and security risks as distinct facets of perceived risks [12]. Perceived risks also have a negative influence on perceived usefulness, user attitude, and intention to use [12]. In situations of higher risk, higher trust can reduce perceived risk [10].

Finally, attitude explores user resistance to an innovation. Resistance is not by definition positive or negative, but can serve as a useful input in exposing system flaws [45]. While resistance itself does not constitute the problem, resistance is often caused by underlying problems and tends to disappear in the presence of satisfactorily conditions. Accordingly, questions asked as part of the interview focus on social pressures to use a given service and levels of user trust.
While not an explicit part of our interview questions, we implicitly cover trust based upon individual’s attitude towards the e-commerce service and the Internet in general. Trust in e-commerce positively influences willingness to buy, further trust mitigates risk [10].

Social pressure, a subjective norm [32][36], influences one’s attitudes specific to intention to use [32][36]. In an online context, social pressure can result from interactions with friends and acquaintances, but also from broader informational social influences [46] like online reviews.

Looking at the PRIMA method [26], all success factors from our literature study appear either directly or indirectly and are shown in Table 1. We include the interview framework in Appendix 1.

### 4. Results

#### 4.1. Microsoft HealthVault

Between 2012 and 2013, we conducted a total of 24 interviews specific to Microsoft HealthVault. Table 2 reports the demographic data for this group of interviewees.

<table>
<thead>
<tr>
<th>Age</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>15-25</td>
<td>14</td>
</tr>
<tr>
<td>25-45</td>
<td>4</td>
</tr>
<tr>
<td>45+</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 2: Demographic data Microsoft HealthVault data.

Relevance focuses on “what is the value for the user of the e-health service”. While a subjective measure, the user both understands what value means to them and when value exists. Accordingly, relevance primarily covers our definition of perceived usefulness. Consistent with our literature study, the usefulness of the e-commerce service closely aligns with the usability of the service [31], resulting in information relevant to both success factors. Hence, posit the following:

**Perceived usefulness will positively influence user adoption of the e-health service.**

**Perceived usability will positively influence perceived usefulness.**

More than half of the interviewees offered negative responses specific to both hypotheses, seeing the primary of use of the system as time consuming, especially inputting of information. The user friendliness, however, was viewed positively. That notwithstanding, more than half of the participants failed to see any added value. Healthy interviewees in particular did not see any underlying need for such a system. For each of the success factors in Table 1 a similar analysis is performed and summarized in Table 3 with a description provided in this section.

The quality of information using Microsoft HealthVault was perceived positive. Most participants found the program useful and possessed a clear understanding of what information was readily accessible, the quality of the stored data, as well as the timeliness of system responses. Some users found the program useful, even participants in good health. Primarily due to the amount of time required to create the record, respondents reported a negative impression of perceived usability.

Table 3 provides highlights and an overview of our results. Noteworthy, despite positive perceptions, almost all interviews expressed a ‘HIGH’ risk. This is potentially directly related low trust in combination with the safety of the system and the certainty of the vendor protecting privacy. Virtually none of the interviewees fully trusted the program with respect to the security of their medical data. As one interviewee...
questioned, “who can see all of this information, who has the right in certain circumstances to be allowed access, who does not?”

<table>
<thead>
<tr>
<th>Success factor</th>
<th>Microsoft Health Vault</th>
</tr>
</thead>
<tbody>
<tr>
<td>Service quality</td>
<td>+</td>
</tr>
<tr>
<td>Information quality</td>
<td>+</td>
</tr>
<tr>
<td>System quality</td>
<td>+</td>
</tr>
<tr>
<td>Trust</td>
<td>-</td>
</tr>
<tr>
<td>Perceived usability</td>
<td>+</td>
</tr>
<tr>
<td>Perceived risks</td>
<td>HIGH</td>
</tr>
<tr>
<td>Perceived usefulness</td>
<td>-</td>
</tr>
<tr>
<td>Social and personal influence</td>
<td>-</td>
</tr>
<tr>
<td>Perceived compatibility</td>
<td>+</td>
</tr>
</tbody>
</table>

Table 3: Overview results Microsoft HealthVault data.

One an essential aspect, namely social and personal influence, merged as important to the future success of the PHR. Our assessment of all of the interviews concluded that a bad result can also be seen as positive. We posit a relationship between the implementation of the application and outside influences. A positive external signal may see more people develop a more favorable impression of the system and thus potentially develop a more positive attitude about the product. In addition, negative signals to the greater marketplace often yield negative impacts.

4.2. Google Health

In 2012, some 27 interviews were conducted with potential users of Google Health. Table 4 reports demographics for this sub-sample.

The information needs describe which information the user likes to receive from the system and should align with the information the service delivers and captures. Our interview explicitly covers the source and completeness of the information, while implicitly reviewing other information quality measures such as accuracy and timeliness. Our literature study demonstrated several connections between information quality and factors directly influencing the adoption of a service, leading to the following hypothesis:

The better the fit between information needs and information provided by the e-health service, the higher the user adoption of the service.

More than half of the interviewees offered negative feedback specific to the hypothesis as (1) the system was only available in English, (2) the systems used too many medical terms, (3) the quality of data input by patients, and the (4) belief that current information along was enough.

For each success factor in Table 1 we again identified specific results, which we summarize in Table 5. Privacy concerns again emerged as the greatest threshold for users with 23 out of 27 noting the issue as a significant concern. Out of all of the interviews emerges the view that users consider health information as very personal with a commercial company like Google untrustworthy when it comes to such information.

<table>
<thead>
<tr>
<th>Age</th>
<th>Amount</th>
<th>Gender</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>15-25</td>
<td>15</td>
<td>Male</td>
<td>11</td>
</tr>
<tr>
<td>25-45</td>
<td>6</td>
<td>Female</td>
<td>16</td>
</tr>
<tr>
<td>45+</td>
<td>6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Demographic data Google Health data.

The usefulness of the Internet-based service emerged as the second significant issue within among the Google among participants. Despite some positive reactions, most of the interviewees failed to see direct value for themselves as result of using Google Health. Most noted not holding their own health information, thus mitigating the need to in the future, consistent with low compatibility with current practices. Moreover, participants noted their own relative good health as a reason to not use such a PHR system.

With respect to barriers to using Google Health, one participant noted, “…in addition to the fact that I don’t have any information to put onto Google Health, I really would want privacy guarantees before putting my information into the system to prevent my information getting public on the internet”. This sentiment illustrates the general opinion emerging from the collective interview data.

We highlight the main problems, or objection points, in Table 5. Users do not see the relevance of Google Health, expressing a predominantly negative attitude toward the product.
5. Analysis

The relevance of new information systems is seen as one of the most important success factors. Our results see perceived usefulness [5], an indicator of relevance, negative in both the Google Health and Microsoft HealthVault cases. Despite positive perceptions of the information quality of Microsoft HealthVault, low relevance makes it potentially difficult for the product to ultimately succeed.

In both cases, Google Health and Microsoft HealthVault, interviewees viewed both system quality and service quality positively. Consistent with Delone and McLean [16], together with the positive information quality, favorable views of system and service bode well for the ultimate success of Microsoft HealthVault. That said, we contend that our study, not grounded in a tethered system context, shows less likely chances for success. Based upon the factors identified within our literature study, we question whether the IS success model should consider trust as fourth causal determinant.

The trust and perceived risk indicators are closely related in both the Google and Microsoft cases with trust negative and risk high. Turban [50] demonstrated that Internet-based information systems depend on trust to ultimately achieve success. Accordingly, as a driver of adoption, high risk and low trust serve as significant hindrances for both Google Health and HealthVault to achieve widespread adoption.

As with many studies, social and personal influence as an influence factor within our study [6]. Our qualitative analysis found social and personal influence negative for both Google Health and Microsoft HealthVault.

Neither system is well known or widely used in The Netherlands, where the current study was conducted. Therefore, few peers or healthcare professionals likely promote either system. Moreover, conducting this work outside of a tethered PHR system context; i.e., not recruiting participants through a healthcare system, hospital, or other healthcare related environment; likely impacts the healthcare provider social and personal influence commonly experienced by participants. Here again, our context mirrors the general population of potential standalone PHR users.

6. Conclusions

As we already know, Google Health did not succeed. Our analyses based upon the literature study and interview data supports the contention that Microsoft HealthVault will similarly fail to reach a critical mass of users, at least as a standalone product offering in a market such as The Netherlands.

The primary reason we see for the lack of success lies in the relevance of both PHR applications. Relevance, perceived usefulness, performance expectancy, relative advantage, and net benefits all focus on the value of a new information system. These factors have failed to draw the attention of vendors developing standalone personal health records applications.

Low trust and high risk emerge as two additional significant reasons for the failure of personal health records, given the analysis of the Google and Microsoft products. Together these factors might be considered as a “new” success factor and validated for standalone as well as tethered applications.

Further, healthy people are unaware of personal health records, and therefore, social and personal influences will likely play a negative role in the success of Microsoft HealthVault and serve as one potential reason for the failure of Google Health. While healthy individuals may not see a need for or interest in PHRs, demonstrating benefits of long-term and accurate PHRs might frame such applications as a beneficial planning tool or later periods in one’s life.

In addition to continuing to examine the PHR phenomenon outside strict tethered system contexts, future studies of personal health records should focus on designing applications that address factors specifically relevant to patient and caregiver end users within untethered contexts. Trust and risk should further be studied and validated as a significant determinant of success under the greater umbrella of reliability.
Limitations
This study aims for a qualitative reasoning why PHR has not been successful so far. The amount of interviews (51) is valid to support this reasoning but not enough to generalize to larger populations or dissimilar contexts.

7. References


the American Medical Informatics Association, 18(2): 118-124.


Miles, M.B., and Huberman, A.M. *Qualitative data analysis: An expanded sourcebook* Sage, 1994.


### Appendix I: Interview framework

<table>
<thead>
<tr>
<th>PRIMA construct</th>
<th>Success factors expected to be measured</th>
<th>Examples of questions asked</th>
</tr>
</thead>
<tbody>
<tr>
<td>Process</td>
<td>Perceived compatibility</td>
<td>Which health records you regularly use? Are you using a fixed sequence of actions? Which alternatives you have to find information?</td>
</tr>
<tr>
<td>Relevance</td>
<td>Perceived usefulness</td>
<td>Which functions of a PHR are most important for you? Which parts of the system you experience as a bottleneck? Do you have suggestions for improvements?</td>
</tr>
<tr>
<td>Information needs</td>
<td>Information quality</td>
<td>Which information you need to get from the service? Do you get sufficient information from the system? Is the information quality sufficient?</td>
</tr>
<tr>
<td>Means and people</td>
<td>Service quality</td>
<td>Do you get sufficient support? Is the system reliable? Does the system offer enough privacy?</td>
</tr>
<tr>
<td>Service quality</td>
<td>System quality</td>
<td>Do you get sufficient support? Is the system reliable? Does the system offer enough privacy?</td>
</tr>
<tr>
<td>Perceived risks</td>
<td>Perceived compatibility</td>
<td>Which health records you regularly use? Are you using a fixed sequence of actions? Which alternatives you have to find information?</td>
</tr>
<tr>
<td>Trust</td>
<td>Social and personal influence</td>
<td>Do you think IT is necessary to improve health information? Do you feel social pressure of using the service? How much time do you want to spend for learning to use the service?</td>
</tr>
<tr>
<td>Social and personal influence</td>
<td>Perceived compatibility</td>
<td>Which health records you regularly use? Are you using a fixed sequence of actions? Which alternatives you have to find information?</td>
</tr>
</tbody>
</table>