Willing Buyer, Willing Seller: Personal Data Trade as a Service

Lindah Kotut
Department of Computer
Science
Virginia Tech
Blacksburg, VA 24060, USA
lkotut@vt.edu

Timothy L. Stelter
Department of Computer
Science
Virginia Tech
Blacksburg, VA 24060, USA
tstelter@vt.edu

Michael Horning
Department of Communication
Virginia Tech
Blacksburg, VA 24060, USA
mhorning@vt.edu

D. Scott McCrickard
Department of Computer
Science
Virginia Tech
Blacksburg, VA 24060, USA
mccricks@vt.edu

Abstract

There is an increased sensitivity by people about how companies collect information about them, and how this information is packaged, used and sold. This perceived lack of control is highlighted by the helplessness of users of various platforms in managing or halting what data is collected from/about them. In a future where users have wrested control of their data and have the autonomy to decide what information is collected, how it is used and most importantly, how much it is worth, a new market emerges. This design fiction considers possible steps prescient companies would take to meet these demands, such as providing third-party subscription platforms offering personal data trade as a service. These services would provide a means for transparent transactions that preserve an owner's control over their data; allowing them to individually make decisions about what data they avail for sale, and the amount of compensation they would accept in trade.

Author Keywords

Design Fiction; Data Ownership; Privacy; Hiking; Ethics; Appalachian Trail

CCS Concepts

•**Security and privacy** → **Economics of security and privacy;**

Data Subscription Service: An Introduction

With the increased digitization of personal information, companies behind mobile applications and online services collect users' data for profit. In hiking communities for example, these practices are useful primarily because of the hikers' unique data needs and also the wide variety of data that they generate while on the trail. By companies packaging and reselling these types of data, the hikers from whom the data was harvested have no understanding and control over what information is exactly collected, who it is sold to and how it is used past the selling.

Hikers' data are useful for variety of applications: mapping terrain, reporting weather conditions, measuring body performance among others [11, 10]. Applications that assist the hiker in achieving these goals also tend to package the hiker's data for sale to other companies. The monetary or equivalent benefit the company receives as a result of collected data is not passed back to the hiker.

There is however an increased awareness of how the user has no control over personal data, knowing they are the commodity for sale [7]. This increased knowledge and understanding of how users are a cog in the wheel of "surveillance capitalism" [12] has shown to lead users to abandon those services that collect the data about them in search of those services that allow them to retain personal control over their data [15]. This trend will only continue as more users disagree with how their data is collected and used. As this trend shows, there is a need for platforms that allow users to control information collected about them and provides absolute transparency of how their information is used.

Our subscription service goes a step further, we allow the users to directly and explicitly *sell* their data. As the "free" services that turn users into products have shown, there is a market for data. We remove the middle-man in the transaction and allow the user to *sell* their information directly to buyers. Depending on perceived need and the amount of compensation, we allow the willing seller to determine whether it is worth a trade of personal information [14].

Addressing Need (Product Design)

Using a popular hiker saying *"The Trail Provides"* [3] used to denote that a hiker's need will be met when they most need it, we propose a data sharing and incentive model. This model relies on *"willing buyer, willing seller"* revised policy that allow buyers to bid for data they wish to know, and/or also allowing the seller to set the price for data they produce, such that both the buyer and the seller are able to exchange data at the fair market rate [19]. Using *The Trail Provides* principle, we can showcase the benefit that our subscription provides to both parties in the trade. A platform to sell data allows the sellers (hikers) to receive compensation for data and also ease the cost of current or future hikes, while at the same time satisfy pre-hikers' (buyers) need for nuanced and contextual information to inform their own curiosity about the trail, or to use to inform their own future hikes.

The Incentive model

Our subscription service offers the option to package the data for sale. For convenience, we provide initial data tiers that the sellers could adopt and allow the buyers to purchase subscriptions for access. We use a simple pricing formula: the more intimate the data, the more expensive it is; the more difficult it is to obtain the data, the more expensive it is. These considerations allow for both a fair return on investments to the sellers and as an incentive model for buyers to access only what they need. Figure 3 provides the entire data packaging that we offer.
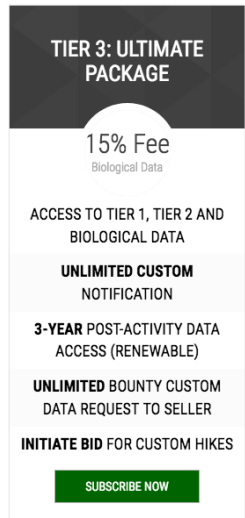

Figure 1: To willing buyers: A conceptual model of the subscription service presented in a tiered data packaging. In this tier, a buyer can purchase access to environmental, situational and biometric data. A 15% fee is applied by the company to offset the cost collecting, warehousing, processing and disseminating data and insights. Figure 3 showcases how the rest of the tiers will be packaged
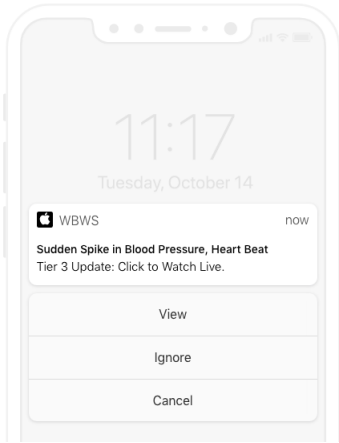
Figure 2: To willing buyers: Live Push notification customized to alert for unexpected events

The willing seller, such as a hiker, also has the option to set their own price, or use recommended pricing depending on market sources. This can be done before the hike commences, during the hike, or even at the end of hike. We also provide a means for third parties to also supply *milestone funds* to incentivize sellers to achieve a set milestone to complete, that is in line with the companies' needs (e.g. sensor data to measure shoe brand performance for a 500-mile hike).

Data Sources and Data Integrity
In order to preserve data integrity of real-time data [17], our service automatically collects from the seller and uploads to our cloud servers at frequent intervals making it easy for the seller to share the data with subscribers (we also provide sensors for rent for a small fee). Our system also provides automatic visualization so that the buyers can either view real-time information on a dashboard (Figure 5) or set alerts that will automatically notify them in real-time whenever a triggering event occurs (Figure 2).

Our company also has sensors and specialized equipment that the seller can purchase or rent. Our sensors use standard universal APIs to enable plug-and-play uses by sponsors and to account for atypical sensors. Figure 4 shows different placement of sensors that would automatically collect and disseminate data without interrupting the hiker's pace and/or trail experience.

While on the hike, light-weight, and non-power-intensive sensors supply data to a connected SIM-enabled command-and-control hub, that subsequently uploads this data at set intervals or streamed live. If connection is lost, uploads would proceed automatically once the connection is re-established. This makes it easier for the seller not to worry about missing updates, while allowing the data to receive just-in-time, reliable information.

Dashboard
Our subscription service provides a seamless interface between the buyer and the seller. We are transparent in how we make money: from sellers, by fees from rented equipment; from buyers, in the form of subscription fees (Figure 3) to process payment, store data and provide just-in-time updates. Importantly, our service provides a means for the the sellers to receive continuous income for data that would otherwise be forgotten after long-term disuse [8].

Both the sellers and the buyers have access to the dashboard where the bidding process for data and terms of use are set and agreed upon. The dashboard provides a visual of the escrow service we provide, to allow for a trustworthy stewardship that also provides for a seamless release of funds at milestone completion that is automatically updated and triggered from the sensor data we receive from the seller.

Both the buyer and the seller can see all the personal information about the seller according to the tier the buyer has subscribed to, such as real-time location updates, last meal eaten, the number of bathroom breaks taken, the elevation gains, the pace etc. The dashboard also provides convenience by offering visualization options for data received from the seller. Figure 5 provides a dashboard from the buyer's point of view that shows how both distance and time can be viewed to provide context on the progress of a seller who is in the process of completing their hike.

To be in the know, the buyer can simply set notification for events of interest that can be received as a notification (Figure 2) with prompts to monitor events live as they happen. We provide add-on services that follow the concept of in-app purchases prevalent in mobile apps. This provides the seller with an additional income stream by setting access to some additional information for a premium, which allow the
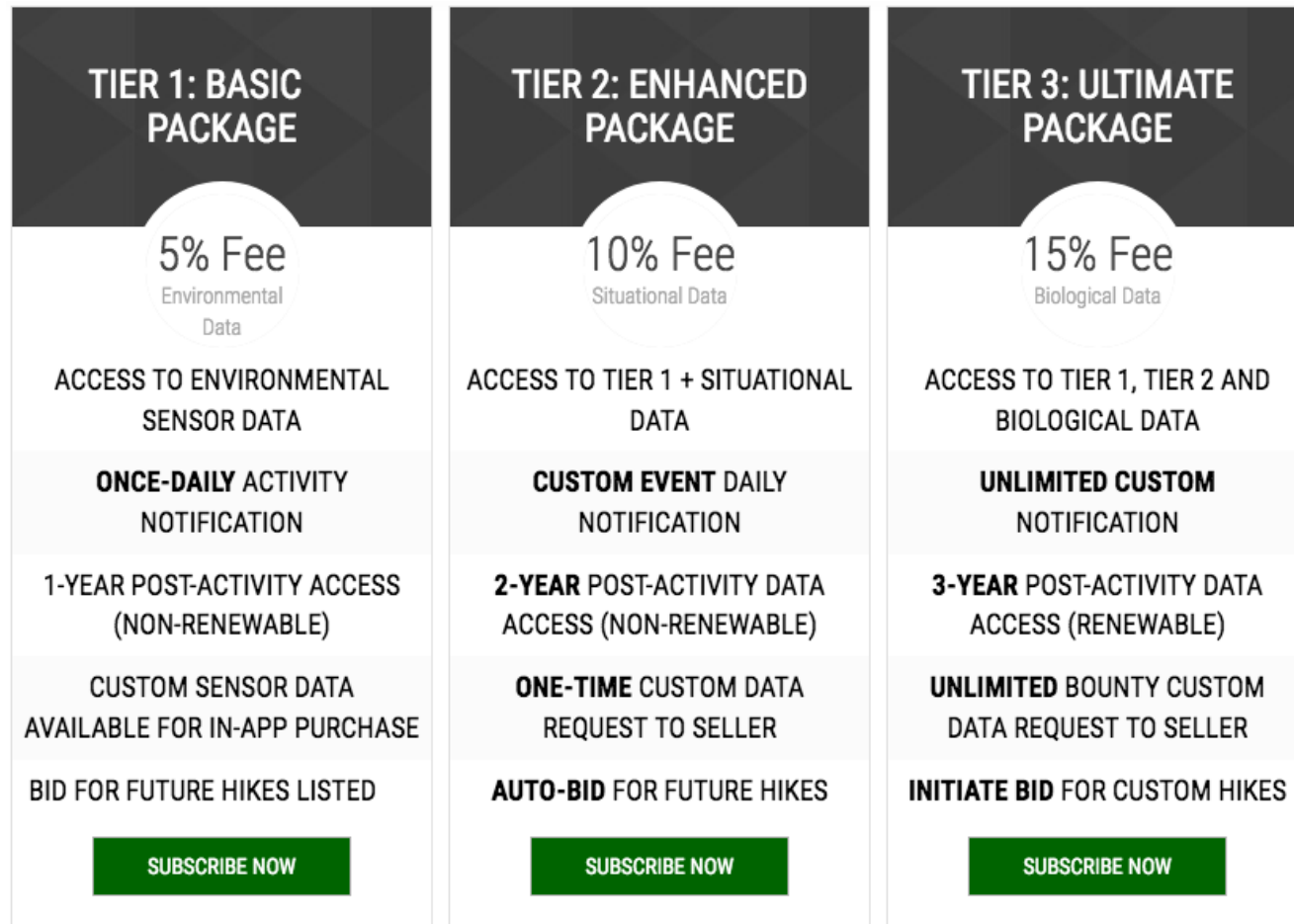
Figure 3: Data pricing and packaging: data tiers presented to willing buyers providing them with increased access to granular data. WBWS charges a nominal fee to defray the cost of accessing, hosting, analyzing and distributing the data from the Willing seller(s) to the buyer(s)

buyers to also opt-in to preferred data updates not available in their current data subscription tier.

Use Case: Long Distance Hikers

Our subscription model allows a range of use. We provide an example of how hikers (both as buyers and sellers) would benefit from our service. Long-distance hikers (thru-hikers who hike a trail end-to-end, and section hikers who spend up to months hiking a section of a long trail) experience unique events while "away" from typical day-to-day life. For instance, hiking a long trail such as the Appalachian Trail is an arduous task that can take a minimum of three months to complete [9]. Appalachian trail hikers do it for various reasons: for a sense of adventure, to conquer self [1] and to collect data [4]. The motivation to hike a trail often has a flow-on effect on how a hiker experiences the trail [22]. Figure 4 is an example of a typical long-distance hiker showcasing the types and placement of sensor devices.

A second consideration around hikers is the case of representation. For instance, there is a need to understand how people with disabilities experience the trail [23]. The under-representation of these populations creates a demand that could be met if there is an incentive model to encourage willing hikers to undertake the hike to meet the need for data to understand these experiences.

A third consideration is doing work. Technology is not prevalent in the wilderness and is an opportunity for hikers to gather data or test some device, individually or in groups. In this case, data collected or generated by hikers could be used to understand remote terrain, how hikers perform health wise in various locations, or testing equipment and gear [21]. The context of the both the hiker and the environment along with the experience makes for valuable insight that could be packaged and sold to the highest bidder.

Figure 4: The Willing Seller: Figure shows some placement of sensors on the hiker's body which borrows from how hikers currently place and use technology . The red circles show current technologies placed on the hiker, and the green circles show some opportunities to have additional sensors. Including sponsored sensors.

Preparation (Pre-Hike State of the Field)
Preparing for a thru-hike is a long-term commitment: future hikers have to make arrangements to take vacation from work, arrange affairs of home while on the trail, purchase trail gear which tend to be expensive, and arrange logistics that are made more complicated with the fact that the future thru-hikers are preparing for the unknown. Regardless of motivation for undertaking the hike, most concerns are valid across hikers of different types [1, 2, 4].

Part of planning for eventualities include discussing trail conditions and seeking advice from experienced and/or previous hikers. This is made easier especially on established trails like the Appalachian Trail. Previous, current and future thru-hikers have a central place such as online communities [1] devoted to the trail, where they can discuss the state of the trail, pose and answer (hypothetical or not) questions about their trail experience.

Considering different hikers, different data is sought: Miles per day, health maintenance (for people with chronic illnesses and those who take daily medication for instance), safety (and how to maintain it), calories taken, weight lost and/or gained through the hike, longevity of gear, trail conditions, weather conditions, water sources, re-supply points etc. These data metrics are important to hikers for their personal reasons and are rich with context and meaning to provide a full picture of what the hiker has gone through. This very context is the data gold mine for select buyers especially if you are a well known seller.

There is a set of hikers (mostly those with name recognition brought about by social media) that undertake sponsored hikes. The sponsorships range from an exchange of free gear for an honest review, to a fully sponsored hike to col-

---

[1]https://www.reddit.com/r/AppalachianTrail/

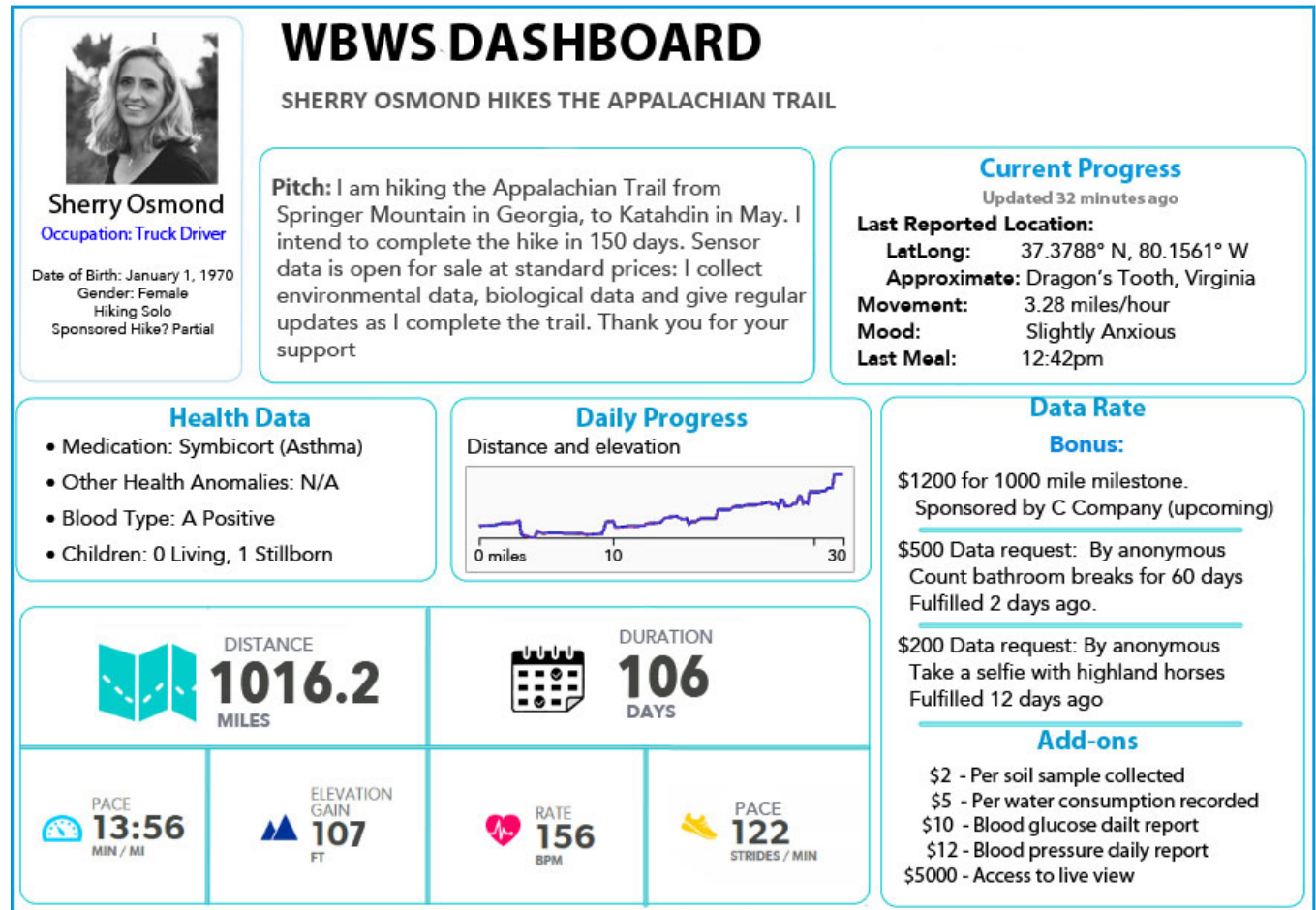**Figure 5:** Data management: The willing Seller provides data that is up for sale including biographical sketch that may influence bidding (medical companies may want to bid to track medical device performance). This dashboard provides a brief visual that allows the willing buyer to consider subscribing to a tier that fits their need, to submit requests for add-ons, and/or provide a sponsorship.

lect data and provide updates via podcasts. Agreements are typically set at the beginning of the hike, and deliverables planned. Our service democratizes the sponsorship process, allowing for a larger population to be able to participate as sponsors to hikers and/or trails that they are interested in.

On-Hike (How Hikers Currently Update)
Long distance hikers typically update their social media letting their followers know of their progress, to solicit help (typically morale-related, or to ask for rides to resupply towns). The communication medium is chosen by the hiker to fit their persona, ease of use, and privacy. Considering those with updates set to public, those on Instagram typically post multiple times a week, while those on YouTube do so once are week or longer. Frequency of posting typically relies on the availability of connectivity and the pre-processing time, so pictures with brief descriptions are faster to output than videos. Some hikers do both. Our service eases the process of updating by providing connectivity even while on the trail (Figure 4). Too, our dashboard provides a central place where updates are automatically incorporated from various social media and subsequently alerting the buyer about new information depending on the notification triggers (Figure 2) they set.

Post-Hike ("Patience, Young Grasshopper")
Post-hike, long distance hikers become senior members of the hiking community with institutional knowledge about the trail. Their role in the community evolves to pass on their knowledge and wisdom about other hikers they met on the trail, together with their trail experience to others in the community and also potential hikers. The experience gained and elevation gives the hiker a chance to partake in bigger selling opportunity where another hike could be performed or taking a group of hikers onto the trail (allow-

ing more data to be collected in a co-located context). In essence, the cycle is completed. This is typically where the hikers would share other data and information they experienced on the trail that was either presented from data posts or not (either due to privacy concerns, the processing involved to showcase the data, or lack of connectivity). The difference in the data provided is affected by time and context: while the accuracy of recall is degraded by the passage of time, the richness of information is increased as a factor of nuance and context.

Our service allows for ease-of-use whenever the hiker is ready to share the data. While our subscription just-in-time service ensures data integrity while the hiker is on the trail, we are nimble enough to allow for data sharing even after the hike is completed. A buyer can evolve to become a seller, or can be a buyer to a different subscription, while being a seller of their own data, allowing for control over personal information.

As we have shown, data control is the future. We offer incentives to our early adopter of our service. For potential sellers, for a brief-period, we offer *milestone bounties* for your activity. For buyers, we offer 2% off of the fees from each tier as our thanks for being an early adopter.

Sign up today!

Author's Statement
This design fiction probes the notion of what happens when personal data is formally turned into a commodity, instead of the indirect way it is currently done, where the users trade their activity data for access to "free services" such as the use of social media accounts. What happens when users control the data collected and what is shared in return for cash? Will over-sharing in exchange for hard currency become the norm, or will it lead users to choose to keep

their information private? We use the "Willing Buyer, Willing Seller" principle to provide transparency for all parties involved by using the activity of hiking as an example. Our use of design fiction also allows us to critically explore the ethics [13] of the future of personal data.

Personal Control
Part of the frustration with the current state in the field in personal information is the lack of control and the lack of knowledge of what exact information is collected, sold [18] or used to re-target the user [6]. The terms of use that users agree with as a condition for using an application are known to use difficult language and are geared towards protecting the providers rather than the user. They also tend to be modified with/without additional consent requested from the user. We consider the move towards users taking control over their data to be inevitable and ponder what it would mean to have the control over what is shared and what is sold.

Extensibility
The use of the hiker term *"The Trail Provides"* allows us to consider a case where money can used as an incentive to make decisions, such as to undertake (un)planned hikes depending on the amount of money available for incentives. As noted earlier in the paper, we considered the case of incentivizing atypical hikers [5] to undertake hikes, or experienced/frequent hikers to hike a specific or unmarked trail. This model can be extended as a dare, allowing users to share information of dare-devils act that is transmitted in real-time opens up a Pandora's box of how the dare-devil culture can be used to incentivize dangerous or illegal tasks.

Gains and Losses: Data for Sale vs Privacy and Ethics
With the "Willing buyer, willing seller" approach, the buyer is provided a just, market-value compensation for their data.

Additionally, they will receive compensation for the life of the data which can also be extended to provide earnings to the heir(s) – as is currently the case with literature and music publishing. This allows the user to obtain benefits that is otherwise kept from them, given that a lot of data collected from the users by current services are never destroyed, and can be re-sold as long as demand persists. We tackle the issue of data manageability of unwieldy data when discussing our dashboard design, which not only allows for a one-stop collection of data, but also provide visualization choices, and therefore context of how the data collected is interpreted.

The use of subscription data also raises an ethical question about ownership, privacy management and consent [20]. Even though the seller is the owner of data and controls what is sold, and for how much; for having paid for the use of the data, do the buyers have a license to the data? What if the seller exercises their right to be forgotten [16] by choosing to no longer make their data available online, what happens to data justly paid for?

Discussion
The *"Willing buyer, willing seller"* principle allows us to first consider the future of data control, and second to consider the ethics of data as a transaction item through the lens of the hiking context. It allows for the discussion on the price of privacy, whether this can be quantified, and the ethics of ownership (For example, do parents own (and therefore receive revenue from) their children's data? Even if the children attain majority?). It also borrows the central thesis from the current state-of-the-art in how companies obtain compensation for selling user data. It is our hope that our use of this principle will provide for an interesting perspective to discuss the ethics of data ownership and the future of data privacy.

REFERENCES

[1] Adam Berg. 2015. "To Conquer Myself": The New Strenuosity and the Emergence of "Thru-hiking" on the Appalachian Trail in the 1970s. *Journal of Sport History* 42, 1 (2015), 1–19.

[2] Karen Berger. 2001. *Hiking the Triple Crown* (1st. ed.). Mountaineers Books, Seattle, WA.

[3] Alyson Browett. 2018. MyATStory: Voices from the Mountain. (2018). Retrieved May 30, 2019 from `http://appalachiantrail.org/myatstory/voices-from-the-mountains`

[4] Bill Bryson. 1999. *A Walk in the Woods: Rediscovering America on the Appalachian Trail* (1st. ed.). Random House, New York, NY.

[5] Merry J. Coburn. 2006. *Walking Home: Women's Transformative Experiences in the Wilderness of the Appalachian Trail*. Ph.D. Dissertation. Institute of Transpersonal Psychology, Palo Alto, California. UMI Order Number: 3221761.

[6] Claire Dolin, Ben Weinshel, Shawn Shan, Chang Min Hahn, Euirim Choi, Michelle L. Mazurek, and Blase Ur. 2018. Unpacking Perceptions of Data-Driven Inferences Underlying Online Targeting and Personalization. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 493, 12 pages. DOI: `http://dx.doi.org/10.1145/3173574.3174067`

[7] Casey Fiesler and Blake Hallinan. 2018. "We Are the Product": Public Reactions to Online Data Sharing and Privacy Controversies in the Media. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 53, 13 pages. DOI: `http://dx.doi.org/10.1145/3173574.3173627`

[8] Mohammad Taha Khan, Maria Hyun, Chris Kanich, and Blase Ur. 2018. Forgotten But Not Gone: Identifying the Need for Longitudinal Data Management in Cloud Storage. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 543, 12 pages. DOI: `http://dx.doi.org/10.1145/3173574.3174117`

[9] Lindah Kotut, Michael Horning, Derek Haqq, Shuo Niu, Timothy L. Stelter, and D. Scott McCrickard. 2018b. Tensions on Trails: Understanding Differences between Group and Community Needs in Outdoor Settings. In *Rural Computing Workshop at CSCW 2018*.

[10] Lindah Kotut, Michael Horning, Steve Harrison, and D Scott McCrickard. 2018c. Opportunity in Conflict: Understanding Tension Among Key Groups on the Trail. In *HCI Outdoors Workshop at ACM CHI 2018*.

[11] Lindah Kotut, Michael Horning, and Scott McCrickard. 2018a. Who's on the Trail: Identifying Trail Uses with Affinity Diagrams. In *Technology on the Trail Workshop at ACM GROUP 2018*.

[12] Yuwei Lin. 2018. #DeleteFacebook is still feeding the beast – but there are ways to overcome surveillance capitalism. (March 2018). Retrieved May 30, 2019 from `https://bit.ly/2NSAxaS`

[13] Joseph Lindley and Dhruv Sharma. 2016. Operationalising Design Fiction for Ethical Computing. *SIGCAS Comput. Soc.* 45, 3 (Jan. 2016), 79–83. DOI: `http://dx.doi.org/10.1145/2874239.2874251`

[14] Leslie S. Liu, Jina Huh, Tina Neogi, Kori Inkpen, and Wanda Pratt. 2013. Health Vlogger-viewer Interaction in Chronic Illness Management. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 49–58. DOI:http://dx.doi.org/10.1145/2470654.2470663

[15] James Morris. 2019. DuckDuckGo: The search engine taking on Google and making the internet 'less creepy' with its privacy mission. (May 2019). Retrieved May 30, 2019 from https://bit.ly/2PZGC7W

[16] Jeffrey Rosen. 2011. The right to be forgotten. *Stan. L. Rev. Online* 64 (2011), 88.

[17] Shruti Sannon, Natalya N. Bazarova, and Dan Cosley. 2018. Privacy Lies: Understanding How, When, and Why People Lie to Protect Their Privacy in Multiple Online Contexts. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 52, 13 pages. DOI:http://dx.doi.org/10.1145/3173574.3173626

[18] Milivoj Simeonovski, Giancarlo Pellegrino, Christian Rossow, and Michael Backes. 2017. Who Controls the Internet?: Analyzing Global Threats Using Property Graph Traversals. In *Proceedings of the 26th International Conference on World Wide Web (WWW '17)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 647–656. DOI:http://dx.doi.org/10.1145/3038912.3052587

[19] Sharon Slade, Paul Prinsloo, and Mohammad Khalil. 2019. Learning Analytics at the Intersections of Student Trust, Disclosure and Benefit. In *Proceedings of the 9th International Conference on Learning Analytics & Knowledge (LAK19)*. ACM, New York, NY, USA, 235–244. DOI:http://dx.doi.org/10.1145/3303772.3303796

[20] Daniel J Solove. 2012. Introduction: Privacy self-management and the consent dilemma. *Harv. L. Rev.* 126 (2012), 1880.

[21] Timothy Stelter and D. Scott McCrickard. 2017. Hiking the Appalachian Trail with Technology. *NatureCHI 2017 Workshop at MobileHCI 2017* (2017). http://people.cs.vt.edu/~tstelter/docs/hiking-appalachian-trail

[22] Cheryl Strayed. 2012. *Wild: From lost to found on the Pacific Crest Trail*. Vintage.

[23] Sarah Trankle. 2013. The Blind Hiker: How one man used technology to conquer the 2,000-mile Appalachian Trail. (Jan. 2013). Retrieved March 30, 2019 from hhttps://bit.ly/2TNHRFH