

# Predictions on Service Adoption and Utilization Meet Reality

## First Results from the Sciebo (Science Box) Project

Raimund Vogl<sup>1</sup>(✉), Holger Angenent<sup>1</sup>, Dominik Rudolph<sup>1</sup>,  
Andreas Wilmer<sup>1</sup>, Anne Thoring<sup>1</sup>, Stefan Stieglitz<sup>2</sup>,  
and Christian Meske<sup>2</sup>

<sup>1</sup> University of Münster, Münster, Germany  
{r.vogl,h.angenent,d.rudolph,a.wilmer,a.thoring}  
@uni-muenster.de

<sup>2</sup> University of Duisburg-Essen, Essen, Germany  
{stefan.stieglitz,christian.meske}@uni-due.de

**Abstract.** A large academic cloud storage service was launched in the beginning of 2015 by the majority of the public research and applied science universities in the German state of North Rhine-Westphalia (NRW) under the brand name “sciebo”. One year after the start, we will examine if the predictions made on service adoption in the preparatory project phase based on the well-known diffusion model by Rogers apply to reality. This is the first study about the adoption of a specific cloud service at several universities. We identify two factors affecting the speed of diffusion: share of technophiles and the use of marketing measures. Organization size does not seem to influence the speed of diffusion .

**Keywords:** Diffusion of innovations · Adoption · Cloud-storage · Project report

## 1 Introduction

Cloud storage services like Dropbox or Google Drive became quite popular in the course of the last half decade, not least in the academic context among students and researchers, making it possible to easily share documents with others and to synchronize data across multiple devices. Those commercial services are very comfortable in use, but security concerns about their data utilization arise, especially after the Snowden disclosures. In 2013, as a consequence, the majority of the public research and applied science universities in the German state of North Rhine-Westphalia (NRW) formed a consortium to start a jointly operated private cloud service for the academic community. This sync and share storage platform should be free of charge, easy to use and, most importantly, it should be hosted on premise at several university data centers to be fully compliant with German data protection regulations [1]. With respect to the software functionality and the required hardware setup for potentially 500.000 users, the system design was grounded on empirical user studies.

A first exploratory survey on the demand for a university operated alternative to Dropbox etc. was conducted among potential users at Münster University in 2012 and extended to a multi-site survey with more than 10.000 participants from three major universities in late 2013 [2]. Both surveys focused on the participants' intention to use such a university operated cloud service, their demand for storage space and client platforms, the type of content (file types) they intended to store, and the communities they wanted to collaborate with using the service's file sharing functionalities. The procurement of the software solution as well as the sizing of the hardware platform were based on the adoption and usage estimates derived from these surveys.

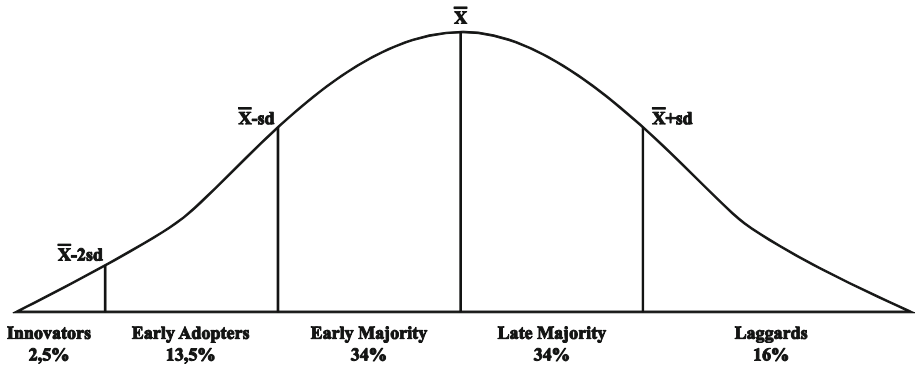
In February 2015, after extensive preparatory work done for the funding proposal, the procurement process, and the system setup and commissioning, the sync and share cloud storage service was launched under the brand name "sciebo – theCampuscloud" (sciebo being short for science box) with three university data centers (Bonn, Duisburg-Essen and Münster) hosting the system platforms on premise. Almost exactly one year after the start, it is now the right time to review how the initial expectations on service adoption and usage correspond with reality. After a year of operation (as of Feb 02 2016), exactly 40.000 users from 24 universities (out of 33 in NRW) and one public research center have signed up for sciebo through the self-enrollment web portal.

The case of sciebo is unique because it allows us to observe the diffusion of a technical innovation from the beginning in a well-controlled setting. There is plenty of literature about the adoption of cloud systems in organizations like SMEs [3–5] or special industries [6–13], but only little is known about the adoption behavior of end-users who can decide freely if they want to use a new cloud service or not [14]. Universities are a special case: On the one hand, they are organizations with a quite uniform population and a manageable size. On the other hand, because of the principle of freedom of research and teaching held high in Germany, there is no possibility to command the use of a system, so users have to be convinced.

## 1.1 Predictions

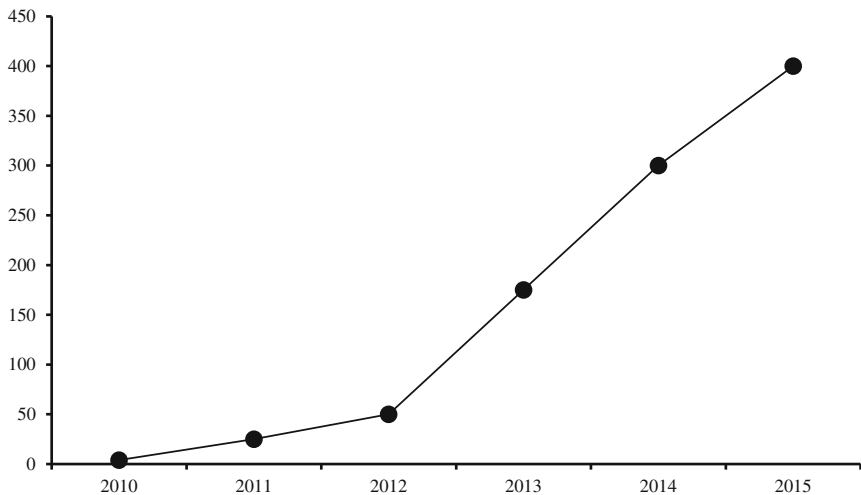
According to Rogers [15], innovativeness, i.e. the readiness and the degree to which a person or an organization adopts an innovation (i.e. a new product) compared with the other members of his population, follows the Gaussian distribution. He identifies five adopter categories (innovators, early adopters, early majority, late majority, laggards; see Fig. 1) who have different characteristics referring to their innovativeness. If you accumulate the adoption decisions of all those adopters over time, you will get an S-shaped curve, the diffusion curve. The faster the innovation is adopted the more steeply this curve will rise. The speed of diffusion depends on the characteristics of the innovation, in particular its relative advantage compared to other existing products, compatibility with existing values and practices, simplicity and ease of use, trial ability and observable results.

As our data from a large user survey conducted in 2013 [2] show, Dropbox is the most used storage service among members of the universities with about 80 percent of market share. This value recurs in another survey conducted in 2015 at the same universities (unpublished work), so we conclude that Dropbox obviously has reached



**Fig. 1.** Adopter Categories according to Rogers [15]

the saturation of demand five years after its inception in 2007 and two years after the release of the first stable version 1.0 in 2010. Examining Dropbox’s worldwide diffusion (see Fig. 2), a flat growth is visible in the first two years and a take-off in the third year.



**Fig. 2.** Diffusion of Dropbox (in mio. users) [16]

For sciebo, we predicted an even faster diffusion because the technology is already known from Dropbox. Moreover, sciebo’s high security standards and bigger free storage space seem to be significant relative advantages as stated by the participants of a survey in 2013 [2]. According to Diffusion Theory, market potential is not the total of all potential users (i.e. all members of the participating universities), but the total of all those persons who will realistically use a new service. In the survey, 92.5 percent of the participants stated that they wanted to use sciebo. Being informed that their usage

authorization would be revoked when leaving university, the count dropped to 65 percent. Thus, 65 percent of all members of the participating universities – that is about 252.000 individuals – constitute the estimated market potential of sciebo.

Based on the distribution of per user storage demands from the survey, we could refine the initial assumption that each user would utilize the planned 30 GB quota to the max and were able to predict an average storage volume of 8 GB (pessimistic scenario) to 16 GB (optimistic scenario) per user and could ascertain that a maximum storage space of 30 GB should fit most users. Assuming that users would switch their academic data from another platform to sciebo in the first days after the registration, we expected a quite linear growth with a 30 percent basis synchronization at the beginning and just small gain of 3 percent a month.

Considering the predictions on service adoption and storage demand, different scenarios were derived to estimate the size of storage systems to be procured and the internet bandwidth required. The total storage volume required for the operation of sciebo in the long term was estimated at 1.7 PB (pessimistic) to 5 PB (optimistic), and the internet connection bandwidth requirement for service operation was estimated at 3 Gbps in the optimistic scenario.

## 2 Findings

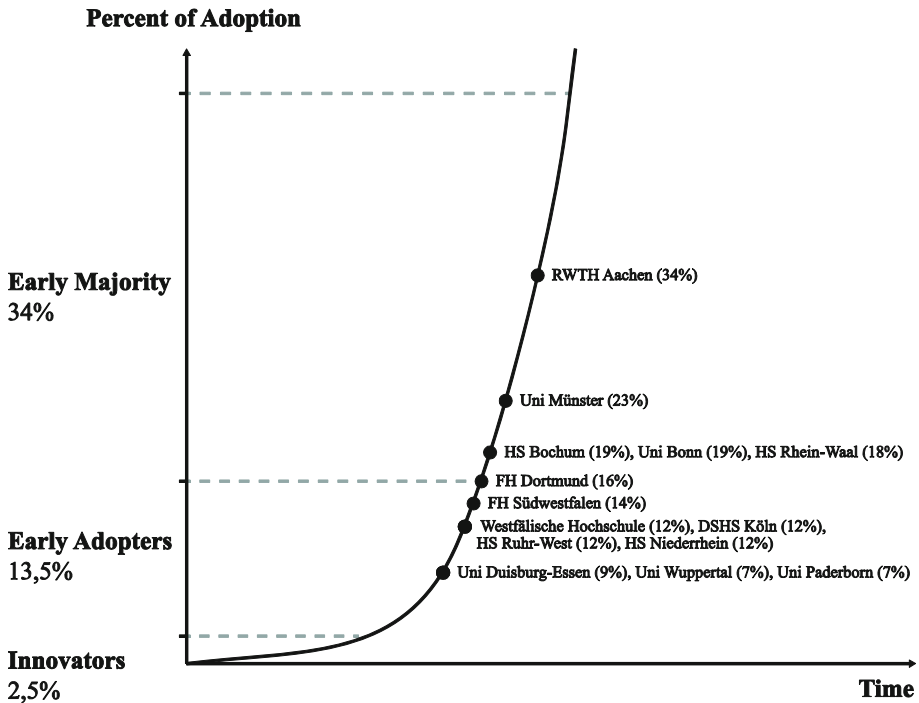
### 2.1 User Diffusion

Nearly one year after the official launch, sciebo hit another milestone with now 40.000 users – this means an actual market share of 17.3 percent. In terms of the Diffusion Theory, this implies that sciebo's diffusion has already reached the early majority phase.

However, diffusion speed varies significantly at the different universities. Figure 3 shows the state of diffusion at the 14 universities that started sciebo in Feb 2015. The spectrum ranges from 6.7 percent at the University of Paderborn to 33.9 percent at RWTH Aachen. University size might serve as one explanation, as information should flow very fast at a small campus with a manageable number of departments. As stated by Rogers [15], diffusion can be seen as a communication process: In smaller and spatially closer populations, communication between the members is much more likely and easier than in a complex university with lots of different departments distributed over the whole city.

Though in theory size suggests itself as a reason for the different diffusion speeds, it does not seem to be a good explanation in our case: Comparing same-sized universities – e.g. the Universities of Münster, Duisburg-Essen and RWTH Aachen with about 44,000 to 49,000 members each (see Fig. 4) – the differences in market share are still evident. Results show a remarkable variance of 24.5 percent between RWTH Aachen (33.9 %) and the University of Duisburg-Essen (9.2 %), with the University of Münster (23.0 %) ranking mid.

Taking all universities into account, RWTH Aachen appears to be an outlier with its high market share. Both, the University of Münster (23.0 %) and the University of Duisburg-Essen (9.2 %), rank much closer to the overall average. One possible



**Fig. 3.** Diffusion of sciebo after one year (only universities starting in Feb 2015)

explanation for RWTH's high performance is that, unlike the Universities of Münster and Duisburg-Essen, RWTH is a technical university with many technophiles. They resemble the innovators described by Rogers and are the first to adopt new technologies. Logically, a technical innovation like sciebo diffuses faster in a technophilic environment than in other populations.

The low performance of the University of Duisburg-Essen, compared with the same-sized universities and the overall average diffusion, is similarly interesting. A closer look reveals that the universities' commitment in terms of marketing activities might be another decisive factor. While RWTH Aachen and the University of Münster, in particular, performed a variety of marketing activities (i.e. direct mailings to all members), the University of Duisburg-Essen did not to that extent. Therefore it is likely that only innovators and early adopters who are interested in innovations and actively search for information on their own account for their share of sciebo users. Further monitoring will show if an early majority can be reached with no marketing and just word of mouth, or if the number of users will be stagnating. According to some authors there is a gap between the early adopters and the early majority which has to be bridged by marketing activities [17, 18], while Rogers [15] considers both groups as a continuum.

Examining the diffusion curves of the different universities (see Fig. 4), deviations from the ideal S-curve of the diffusion model are clearly visible. Usually, they are

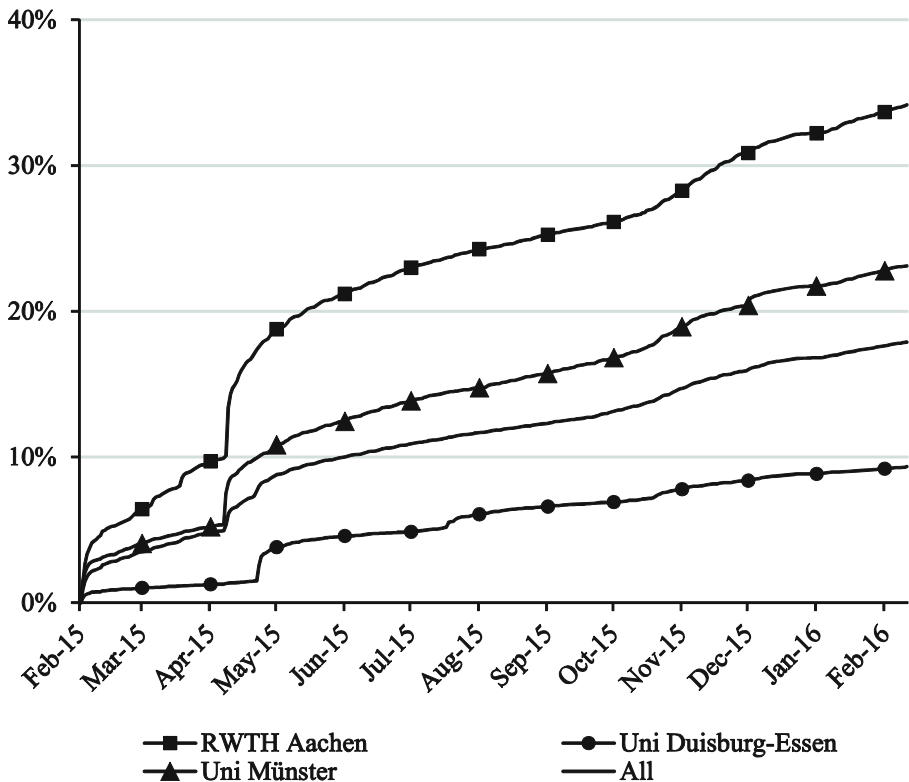


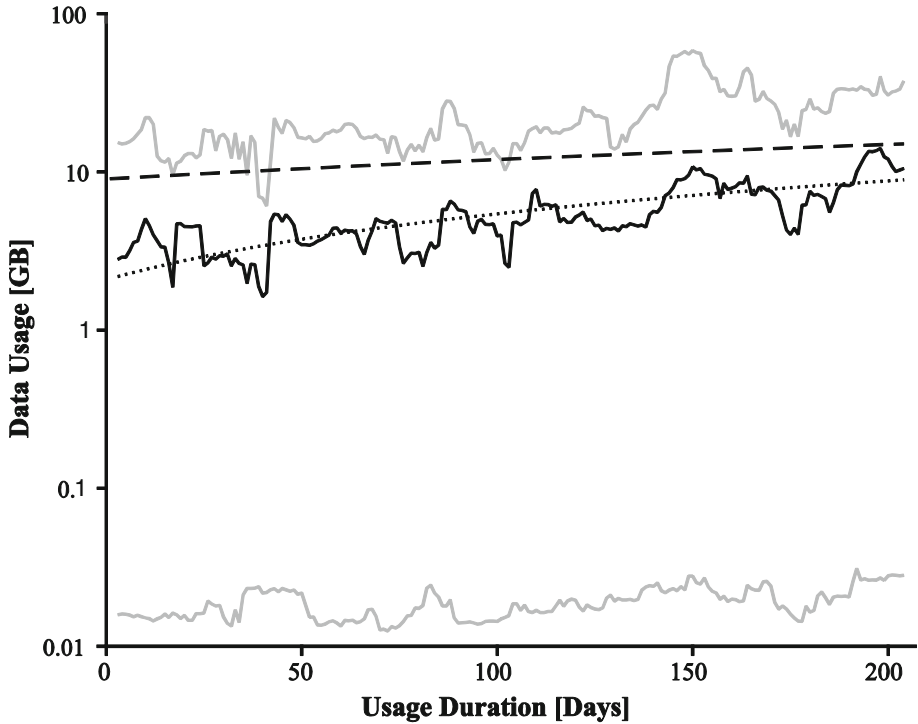
Fig. 4. Diffusion curves of chosen universities with same size starting in Feb 2015

caused by special events. The first boost in February 2015 is the official launch of the service. In the run-up we realized a large Facebook campaign with posts in over 400 user groups related to the participating universities. Also, test users were now added to the statistics. The second and largest user increase in April, at the start of the summer term in Germany, is triggered by direct mailing, that most participating universities did send to their members. The diffusion curves of those universities passing up this opportunity show no such steep rise. In October most universities welcome their largest share of new students for the winter term, explaining the next boost. In December, some universities used direct mailing to promote an online survey related to sciebo, again gaining attention and an additional boost for new users for sciebo.

As regards storage space, we initially expected 9 GB (30 % of the intended per user quota limit of 30 GB) right after registration and a monthly growth of 3 percent (until the quota limit is reached). Currently, the average volume needed by an active user (i.e. a user who uploaded some data) is 3.3 GB, amounting to a total of 99.8 TB storage space used in sciebo. Those universities which grant access to sciebo only to their staff have a substantially higher storage demand per active user (e.g. Düsseldorf University with 7.4 GB) than most other universities where usually three out of four active users are students (e.g. Münster University with 4.7 GB).

## 2.2 Data Storage

In Fig. 5 we analyzed the storage load on an individual user basis. In particular, we looked at the dependency between the consumed storage space of a single account and its age. Shown is the mean used disk storage for user accounts in dependency of the account lifetime (solid black line), the 0.05-quantile (lower grey line) and the 0.95-quantile (upper grey line) in a logarithmic plot. The broken black line represent the expected and the dotted black line the observed linear model of the user behavior.



**Fig. 5.** Storage load on individual user basis per time vs. model (broken line)

Altogether 6,581 user accounts were analyzed on a day-wise basis. The statistical values were computed across an ensemble of user accounts for a specific account age. In addition, two data-sets from different time points are independent from each other because accounts were not tracked over time. The analysis was restricted to active accounts with a used storage capacity of more than 10 MB. Thus, inactive accounts from seasonal side effects, such as beginning of a new semester, are excluded from analysis. In addition, a moving average with a window size of 7 days was used to accumulate the number of user accounts for statistical analysis, i.e. in average  $N = 225 \pm 92$  accounts were analyzed for each day.

We observed two main findings: First, we predicted that on average an account initially requires 30 percent of its full 30 GB quota and grows in a linear fashion with 3

percent of its quota per month. One can rewrite this assumption to a linear equation of the form  $f(x) = A + Bx$  with the function  $f$  describing the disk usage in dependency of the time  $x$  and the coefficient  $A$  as the initial off-set,  $B$  as the slope of the function, i.e. we expected  $A_{exp} = 9000$  [MB] and  $B_{exp} = 29.6$  [MB/Day]. However, we observed an offset  $a_{obs} = 2077.1 \pm 239.3$  [MB] and a slope  $b_{obs} = 33.5 \pm 2.0$  [MB/Day] with a linear Least-Squares Fit ( $p < .001$  and adjusted R-Squared 0.578). The observed results show that on average a user synchronizes less data directly after the subscription than initially expected, but fairly consistent with 30 % of the average storage space per user of 8 GB in the pessimistic scenario deduced from the survey findings. However, the growth of the data synchronized is higher than expected.

Second, sciebo has to handle a variety of usage scenarios. In Fig. 5 the close distance of the average to the 0.95-quantile indicates a positively skewed underlying distribution, which is caused by an extensive disk usage of some few accounts. This indicates on the one hand that usage scenarios will differ in strong fashion between users and, on the other hand, that sciebo is capable to deal with a wide variety of use cases.

### 2.3 Bandwidth

The initial estimates of bandwidth requirements were essential to make sure that the internet connection bandwidth of the three university data centers hosting the sciebo platform was not entirely consumed by the new sciebo service. Based on simply models of service utilization (up- and downloads) an overall limit of 3 Gbps sustained for the whole sciebo system, thus approx. 1 Gbps for each of the datacenter sites, was predicted as being sufficient.

One year after the start of operation, this sustained data rate has not been reached by far, but temporary bandwidth peaks at each of the three sites are in the 800 Mbps range (see Fig. 6). With continuous growth of the sciebo user base and storage volume, bandwidth demands will necessarily grow, but negative effects on the internet connectivity of the hosting universities (each currently has a 10 Gbps internet link) are, as initially predicted, not to be expected, especially since traffic policies limiting the bandwidth allocated to individual connection could still be imposed. The mutual data backups between the three sites are schedules in the 12am to 6am timeframe where service utilization is low and thus do not negatively impact the bandwidth budget.

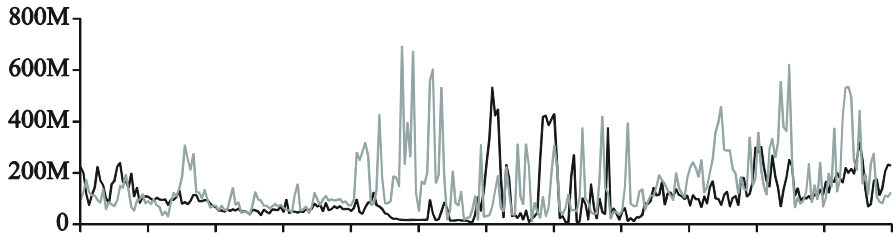
### 2.4 Additional Findings

Apart from those findings related to our predictions, some additional outcomes are worth mentioning. The first finding broaches the issue of user activity: 38 percent of the registered users are inactive, i.e. they have not uploaded any data to sciebo yet. Based on Rogers' Diffusion Theory [15], this inactivity of a substantial user fraction could be interpreted as either a prolonged phase of decision making or as discontinuance (without having used the service apart from signing up) [19, 20]. This finding needs further research.

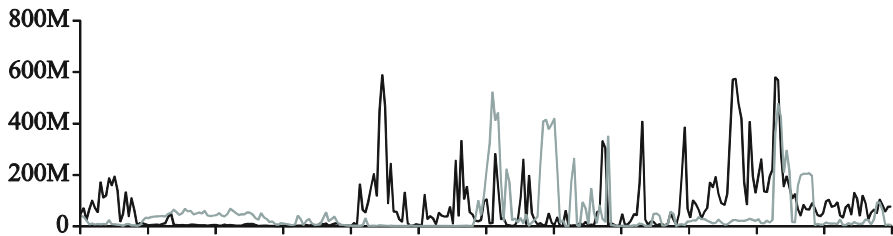


**HP59SNS-BN-Bridge-Aggregation10-972**

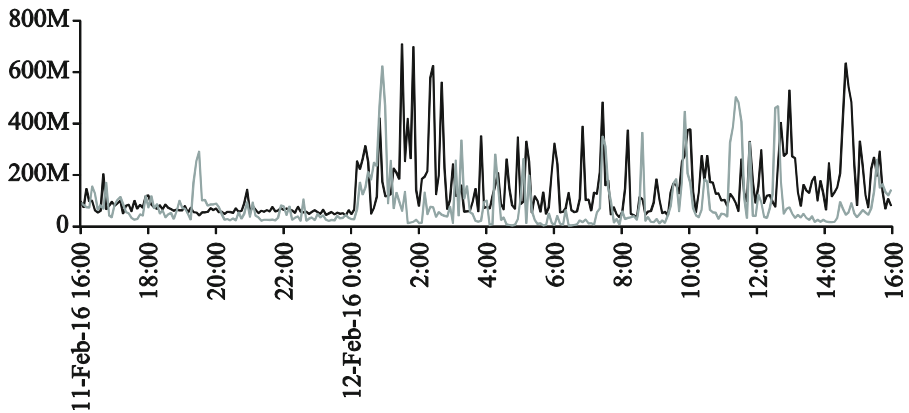
BW: 20.0 G

**HP59SNS-DU-Bridge-Aggregation10-972**

BW: 20.0 G

**HP59SNS-MS-Bridge-Aggregation10-973**

BW: 20.0 G

**Time**

**Fig. 6.** Internet bandwidth (in = black, out = grey) consumed by the 3 sciebo sites at Bonn (BN), Duisburg-Essen (DU) and Münster (MS) on a typical day after one year of operation. Correlated in/out peaks between two sites in the 12am – 6am timeframe are due to mutual backup copy operation.

The second finding focuses on the key collaboration feature of sciebo – sharing data with other sciebo users or externals (share via hyperlink). With an overall average of 2.4 shares per active user, this feature is not used very strongly yet. Folders (66.5 %) are shared more often than files (33.5 %). Approximately 50 percent of all shares are

performed via link (primarily intended for external exchange), contrary to expectations from the survey [2], where 65 percent of the participants intended to share within their university and only 21 percent intended to share with externals.

### 3 Conclusion

These first results shows that the predictions phrased in the aftermath of the 2013 survey [2] are – up to now – in line with the service’s adoption, and, moreover, Rogers’ diffusion theory has proved to be an adequate model. We could identify two factors influencing the speed of diffusion of the sciebo cloud-service:

1. Share of technophiles in the organization
2. Use of marketing measures

Both findings are supported by the diffusion model. As known from the diffusion literature, an innovation is more likely to be adopted if it is not too complex and consistent with known products. Consequently, technophiles who understand a technical innovation much better and usually find it less complex than other people, will be more likely to adopt an innovation quickly. As noted by some authors, there might be a gap – in terms of missing peer-to-peer connections – between innovators and early adopters on the one hand and the early majority on the other hand, because of the significant differences between those groups [17]. Marketing measures like direct mailings, Facebook posts, YouTube videos etc. can bridge this gap by informing the early majority about a new service, and thus speed up the diffusion process. According to our data, organization size does not influence the diffusion speed.

Finally, the universities’ heterogeneous rate of adoption and the high fraction of inactive users leave a wide field for further research. In the upcoming months, analyzing the reasons for discontinuance of use will be a key focus.

### References

1. Vogl, R., Angenent, H., Bockholt, R., Rudolph, D., Stieglitz, S., Meske, C.: Designing a large scale cooperative sync&share cloud storage platform for the academic community in northrhine-westfalia. In: Sukovski, U. (ed.) *ICT Role for Next Generation Universities - 19<sup>th</sup> European University Information Systems*. Riga Technical University, Riga (2013)
2. Stieglitz, S., Meske, C., Vogl, R., Rudolph, D.: Demand for cloud services as an infrastructure in higher education. In: *Proceedings of International Conference on Information Systems, ICIS 2014* (2014)
3. Trigueros-Preciado, S., Pérez-González, D., Solana-González, P.: Cloud computing in industrial SMEs: identification of the barriers to its adoption and effects of its application. *Electron. Mark.* **23**, 105–114 (2013)
4. Tehrani, Shima Ramezani, Shirazi, Farid: Factors influencing the adoption of cloud computing by small and medium size enterprises (SMEs). In: Yamamoto, Sakae (ed.) *HCI 2014, Part II. LNCS, vol. 8522*, pp. 631–642. Springer, Heidelberg (2014)

5. Alshamaila, Y., Papagiannidis, S., Li, F.: Cloud computing adoption by SMEs in the north east of england: a multi-perspective framework. *J. Enterp. Inf. Manage.* **26**, 250–275 (2013)
6. Oliveira, T., Thomas, M., Espadanal, M.: Assessing the determinants of cloud computing adoption: an analysis of the manufacturing and services sectors. *Inf. Manage.* **51**, 497–510 (2014)
7. Moryson, H., Moeser, G.: Consumer adoption of cloud computing services in germany: investigation of moderating effects by applying an UTAUT model. *Int. J. Mark. Stud.* **8**, 14 (2016)
8. Low, C., Chen, Y., Wu, M.: Understanding the determinants of cloud computing adoption. *Ind. Manage. Data Syst.* **111**, 1006–1023 (2011)
9. Lian, J.-W., Yen, D.C., Wang, Y.-T.: An exploratory study to understand the critical factors affecting the decision to adopt cloud computing in taiwan hospital. *Int. J. Inf. Manage.* **34**, 28–36 (2014)
10. Khanagha, S., Volberda, H., Sidhu, J., Oshri, I.: Management innovation and adoption of emerging technologies: the case of cloud computing. *Eur. Manage. Rev.* **10**, 51–67 (2013)
11. Ivanova, M., Ivanov, G.: Cloud computing for authoring process automation. *Procedia-Social Behav. Sci.* **2**, 3646–3651 (2010)
12. Gupta, P., Seetharaman, A., Raj, J.R.: The usage and adoption of cloud computing by small and medium businesses. *Int. J. Inf. Manage.* **33**, 861–874 (2013)
13. Cegielski, C.G.: Allison Jones-Farmer, L., Wu, Y., Hazen, B.T.: Adoption of cloud computing technologies in supply chains: An organizational information processing theory approach. *Int. J. Logistics Manage.* **23**, 184–211 (2012)
14. Shin, J., Jo, M., Lee, J., Lee, D.: Strategic management of cloud computing services: focusing on consumer adoption behavior. *IEEE Trans. Eng. Manage.* **61**, 419–427 (2014)
15. Rogers, E.M.: *Diffusion of Innovations*. Free Press, New York (2003)
16. Dropbox, <https://www.dropbox.com/>
17. Moore, G.A.: *Crossing the Chasm. Marketing and Selling Disruptive Products to Mainstream Customers*. Harper Business, New York (2014)
18. Christensen, C.M.: *The Innovator's Dilemma. The Revolutionary Book that Will Change the Way You do Business*. Harper Business, New York (2011)
19. Parthasarathy, M., Bhattacharjee, A.: Understanding post-adoption behavior in the context of online services. *Inf. Syst. Res.* **9**, 362–379 (1998)
20. Black, W.: Discontinuance and Diffusion: Examination of the Post Adoption Decision Process. *Adv. Consum. Res.* **10**(1), 356 (1983)