Atlas of Caregiving Pilot

The Process of Developing 237 Diagrams

—

Dubberly Design Office
Version 3.0
April 18, 2016

Table of Contents

Introduction

Implications of IoT for Family Care Giving

A new wave of information technology is barreling towards us—the Internet of Things (IoT). Experts predict this wave will create more change and deeper change than the original commercialization of the Internet and the PC revolution.

At heart, IoT is about adding sensors and processors to everything—creating smart devices, which collect, store, and analyze data, so that they can act "intelligently."

IoT is also about connecting those smart devices to each other and to the Internet—especially to cloud-based applications and services, which aggregate data and "mine" it, looking for relationships and patterns. Smart, connected devices and their supporting cloud-based apps create product systems. These systems can also connect with other systems to form vast product-service ecologies, making the world aware of itself, at least in a sense.

The IoT revolution is underway across many sectors of industry, and it's now moving into consumer products and the home. And into healthcare.

The Study

The Robert Wood Johnson Foundation Family Care Giving (RWJF FCG) study is premised on the idea that sensors can aid in understanding care giving, particularly in locating sources of stress and perhaps measuring amount of stress. The technology for measuring a person's stress level is well advanced. Several examples have been commercialized and are entering the market. These technologies are proven in lab settings and even in closely controlled field studies, not just in academic projects but also in commercial marketing research projects for companies such as Lego and Lowes.

What hasn't been proven yet—and a key part of our study—is the use of sensors in homes over an extended period of several hours.

The main goal of the study was to understand if and how stress sensors might supplement traditional ethnography. In addition to stress sensors, we employed a number of other sensors in the study.

The complete study report can be found at http://atlasofcaregiving.com/wp-content/uploads/2016/03/Study_Report.pdf

The Diagrams

Another premise of the study is that we do not have a clear picture of "Family caregiving" and the large amount of work involved. We need to map the territory. That thought led to the idea of creating an "Atlas of Caregiving."

A major component of the final report—and a step towards the "Atlas"—is the 237 diagrams we created for the report. They are contained in an addendum to the report titled Data Visualizations, available here http://atlasofcaregiving.com/wp-content/uploads/2016/03/Data_Visualizations.pdf

Related Publications

Diagram Development

We created a process document which details the development of the 237 diagrams. It covers everything from recruiting and interviewing, to data downsampling and filtering, available here http://staging.dubberly.com/atlas/160504-Final_Documents/AoC_Development_160505a.pdf

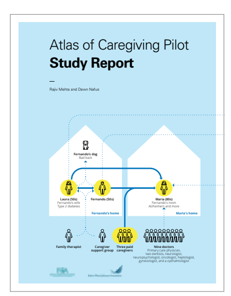
Care Network Diagram Rationale

We created a design rationale document which details the design process for the care network diagram, available here http://staging.dubberly.com/atlas/160504-Final_Documents/AoC_Care_Network_Diagram_Style_160505a.pdf

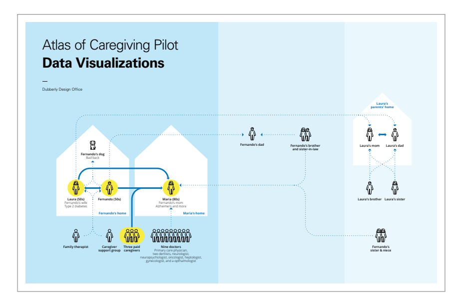
Website

The project is also available online. All of the diagrams are available in SVG format and can be enlarged to show more detail. You can find the website here http://atlasofcaregiving.com

SmartThings Motion Sensor



Study Report



Data Visualizations

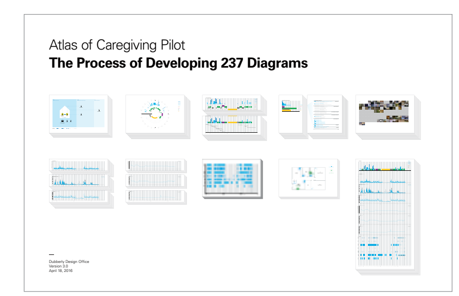
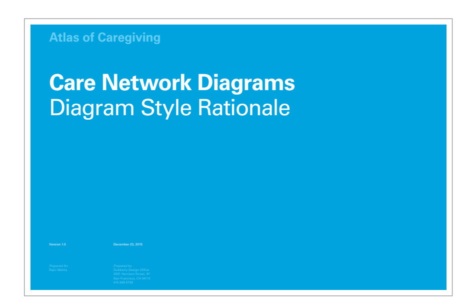
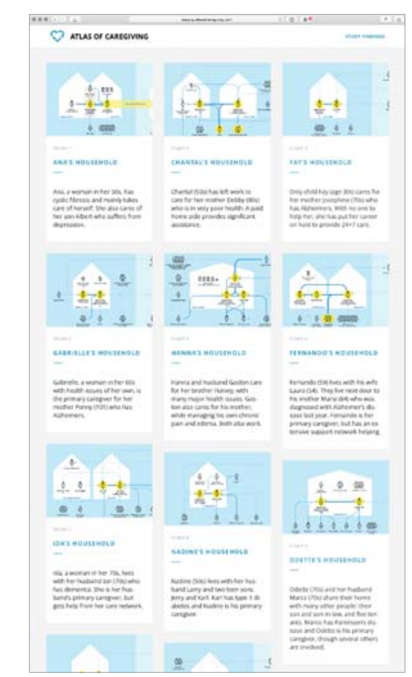


Diagram Development



Care Network Diagram Rationale



Website Studies Page

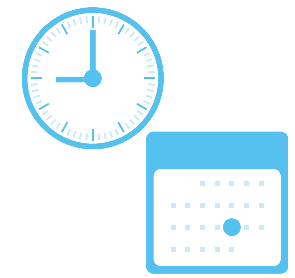
Overview of the Project

14 households

× 2 environmental sensors

CO2
Sound
Humidity

measured 3 factors

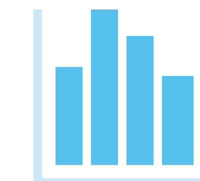


over 36 hours (over 6 months)

Gigabytes

yielding 5 GB of data

or nearly half a billion data points



turned into 237 diagrams

with 20 participants

× 2 wearable sensors

Acceleration X, Y, Z
(Average Motion*)
Blood Volume Pulse
(Heart Rate*)
Electro Dermal Activity

measured 5 factors

* calculated 2 derivative factors

with 21 chronic illnesses

× interviewing + observation

× self reporting

Overview of Participants

Fourteen households were recruited for the study.
Three of the households had two participants.
One of the households had three participants.

There were twenty participants in total.
(All names are fictitious to protect participant privacy.)

Ana

Study 1 – Ana's Household
Ana, a woman in her 50s, has cystic fibrosis and mainly takes care of herself. She also cares of her son Albert who suffers from depression.

Chantal

Study 2 – Chantal's Household
Chantal (50s) has left work to care for her mother Debby (80s) who suffers from dementia, hip and back pain, allergic pneumonia, congestive heart failure, type-2 diabetes, and kidney and thyroid pain.

Fay

Study 3 – Fay's Household
Only-child Fay (30s) cares for her mother Josephine (70s) who has been diagnosed with Alzheimer's disease. With no one to help her, she has put her career on hold to provide 24×7 care.

Gabrielle

Study 4 – Gabrielle's Household
Gabrielle, a woman in her 60s with health issues of her own, is the primary caregiver for her mother Penny (101) who has Alzheimer's.

Hanna Gaston Harvey

Study 5 – Hanna's Household
Hanna (50s) and husband Gaston (50s) care for her brother Harvey (50s), whom has epilepsy, cognitive and motor decline, and is prone to pneumonia/sepsis. Gaston also cares for his mother, while managing his own chronic pain and edema. Both also work.

Fernando Laura

Study 6 – Fernando's Household
Fernando (50s) lives with his wife Laura (50s). They live next door to his mother Maria (80s) who was diagnosed with Alzheimer's disease last year. Fernando is her primary caregiver, but has an extensive support network helping.

Ida

Study 7 – Ida's Household
Ida, a woman in her 70s, lives with her husband Ian (70s) who has dementia. She is her husband's primary caregiver, but gets help from her care network.

Nadine

Study 8 – Nadine's Household
Nadine (50s) lives with her husband Larry and two teen sons Jerry and Karl. Karl has type-1 diabetes and Nadine is his primary caregiver.

Odette

Study 9 – Odette's Household
Odette (70s) and her husband Marco (70s) share their home with many other people: their son and son-in-law, and five tenants. Marco has Parkinson's disease and Odette is his primary caregiver, though several others are involved.

Nate Patty

Study 10 – Nate's Household
Nate (30s) lives with his wife Patty (30s). Nate has a brain tumor and Patty has MS. They are each others primary caregivers.

Sally

Study 11 – Sally's Household
Sally (50s), a former lawyer, lives with her son Pablo (20s). 3 months into the pregnancy with Pablo, Sally suffered a ruptured amnio,which led to the discovery that Pablo had XYY chromosome disorder. Sally provides full time care for Pablo.

Tammy Rafael

Study 12 – Tammy's Household
Tammy's home includes her husband Rafael, and their two pre-teen children Wanda and Sam. Wanda has severe epilepsy; cerebral palsy; and is effectively quadriplegic. Sam has severe autism and requires constant supervision.

Teddy

Study 13 – Teddy's Household
Teddy (40s) lives with his wife and two sons, Walter and Van. Van has Aspergers, and his parents function as his primary caregivers.

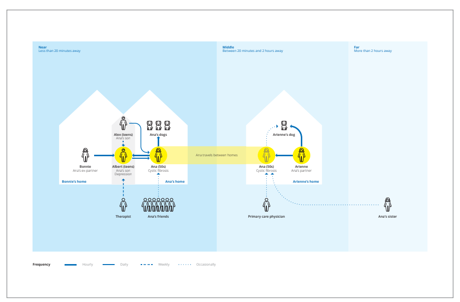
Omar Cindy

Study 14 – Omar's Household
Omar (40s) and his separated wife Cindy (40s) share a home with their son Bob (pre-teen). Bob has Aspergers and his parents are his primary caregivers.

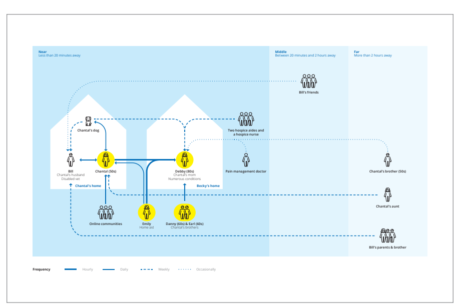
Overview of Care Networks

Care network diagrams were created for each household. Care network diagrams show all of the people involved in receiving and or giving care including details such as their name, age, illness, and their relationship to each other. The diagrams group people by distance, as indicated by the blue backgrounds. Different line styles indicate the frequency of care.

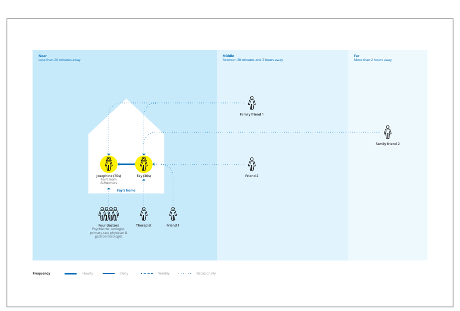
Study 1 – Ana's Household



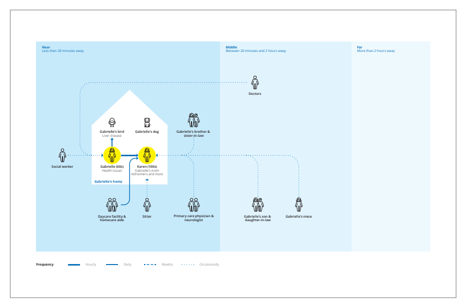
Study 2 – Chantal's Household



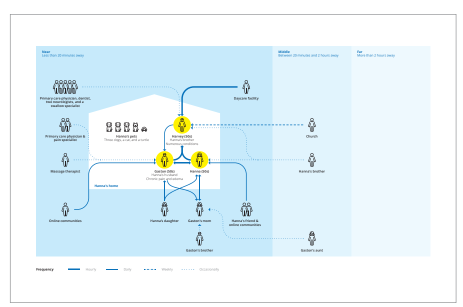
Study 3 – Fay's Household



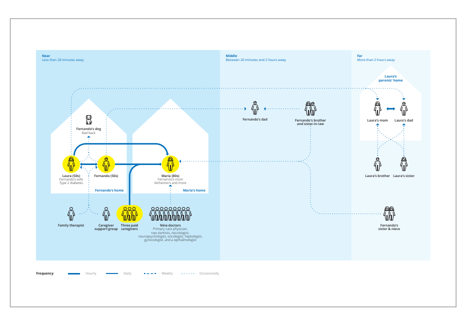
Study 4 – Gabrielle's Household



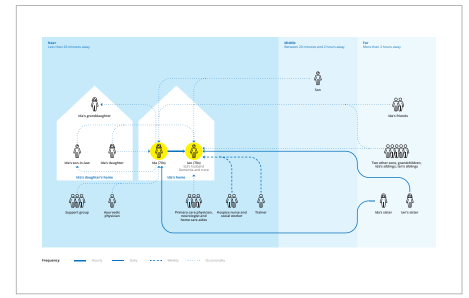
Study 5 – Hanna's Household



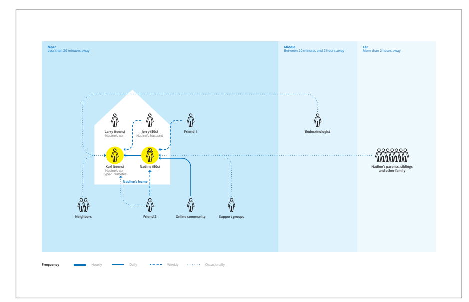
Study 6 – Fernando's Household



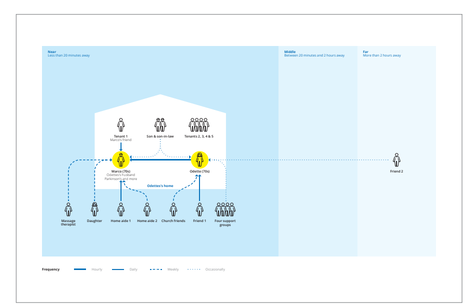
Study 7 – Ida's Household



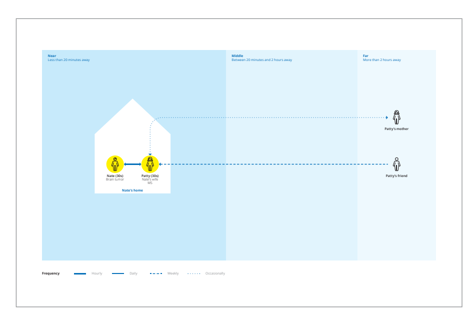
Study 8 – Nadine's Household



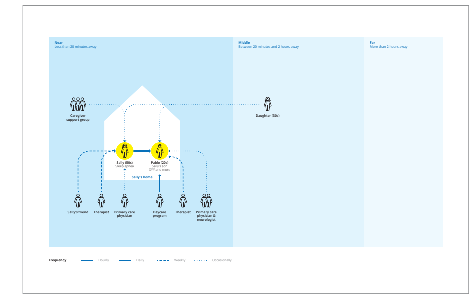
Study 9 – Odette's Household



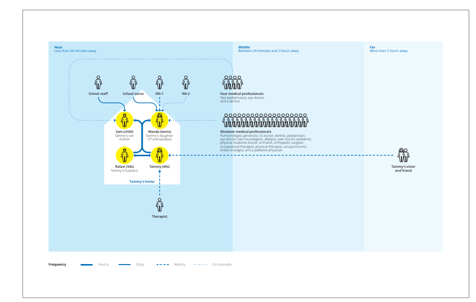
Study 10 – Nate's Household



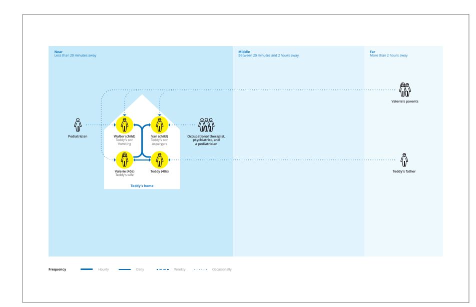
Study 11 – Sally's Household



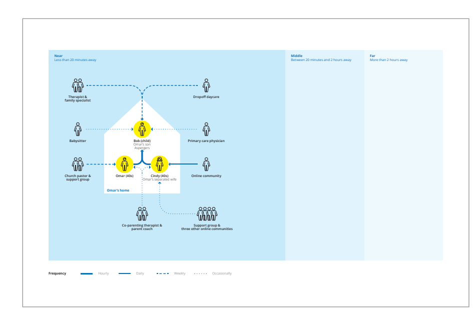
Study 12 – Tammy's Household



Study 13 – Teddy's Household



Study 14 – Omar's Household



Overview of Floorplans

Floorplans were created for household's which had environment and motion sensors placed inside. (You can see that not all participant's agreed to deploy in-home sensors, or data was not collected due to technical difficulties.) Blue icons indicate the presence of a SmartThings motion sensor and its direction. Black text labels the location in the house and number of the sensors. Green icons indicate Netatmo weather stations. Green text clarifies the outdoor/primary unit from the indoor/secondary unit.

Study 1 – Ana's Household

NO DATA COLLECTED

Study 2 – Chantal's Household

Study 3 – Fay's Household

Study 4 – Gabrielle's Household

Study 5 – Hanna's Household

NO DATA COLLECTED

Study 6 – Fernando's Household

NO DATA COLLECTED

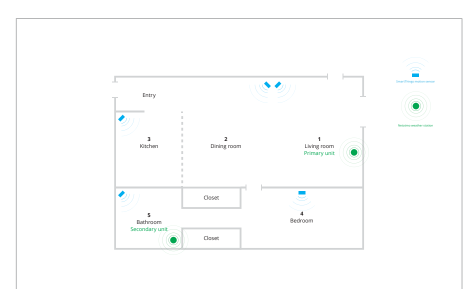
Study 7 – Ida's Household

Study 8 – Nadine's Household

NO DATA COLLECTED

Study 9 – Odette's Household

Study 10 – Nate's Household



Study 11 – Sally's Household

NO DATA COLLECTED

Study 12 – Tammy's Household

Study 13 – Teddy's Household

NO DATA COLLECTED

Study 14 – Omar's Household

NO DATA COLLECTED

Overview of Methods

Four methods were used to collect data:

1. Interviewing + Observation
- Online questionnaires
- In-person interviews

2. Self Reporting
- Activity log

3. Wearable Sensors
- Narrative Clip
- Empatica E4
- SmartSense Presence

4. Environmental Sensors
- SmartSense Motion Sensors
- SmartThings Hub
- Netatmo Weather Stations

Audio Visual Sensors
Sound and video recording were also considered but ultimately rejected because the Atlas of Caregiving core team and advisors felt it would have been too invasive.

Look for these icons throughout this document to indicate the method type.

Overview of Results

At right we show a complete set of 10 diagrams for a single study participant (Fay).

Some diagrams are made of multiple data plots. Below is a tally of each plot for every participant.

14 Care Network
18 24-hour Log
36 36-hour Log
36 Activities
20 Photo Log
57 Body
24 Environment
7 Motion
7 Floorplan
18 Summary

237 total diagrams

Care Network × 14 participants

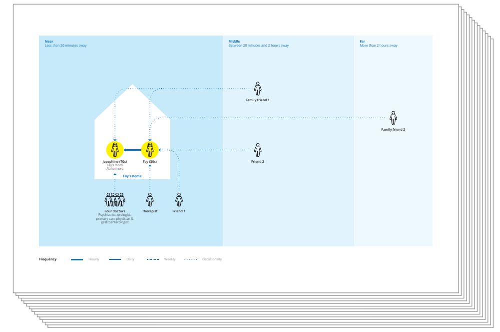
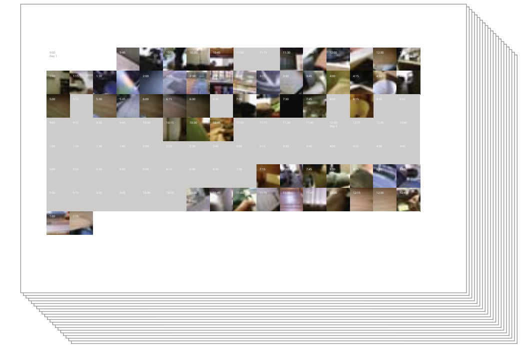
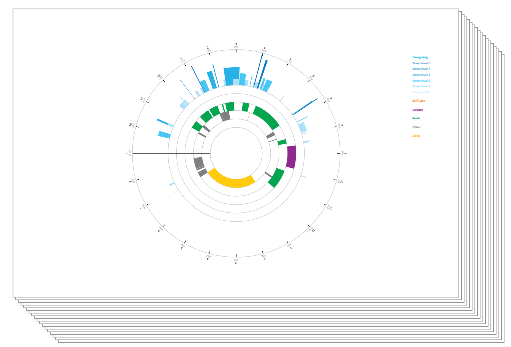


Photo Log × 20 participants

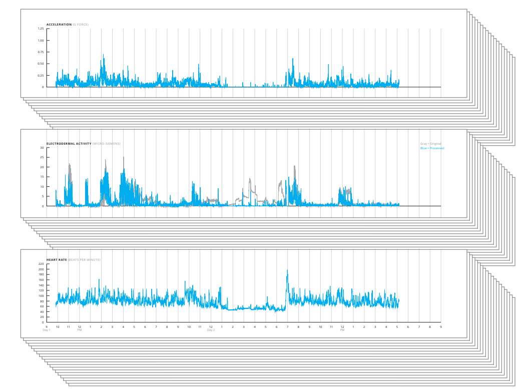


Floorplan × 7 participants

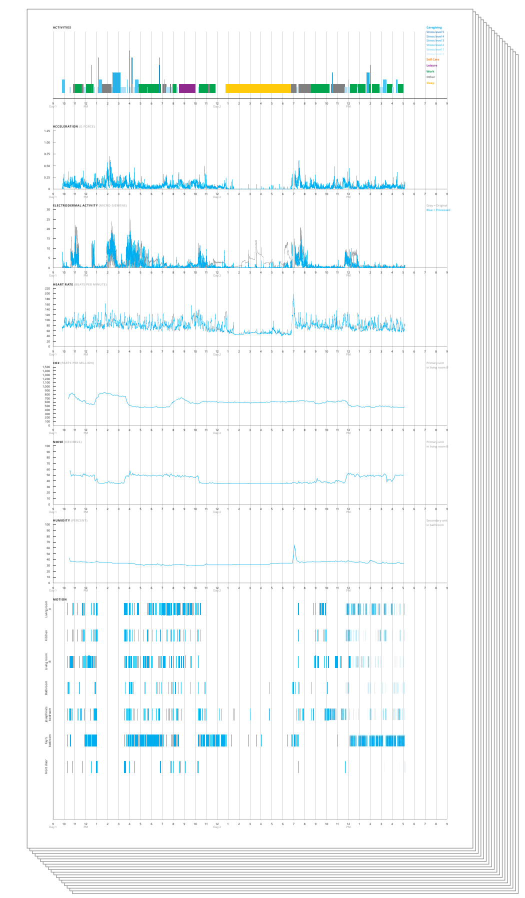
24-hour Log × 18 participants



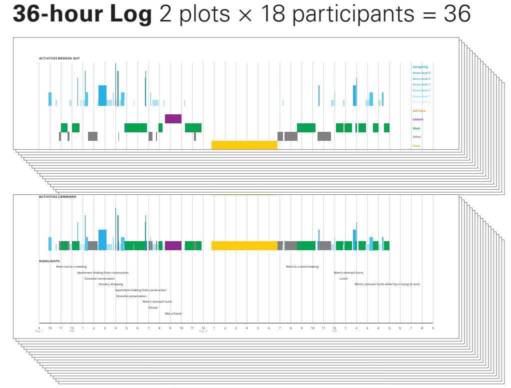
Body 3 plots × 19 participants = 57



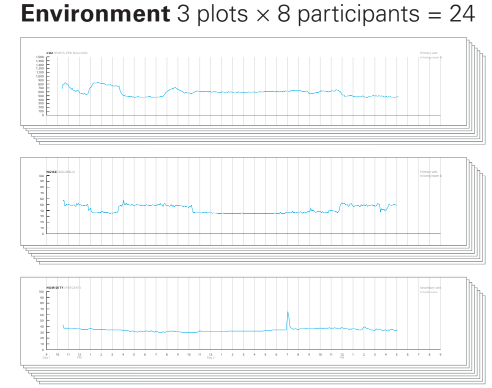
Summary × 18 participants



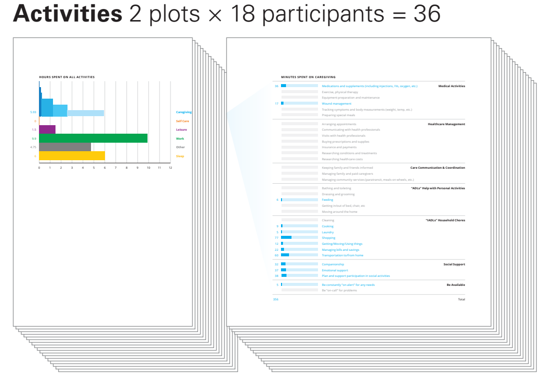
36-hour Log 2 plots × 18 participants = 36



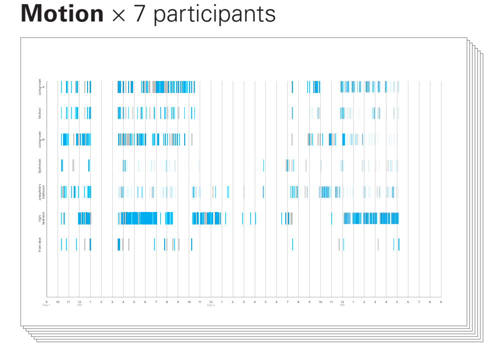
Environment 3 plots × 8 participants = 24



Activities 2 plots × 18 participants = 36



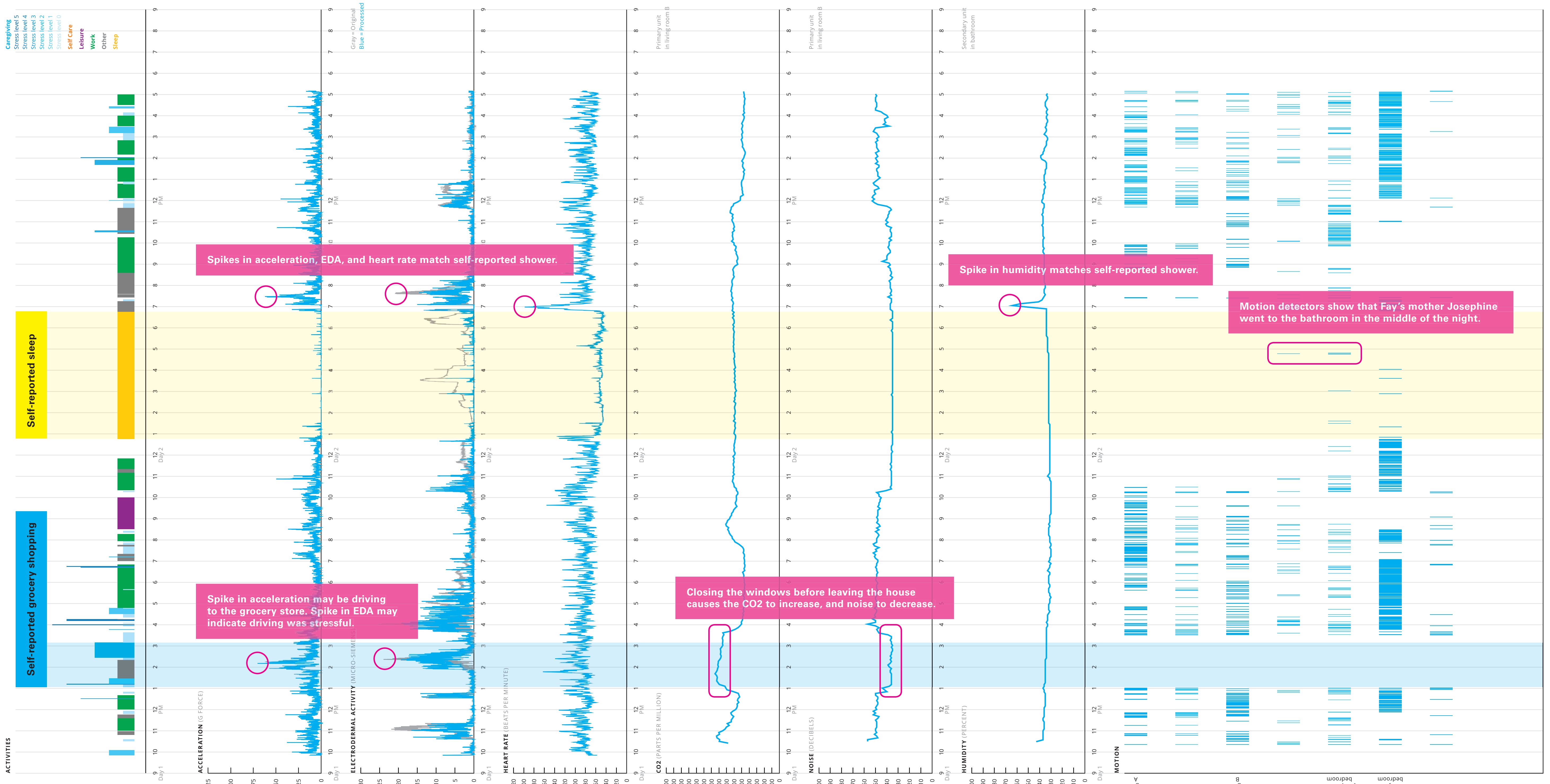
Motion × 7 participants



How to Read the Diagrams

Many interesting insights are present in the diagrams, but they might not immediately jump out at you. Below we present the summary diagrams for Fay, and we've added overlays which illustrate our reading of the data and provide possible insights.

The yellow strip represents sleep time as reported by the participant. The blue strip represents a trip to the grocery store, also as reported by the participant. The magenta circles and descriptive text highlight points of interest.



Part 1

Planning the Study

Recruiting Participants

Participants were recruited through health organizations, affinity groups, friends, and associates, by the study's principle investigators.

The intention was to create a qualitative study—sampling a range of different care situations with people of different ages and conditions.

Project Description

ATLAS OF CAREGIVING

BUILDING AN ATLAS OF CAREGIVING

We know that family caregiving can be rewarding. But we also know that it is hard work. It can be physically and emotionally exhausting. It can have financial, social and health consequences for people who do it. It can be 24/7. To date, traditional research efforts have only provided us a high level view of what it means to be a caregiver. We need to know more.

We're building an Atlas of Caregiving to understand the actual, lived experience of daily caregiving. Armed with this information, we can create new technologies, improve services and develop policies to make caregiving both easier and more effective.

A DAY IN THE LIFE

The Family Caregiver Alliance is undertaking a novel research project, led by Rajiv Mehta of Bhageera, Inc. and Hugh Dubberly of Dubberly Design Office, to look deeply into a family's day and map the life of a caregiver.

Using both traditional ethnographic research approaches alongside new wearable technologies that can track activities and physiological signals such as heart rate and movement in real time, we will collect data from a diverse group of families in the San Francisco Bay Area.

We will analyze that data and seek to answer: Who is involved in family caregiving? What do they do? What are the impacts of these activities? This information will be gathered into an Atlas of Caregiving—a collection of stories, diagrams, data and models—that exposes the day-to-day realities of caregiving and provides rich insights into the challenges caregivers face.

A TIME TO ACT

Society is already highly dependent on family caregivers, as they do the bulk of the work of caring for people's health. As the population ages, as medical advances enable people to live much longer with serious illnesses, and as medical care moves from a hospital-based system into the home and community, a growing number of Americans will find themselves providing unprecedented levels and types of care. We need to better understand family caregiving if we are going to be ready to manage these demands.

Visit the project website for updates and emerging insights: www.atlasofcaregiving.com

Collaborators for the project include individuals from Caring Across Generations; Cincinnati Children's Hospital; Columbia University; Cystic Fibrosis Foundation; Dartmouth University; Health 2.0; Institute for the Future; Intel; Medrise; Quantified Self; ReACT; SmartPatients; and United Hospital Fund. The Atlas of Caregiving Pilot Project is supported by a grant from the Robert Wood Johnson Foundation. The Family Caregiver Alliance is a national nonprofit organization that provides direct caregiver support services, public policy development, research, and public awareness regarding family caregivers.

Robert Wood Johnson Foundation

FAMILY CAREGIVER ALLIANCE
National Center on Caregiving

ATLAS OF CAREGIVING

Atlas of Caregiving Pilot
Details about Caregiver Participation in Research

Who can participate in the research?
The project is seeking people actively involved (more than 2 hours per day) in the day-to-day care of a family member or friend. Participants may also be taking care of their own health needs.

By "care" we mean anything you do to help the other person due to their illness, disability or aging — medical activities, household chores, keeping company, etc. So, normal parenting doesn't count, but all the activities due to that child's chronic condition would count.

We are especially interested in situations where multiple family members are involved, and willing to participate. For example:

- A couple caring for each other — 2 participants.
- A couple caring for a parent — 2 (or 3) participants.

We are seeking a diverse group of families, with a variety of health conditions, social backgrounds, and family situations.

Families should be in the San Francisco / Silicon Valley area (generally the counties of Alameda, Contra Costa, San Francisco, San Mateo, and Santa Clara).

People do *not* need to be "tech-savvy" to participate. However, they need to be able to communicate clearly, in-person and via email. And, they need to be able to wear and take on/off a watch-like device.

What will I be asked to do?
Your participation will include a survey, two interviews (at your home), and tracking your care activities using a log and wearable technologies.

Survey: Initially, a survey will collect basic information about yourself (age, sex, etc.), your life (job, home, etc.), and about your care situation (the age, health, and relationship to the person you care for, etc.).

Interviews: This will be followed by your participation in an in-home study, which will take place over 3 days. A trained researcher will interview you at home, to learn more about yourself, your life and care situation. With your permission, the researcher will tape-record the interview, and take photos of your care environment. Following the tracking period (next paragraph), you will participate in a second interview during which you and the researchers will discuss your log and the data from the devices.

Tracking: You will be loaned some wearable technologies—a watch-like device and a small camera that can be clipped onto your shirt—that you will wear for 24-36 hours. (We will ask you to wear them as much as possible, but you are free to take them off when you need to.) During that same 24-36 hour period you will also keep a log of your care activities, using a special notebook. Other devices may be placed in your home to collect environmental data (humidity, sound level, etc.).

What data will be collected?
We are interested in learning: Who is involved in family caregiving? What do they do? What are the impacts of these activities?

Through the survey, interviews, and log we will learn:

- Basic demographic information (age, sex, etc.)
- Who in your family is taking care of whom, and what their health needs are.
- What specifically you did, what your care activities were.
- How much time these activities took.
- How physically demanding and emotionally stressful you felt these activities were.

Images from the small camera will be used to help you remember and discuss your activities during the second interview.

The watch-like device collects information about your heart rate, perspiration, movement, and skin temperature. These data will be analyzed to make rough numerical assessments on how physically demanding and emotionally stressful (or positive) your activities were.

How will this data be used?
We will analyze what we learn from you to create some diagrams to describe your care situation. These diagrams, along with a textual description, will be published on the project website, in project reports, and presented at conferences. Such published data will *not* include your name or other identifying information (see next section about privacy).

We expect that this information will help those seeking to develop products, services and policies to help caregivers. We also expect the news media will write stories about the project and what we have learned.

How will my privacy be protected?
Your personally identifiable data (name, address, email, photos, etc.) will only be known within the research team. Otherwise, your personally identifiable data will be kept confidential. Published information will use substitute names, and generalized locations.

If we think it will be very helpful to use personally identifiable data publicly, we will explicitly ask for your permission first.

We will also take reasonable and currently available efforts to keep your data secure on our computers. However, there is no way to guarantee that your data will be perfectly safe.

What harms may come from participating?
As you know, caring for yourself and your family can be hard. Talking about your efforts and seeing "proof" of how much time, physical and mental energy you spend may make you more aware of how hard caregiving can be, and so make you feel anxious or worried. If this happens, we will refer you to support resources in your area.

What benefits may come from participating?
In similar previous studies, participants found it interesting and helpful to learn about themselves. They saw something new, something not consciously noticed before, from the log, from the sensor data, and from the interview. Understanding what causes you stress or gives you pleasure may lead to finding appropriate help or ways to change the situation.

You will also be helping families everywhere as what is learned from this study will contribute to improved support for family caregiving.

Is there any compensation?
Yes, you will receive $200 for participating in the study.

What else should I know?
Your participation is voluntary, and you can stop at any time — either temporarily, by turning off or removing the wearable devices, or completely. When we pick up the devices, we will also ask if there is any time period for which you want us to *not* look at the data, for example, to protect your family member's privacy, and we will honor your request.

Who can I call with questions?
You can contact Rajiv Mehta, the project director for the Atlas of Caregiving Pilot, at (650) 823-3274 or rajiv@bhageera.com

Ethnography for Pilot

Summary of Instructions for Interviewers
We will record interviews (of family caregivers) using audio, using a smartphone or voice-recorder, and possibly one more device for backup.

Interviewers will also take written notes while interviewing.

As soon after the interview as possible, the interviewers will make detailed notes, including impressions, using the voice-recording and contemporaneous notes as supplement if necessary.

We will not do video-recording of the interviews.

We will also take photos of the home and family (assuming permission).

Overall Flow
Participant screening via phone/web/email
Consent & scheduling via phone/email

1st appointment (day 1) — 90-120 minutes
- Check-in r.e. consent—5 minutes
- interview for context & ecosystem — 60-90 minutes
- provide/set-up sensors/logs & training — 30 minutes

2nd appointment (day 2) — 15 minutes
- pick-up sensors/logs
- ask if anything we should know as we look at data (equipment problems, etc.)

3rd appointment (day 3) — 60 minutes
- interview for reflections on experience
- interview for data refinement & interpretation
- re-check r.e. consent—5 minutes

Later discussion of "final" analysis via phone.

1st Appointment Interview
Review 2-3 highlights of the consent form.

Understand overall context: life, health, family, work, etc., with questions like:
- Tell me about yourself and your family.
- How long have you been in this home/community?
- What are the health issues you're dealing with?
- Do you work? What do you like/dislike about your work?
- How do work and caregiving fit?
- What do you do for leisure?
- What would you like to do if you had more time?

Care ecosystem:
- Who do you care for?
  - What are their health issues?
  - How long have you been providing care?
- Continuous? Stop/start?
  - What exactly do you do as regards caregiving?
  - How does this impact you?
- Who else cares for them?
  - What do they do?
- Who cares for you?
  - What do they do?
- Look for indication of care work that the person doesn't think of as "care work" but still needs doing.
- What is involved with the orchestration of everything?
- Communication and coordination of everyone?

Explain/demo/train on logging and sensing.

Schedule equipment pick-up and 2nd interview.

2nd Appointment Interview
Ask for their reflection about the day in question
- How was the experience of logging and using the sensors?
- How was the day overall (similar to or different from other days)?
- Did they learn anything new?

Review collected data
- Review one stream of info at a time (e.g. written log, Narrative Clip, sensor data, etc.).
- Anything we've learned that you'd rather we didn't use?

On making sense of the sensor data
Once we have a good data from a few families, we'll want to discuss our findings with people expert in analyzing the kinds of data collected by the Fitbit/Basis/Empatica devices (e.g. HR, EDA, accelerometer, etc.).

We received valuable advice and help from John Cain, VP Market Analytics, Sapient-Nitro and his staff at Iota, especially Peter Binggeser, author of Datadeer, and Aniket Bhatnagar and Pasindu Banda Wewegama of Sapient India. John and the Iota team are experts in using sensors to support ethnography.

We also received valuable counsel from Elliott Hedman, founder of mPath, a research consultancy focused on understanding user emotions. Elliott recently received his PhD from MIT and is a world expert in collecting and assessing EDA data.

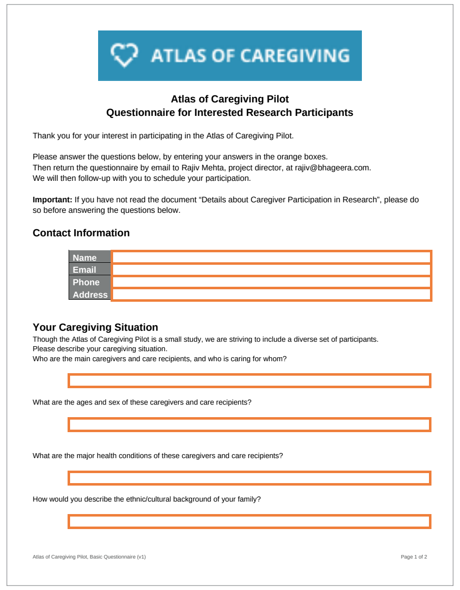
Patch Kessler, a UC Berkley PhD, and expert in using MatLab, provided valuable advice on data filtering.

Part 2

Collecting Data

Interviewing + Observation

Online Questionnaire



ATLAS OF CAREGIVING

Atlas of Caregiving Pilot
Questionnaire for Interested Research Participants

Thank you for your interest in participating in the Atlas of Caregiving Pilot.

Please answer the questions below, by entering your answers in the orange boxes.
Then return the questionnaire by email to Rajiv Mehta, project director, at rajiv@bhageera.com.
We will then follow-up with you to schedule your participation.

Important: If you have not read the document "Details about Caregiver Participation in Research", please do so before answering the questions below.

Contact Information

Name	
Email	
Phone	
Address	

Your Caregiving Situation
Though the Atlas of Caregiving Pilot is a small study, we are striving to include a diverse set of participants.
Please describe your caregiving situation.
Who are the main caregivers and care recipients, and who is caring for whom?

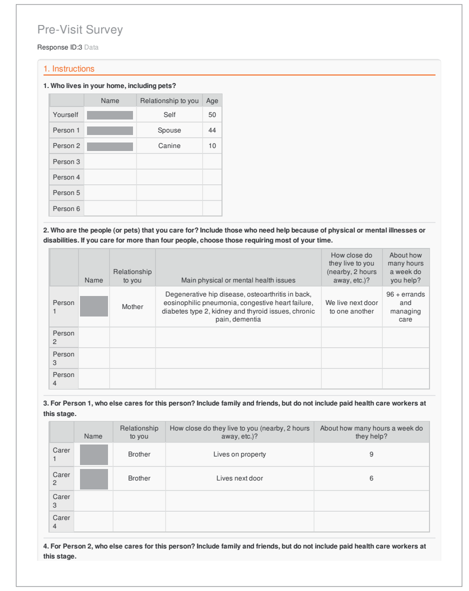
What are the ages and sex of these caregivers and care recipients?

What are the major health conditions of these caregivers and care recipients?

How would you describe the ethnic/cultural background of your family?

Atlas of Caregiving Pilot, Basic Questionnaire (v1) Page 1 of 2

Pre-Visit Survey



Pre-Visit Survey

Response ID:3 Data

1. Instructions

1. Who lives in your home, including pets?

	Name	Relationship to you	Age
Yourself		Self	50
Person 1		Spouse	44
Person 2		Canine	10
Person 3			
Person 4			
Person 5			
Person 6			

2. Who are the people (or pets) that you care for? Include those who need help because of physical or mental illnesses or disabilities. If you care for more than four people, choose those requiring most of your time.

	Name	Relationship to you	Main physical or mental health issues	How close do they live to you (nearby, 2 hours away, etc.)?	About how many hours a week do you help?
Person 1		Mother	Degenerative hip disease, osteoarthritis in back, eosinophilic pneumonia, congestive heart failure, diabetes type 2, kidney and thyroid issues, chronic pain, dementia	We live next door to one another	96 + errands and managing care
Person 2					
Person 3					
Person 4					

3. For Person 1, who else cares for this person? Include family and friends, but do not include paid health care workers at this stage.

	Name	Relationship to you	How close do they live to you (nearby, 2 hours away, etc.)?	About how many hours a week do they help?
Carer 1		Brother	Lives on property	9
Carer 2		Brother	Lives next door	6
Carer 3				
Carer 4				

4. For Person 2, who else cares for this person? Include family and friends, but do not include paid health care workers at this stage.

Online Questionnaire
Caregivers interested in participating were first asked to complete an "Atlas of Caregiving Questionnaire". If selected, they were asked to complete a "Pre-Visit Survey".

Atlas of Caregiving Questionnaire
This questionnaire was used to screen interested participants. The intent was to choose participants who spent at least 2 hours per day on caregiving activities, and who as a group provided a diverse set (ages, conditions, urban/rural, ethnic/cultural background).

This brief survey asked:
- Name, email, phone, address?
- Who are the main caregivers and care recipients, and who is caring for whom?
- What are the ages and genders of these caregivers and care recipients?
- What are the major health conditions of these caregivers and care recipients?
- How would you describe the ethnic/cultural background of your family?
- How much time do you spend on caregiving activities each day, in general?
- Which of the caregivers and care recipients you described above will participate in the research?
- Preferred dates for participation?

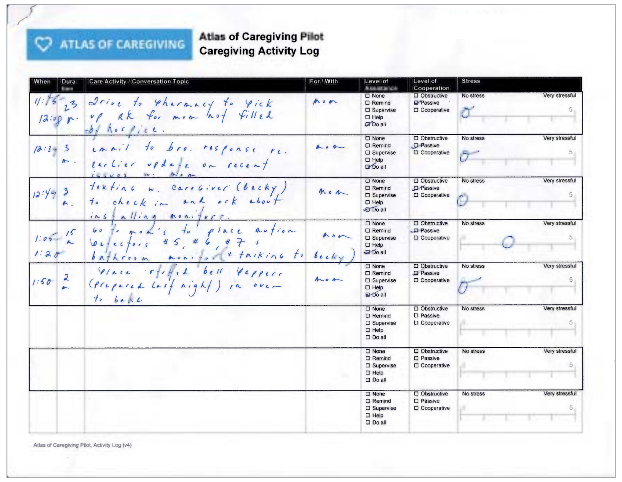
As it turned out, there were only three caregivers who filled out the questionnaire who did not participate. One because the care recipient changed his mind about participating, one because his care burden increased such that he felt he couldn't take the time to participate, and one because of scheduling difficulties.

Pre-Visit Survey
This questionnaire was meant to help the researchers prepare for the interviews by knowing a little bit about the various people and services involved. Participants were asked to fill this out quickly, to provide top-of-mind answers rather than worry about accuracy or completeness.

The survey asked:
- Who lives in your home, including pets?
- Who are the people (or pets) that you care for? Include those who need help because of physical or mental illnesses or disabilities. If you care for more than four people, indicate those requiring most of your time.
- For each of the four care recipients noted in the previous question—Who else cares for this person? Include family and friends, but do not include paid health care workers at this stage.
- Who helps care for you, or provides practical and emotional support?
- Do you belong to any online forums, or in-person networks (such as support groups or religious organizations) that provide support or advice?
- Do you have any home health aides, nurses or other paid assistants who come in to the home?
- Do you use any kind of paid assistance, such as cleaning service, food delivery service, dog walkers, child care providers, etc.?
- What medical professionals (nurses, doctors, nutritionists, etc.) have you interacted with in the past three months?
- Is there anybody else involved in your caregiving situation who we should know about?

Self Reporting

Caregiving Activity Log


ATLAS OF CAREGIVING
Atlas of Caregiving Pilot
Caregiving Activity Log

Caregiving Activity Log
Participants were given a clipboard with 10 pages of a printed log table to fill-out. There was enough space for 80 entries; only one participant came close to running out of space.

During interview #1, researchers reviewed the log with the participants. They were also given printed instructions, to refer to if necessary.

Participants were asked to keep track of all their caregiving and self-care activities and conversations. They were encouraged to provide as much detail as they could, and to break up activities into smaller tasks if possible. They were encouraged to think broadly as to what might constitute a "caregiving" activity, and were provided with a page-long list of example caregiving activities. Participants were also encouraged to note any non-caregiving activity (such as sleep, work, watching TV, etc.) that felt noteworthy or lasted more than 15 minutes. An example filled-in care log was also provided.

The log included several columns for each entry:

When
- The time of the activity/conversation

Duration
- How long they spent on the activity/conversation

Description
- For an activity: what was done
- For a conversation: the topic

For/With
- For an activity: participants (such as "me", "mom", etc.)
- For a conversation: who else was involved

Level of Assistance
- Only for caregiving activities. There were 5 checkboxes to note how much assistance was provided: None, Remind, Supervise, Help, and Do all.

Level of Cooperation
- Only for caregiving activities. There were 3 checkboxes to note the level of cooperation of the care recipient: Obstructive; Passive; Cooperative

Stress
- Participants were asked to note their level of stress at the time of the activity, using a 0–5 scale, where 0 = no stress, and 5 = very stressful

Wearable Sensors

Participants were asked to wear three devices during the study.

Narrative Clip
The Narrative Clip is a small wearable life logging camera that automatically takes one picture every 30 seconds throughout the day.

Photos are time-stamped and include rough GPS positioning information (e.g., nearest street). Photos can be viewed (after uploaded) on the Narrative website or mobile app.

Participants were told to clip it to their shirt or jacket and instructed how to disable it for privacy or sleep. (There's no on/off switch, but instead a light sensor. Participants were instructed to turn the camera face down to block light from reaching the sensor, effectively turning it "off".)

Empatica E4
Participants were asked to wear an Empatica E4, a wearable monitor with multiple sensors for:
- blood volume pulse (BVP)
- heart rate inter-beat interval
- motion (3-axis accelerometer)
- temperature
- electrodermal activity (EDA)

All data is stored in the device memory and later downloaded by the study's principle investigator. After uploading to Empatica's cloud, data can be previewed with the Empatica Connect Web Application.

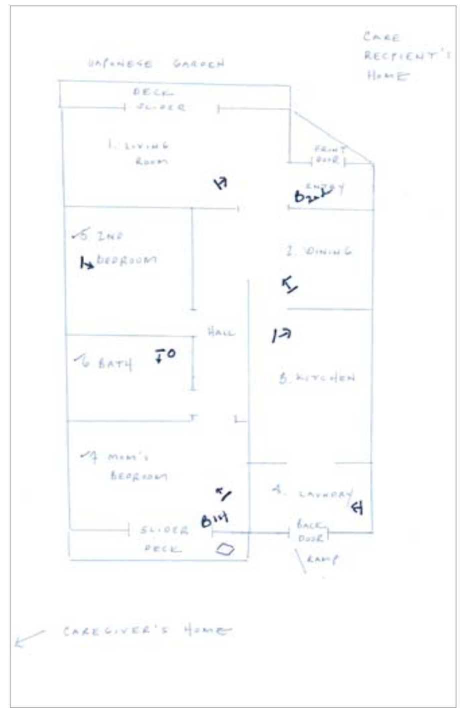
SmartSense Presence
The sensor sends a notification when it is within 500 feet of the SmartThings Hub. If worn by the participant it can be used as a proxy for presence at home.

To be effective it must be worn by a participant or attached to a key chain. Not all participants used the device and the data from it was inconclusive and not used in the study, which is why it is grayed out here.

Environmental Sensors

Participants were asked if as many as 10 sensors could be placed throughout their homes.

Participant's Home Setup
To minimize disruption and anxiety for the participants, a self-contained mobile hotspot was deployed with all of the connected devices, in a small wicker basket. The basket contained an AT&T mobile hotspot, a Netgear signal boosting and charging dock for the hotspot, an indoor Netatmo module, a SmartThings Hub, and an Apple Airport Express (to connect the hub to the Wi-Fi network). All of the devices were connected to a power strip. Plugging in the power strip to an outlet turned on all pre-configured devices.



Floorplan with locations of sensors
A hand drawn floor plan, indicating placement, and number of each sensor, is created for each household.

SmartSense Motion Sensors (×7)
The motion sensor sends a signal when movement is detected in its range.

The small, battery powered motion detectors were placed around the home, connected by Wi-Fi to the SmartThings hub.

We used SmartThing's API to connect to the Sapient-Nitro data pipeline and import into Datadeer.

SmartThings Hub (×1)
The SmartThings Hub connects all of the SmartThings sensors in the participant's home and sends the data to SmartThing's cloud. Data can be previewed with SmartThings website and mobile app.

Netatmo Weather Stations (×2)
The weather station has one indoor and one outdoor module. The indoor module can be connected to a Wi-Fi network to send data to Netatmo's cloud. Data can be previewed with Netatmo's website and mobile app.

The indoor module measures temperature, humidity, CO_2, barometric pressure, and sound. We found that CO_2 and sound data were good indicators of the participants presence in the home.

The outdoor module measures temperature and humidity. Placing the outdoor module in the bathroom, and observing humidity levels, enabled the detection of the participant showering.

We used Netatmo's API to connect to the Sapient-Nitro data pipeline and import into Datadeer.

Comparison of Sensor Size

Part 3

Wrangling Data

Scope of Production

Project Management
In addition to planning, data collection, and design, producing the diagrams was a huge amount of work. Organization was of paramount importance. We created a matrix which provided a top-level view of where we were in the process.

There are 20 columns; one for each participant. There are 14 rows; one for each diagram type. Inside each cell are details of what steps are required. Finally, the details in each cell were color-coded to show progress:
Cyan = done
Blue = in process
Magenta = to do
Gray = duplicate
Green = note

The final version is shown at right, so all the text is cyan.

Diagram Count
20 participants × 16 diagrams = 320 assets

Some participants don't have data for some diagrams and/or share households: 320 - 83 = 237 assets

The diagrams are shown at 2 sizes in 2 contexts*:
237 × 4 = 948 assets

To put a fine point on it:
- Print diagrams are shown twice
  - Small thumbnails in the 8.5-by-11 inch report
  - Full-size in the 17-by-11 inch appendix
- Web diagrams are shown twice
  - Previews in webpages
  - Full size in pop-out windows

Diagram Production Matrix

This matrix captures all participants and all diagrams on a single page, with brief descriptions of what is done, in process, and to do for each.
It presents a high-level view of project status and could continue to be updated throughout the project.

Key
Cyan = done
Blue = in process
Magenta = to do
Gray = duplicate
Green = note

	1 Ana	2 Chantal	3 Fay	4 Gabrielle	5 Hanna	6 Gaston	7 Harvey	8 Fernando	9 Laura	10 Ida	11 Nadine	12 Odette	13 Nate	14 Patty	15 Sally	16 Tammy	17 Rafael	18 Teddy	19 Omar	20 Cindy
1 **Care Network**	Sketch Draw Design	Sketch Draw Design	Sketch Draw Design	Sketch Draw Design	Sketch Draw Design			Sketch Draw Design		Sketch Draw Design	Sketch Draw Design	Sketch Draw Design	Sketch Draw Design		Sketch Draw Design	Sketch Draw Design		Sketch Draw Design	Sketch Draw Design	
2 **24-hour Log**	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	No data collected	CSV Design Upload to WP Download SVG	No data collected	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG
36-hour Log 3 Broken out 4 Combined	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	No data collected	CSV Design Upload to WP Download SVG	No data collected	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG
Activities 5 Hours spent 6 Minutes spent	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	No data collected	CSV Design Upload to WP Download SVG	No data collected	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG	CSV Design Upload to WP Download SVG
7 **Photo Log**	Design Run script Export SVG	Design Run script Export SVG	Design Run script Export SVG	Design Run script Export SVG	Design Run script Export SVG	Design Run script Export SVG	Design Run script Export SVG	Design Run script Export SVG	Design Run script Export SVG	Design Run script Export SVG	Design Run script Export SVG	Design Run script Export SVG	Design Run script Export SVG	Design Run script Export SVG	Design Run script Export SVG	Design Run script Export SVG (+ Wanda's room)	Design Run script Export SVG	Design Run script Export SVG	Design Run script Export SVG	Design Run script Export SVG
Body 9 EDA 9 Heart Rate 10 Movement	CSV Design Process in MatLab	CSV Design Process in MatLab	CSV Design Process in MatLab	CSV Design Process in MatLab	CSV Design Process in MatLab	CSV Design Process in MatLab	CSV Design Process in MatLab	CSV Design Process in MatLab	No data collected	CSV Design Process in MatLab	CSV Design Process in MatLab	CSV Design Process in MatLab	CSV Design Process in MatLab	CSV Design Process in MatLab	CSV (3 files) Design Process in MatLab	CSV Design Process in MatLab	CSV Design Process in MatLab	CSV Design Process in MatLab	CSV Design Process in MatLab	CSV (battery died) Design Process in MatLab
Environment 11 CO2 12 Humidity 13 Noise	No data collected	Data in DD Design Download SVG	Data in DD Design Download SVG	Data in DD Design Download SVG	No data collected (unknown technical problems)			No data collected		Data in DD Design Download SVG	Data in DD Design Download SVG	Data in DD Design Download SVG	Data in DD Design Download SVG		No data collected	Data in DD Design Download SVG		No data collected	No data collected	
Motion 14 Motion 15 Floorplan	No data collected No floorplan	Data in DD Design Download SVG Floorplan	Data in DD Design Download SVG Floorplan	Data in DD Design Download SVG Floorplan	No data collected No floorplan			No data collected No floorplan		Data in DD Design Download SVG Floorplan	No data collected (hotspot failure?) No floorplan	Data in DD Design Download SVG Floorplan	Data in DD Design Download SVG Floorplan		No data collected No floorplan	Data in DD (only 1 day) Design Download SVG Floorplan		No data collected No floorplan	No data collected No floorplan	
16 **Summary**	Design - 36-hour Log - Body - Environment - Motion	Design - 36-hour Log - Body - Environment - Motion	Design - 36-hour Log - Body - Environment - Motion	Design - 36-hour Log - Body - Environment - Motion	Design - 36-hour Log - Body - Environment - Motion	Design - 36-hour Log - Body - Environment - Motion	Design - 36-hour Log - Body - Environment - Motion (insufficient data)	Design - 36-hour Log - Body - Environment - Motion	Design - 36-hour Log - Body - Environment - Motion (insufficient data)	Design - 36-hour Log - Body - Environment - Motion	Design - 36-hour Log - Body - Environment - Motion	Design - 36-hour Log - Body - Environment - Motion	Design - 36-hour Log - Body - Environment - Motion	Design - 36-hour Log - Body - Environment - Motion	Design - 36-hour Log - Body - Environment - Motion	Design - 36-hour Log - Body - Environment - Motion	Design - 36-hour Log - Body - Environment - Motion	Design - 36-hour Log - Body - Environment - Motion	Design - 36-hour Log - Body - Environment - Motion	Design - 36-hour Log - Body - Environment - Motion

Study Length

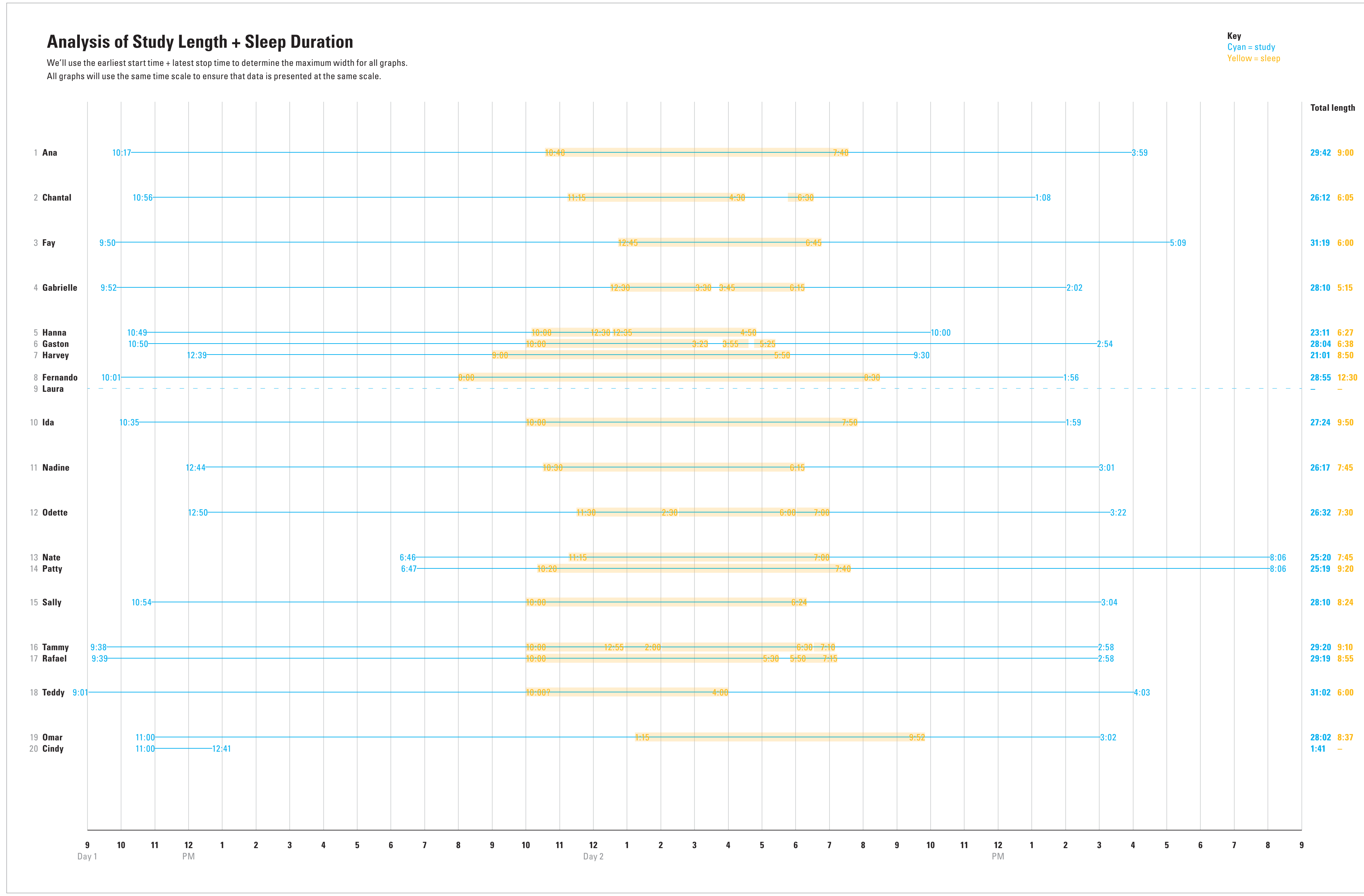
We needed to find a consistent time interval on which to plot the data for all data streams and participants. In order to do that we had to find the earliest start time and latest stop time, for each study.

We inspected each Empatica E4 data file and noted the start time (first row), and end time (last row). Out of all the data we had, we used the Emaptica data because we had data from every participant (not all participants allowed us to place environment sensors in their homes). We learned that no study started before 9 AM the first day, and none ended after 9 PM the second day. All of the data could be displayed consistently on a 36 hour scale. The diagram at right shows the varying study lengths for each participant (blue lines).

To add more information, we inspected the self-reported log spreadsheets to find sleep duration (yellow highlight lines).

Hour Width
If we start with 36 hours and spread them across 1080 points, then each hour is 30 points wide. A nice round number.

15 inches wide at 72 dpi = 1080 points
1080 points / 36 hours = 30 points per hour



Data Volumes

The IoT and our sensors are generating vastly more data than we have seen before.

Just one device, the Empatica, has six physical sensors:
- A blood volume pulse (BVP) sensor
- Three accelerometers
  (one arrayed in each of three axes: x, y, and z)
- An electrodermal activity (EDA) sensor
- A skin temperature sensor

The BVP sensor is running at 64 Hz. That means it makes a reading every 1/64th of a second. 60 seconds comprise a minute; 60 minutes comprise an hour; and 36 hours is the maximum duration of one of our study sessions. In other words, one study session comprises 2,160 minutes, and just one of the sensors is collecting 3,840 samples per minute. That's 8,294,400 samples collected over the course of one 36-hour session.

Add to that samples for the other five sensors:
8,294,400 samples for BVP (at 64 Hz)
4,147,200 samples for x axis acceleration (at 32 Hz)
4,147,200 samples for y axis acceleration (at 32 Hz)
4,147,200 samples for z axis acceleration (at 32 Hz)
518,400 samples for EDA (at 4 Hz)
518,400 samples for skin temperature (at 4 Hz)

21,772,800 samples of raw data for one participant

That's not all. The raw data needs to be processed to make useful data, which leads to still more data. For example, the x, y, z acceleration data is used to calculate motion, and BVP is used to calculate heart rate. (We did not use skin temperature.)

To understand just how extraordinary these numbers are, compare this fact. Google, one of the largest data processors on the planet, builds special systems for processing vast amounts of data. The idea that one of Google's systems would add 10 or 15 million new rows of data per day is considered large, though not surprising. But that's for ALL of Google, in which there may be tens of such systems. Doing so for hundreds, thousands, or millions of people—that's unheard of. And yet, we are on the verge of such a change.

The shear amount of data generated in this study gives a flavor. 14 studies with 20 participants. Roughly 270 MB of data per participant. That's 5,400 MB, roughly 5 GB. That's equivalent to downloading 2-5 feature films.

And so far, we've only discussed one device. In the study, we also included several other devices with sensors.

Coping with the unprecedented quantity of data—simply "wrangling" it, storing it, moving it around, and visualizing it—will require a new generation of tools. Unfortunately, Microsoft Excel just wasn't up to the task of opening the full 270 MB Empatica file generated by one participant. And even Tableau, a specialized analytics tool, proved unwieldy. In the end, we had to turn to scripting—to writing code to process the data—using Python and MatLab. Clearly, not having powerful and easy-to-use tools will hold back analysis and progress.

Screenshot illustrating sampling rates

Here's 1 second (64 rows) from an Empatica data file. One can see the how the sampling rates differ from column to column.

BVP	EDA	HR*	X	Y	Z	AVG*
@64 Hz	@4 Hz	@1 Hz	@32 Hz	@32 Hz	@32 Hz	@1 Hz
Every row	Every 15 rows	Every 64 rows	Every 2 rows	Every 2 rows	Every 2 rows	Every 64 rows

Date & Time	BVP	EDA	HR	x	y	z	moving AVG ACC
2015-08-31 09:50:47	0	0		-12	16	60	
2015-08-31 09:50:47	0						
2015-08-31 09:50:47	0			-12	16	60	
2015-08-31 09:50:47	0						
2015-08-31 09:50:47	0			-12	16	60	
2015-08-31 09:50:47	0						
2015-08-31 09:50:47	0			-11	16	60	
2015-08-31 09:50:47	0						
2015-08-31 09:50:47	0			-11	16	61	
2015-08-31 09:50:47	0						
2015-08-31 09:50:47	0			-10	15	60	
2015-08-31 09:50:47	0.01						
2015-08-31 09:50:47	0.01			-10	15	61	
2015-08-31 09:50:47	0						
2015-08-31 09:50:47	-0.03			-12	15	60	
2015-08-31 09:50:47	-0.05						
2015-08-31 09:50:47	-0.05	0.836457		-12	16	59	
2015-08-31 09:50:47	0						
2015-08-31 09:50:47	0.13			-13	15	60	
2015-08-31 09:50:47	0.36						
2015-08-31 09:50:47	0.66			-13	15	60	
2015-08-31 09:50:47	1.01						
2015-08-31 09:50:47	1.37			-12	15	60	
2015-08-31 09:50:47	1.72						
2015-08-31 09:50:47	2.06			-10	15	61	
2015-08-31 09:50:47	2.4						
2015-08-31 09:50:47	2.79			-10	16	60	
2015-08-31 09:50:47	3.24						
2015-08-31 09:50:47	3.8			-11	16	59	
2015-08-31 09:50:47	4.42						
2015-08-31 09:50:47	5.06			-12	15	60	
2015-08-31 09:50:47	5.66						
2015-08-31 09:50:47	6.2	1.18386		-12	15	60	
2015-08-31 09:50:47	6.68						
2015-08-31 09:50:47	7.15			-11	15	61	
2015-08-31 09:50:47	7.68						
2015-08-31 09:50:47	8.31			-10	15	61	
2015-08-31 09:50:47	9.14						
2015-08-31 09:50:47	10.09			-11	15	60	
2015-08-31 09:50:47	11.04						
2015-08-31 09:50:47	11.88			-12	15	60	
2015-08-31 09:50:47	12.56						
2015-08-31 09:50:47	13.14			-11	15	60	
2015-08-31 09:50:47	13.76						
2015-08-31 09:50:47	14.65			-11	16	60	
2015-08-31 09:50:47	15.96						
2015-08-31 09:50:47	17.91			-12	16	60	
2015-08-31 09:50:47	20.32						
2015-08-31 09:50:47	22.9	1.622013		-13	15	60	
2015-08-31 09:50:47	25.38						
2015-08-31 09:50:47	27.58			-13	15	61	
2015-08-31 09:50:47	29.57						
2015-08-31 09:50:47	31.68			-12	16	60	
2015-08-31 09:50:47	34.32						
2015-08-31 09:50:47	37.81			-11	15	60	
2015-08-31 09:50:47	42.54						
2015-08-31 09:50:47	48.15			-11	15	60	
2015-08-31 09:50:47	54.1						
2015-08-31 09:50:47	59.83			-11	15	60	
2015-08-31 09:50:47	64.97						
2015-08-31 09:50:47	69.58			-13	14	59	
2015-08-31 09:50:47	74.14						
2015-08-31 09:50:47	79.34			-12	16	59	
2015-08-31 09:50:47	85.71						

*HR is derived from BVP.
*AVG is derived from XYZ.

Raw Data Count

The Empatica device has six sensors, sampling at different rates, generating a vast amount of data.

BVP Blood Volume Pulse	**@ 64 Hz** 64 samples per second	**= 3,840 rows** per minute	**× 2160 minutes** 60 minutes per hour 36 hours	**= 8,294,400 rows**	+
X accelerometer	**@ 32 Hz** 32 samples per second	**= 1,920 rows** per minute	**× 2160 minutes** 60 minutes per hour 36 hours	**= 4,147,200 rows**	+
Y accelerometer	**@ 32 Hz** 32 samples per second	**= 1,920 rows** per minute	**× 2160 minutes** 60 minutes per hour 36 hours	**= 4,147,200 rows**	+
Z accelerometer	**@ 32 Hz** 32 samples per second	**= 1,920 rows** per minute	**× 2160 minutes** 60 minutes per hour 36 hours	**= 4,147,200 rows**	+
EDA Electro Dermal Activity	**@ 4 Hz** 4 samples per second	**= 240 rows** per minute	**× 2160 minutes** 60 minutes per hour 36 hours	**= 518,400 rows**	+
Skin Temp Blood Volume Pulse	**@ 4 Hz** 4 samples per second	**= 240 rows** per minute	**× 2160 minutes** 60 minutes per hour 36 hours	**= 518,400 rows**	

= 21,772,800 rows
per participant

× 19 participants
Ana
Chantal
Fay
Gabrielle
Hanna
Gaston
Harvey
Fernando
Ida
Nadine
Odette
Nate
Patty
Sally
Tammy
Rafael
Teddy
Omar
Cindy

= 413,683,200 rows
nearly half a billion data points

Raw Data Processing

The raw data needs to be processed to make useful data, which leads to still more data. For example, BVP is used to calculate heart rate, and the x, y, z acceleration data is used to calculate motion. (We did not use skin temperature.)

Table showing how source data was processed
Gray arrow indicate processes done by Empatica, blue arrows indicate processes done by DDO.

BVP	**Accelerometers**	**EDA**	**Skin Temp**
Blood Volume Pulse @ 64 Hz	X, Y, Z dimensions @ 32 Hz	Electro Dermal Activity @ 4 Hz	Blood Volume Pulse @ 4 Hz
used to calculate	used to calculate	filter	not used
Heart Rate @ 1 Hz	**Moving Average** @ 1 Hz	**Filtered EDA** @ 1 Hz	
	convert units		
	Unit Conversion from 1/64 G-force to decimal G-force		

Downsampling

Let's start with our highest sampling rate—64 Hz for BVP—because that's the most data we need to accommodate. All other data is sampled at less frequent rates.

64 times per second
× 60 seconds per minute
× 60 minutes per hour
× 36 hours in a single study
= 8,294,400 data points

Many more data points than we could ever show. So we have to downsample the data, but how? What's the right balance between showing what's useful and usable?

If we downsample from 64 times per second
to 1 time per second (1 Hz):

1 time per second
× 60 seconds per minute
× 60 minutes per hour
× 36 hours in a single study
= 129,600 data points

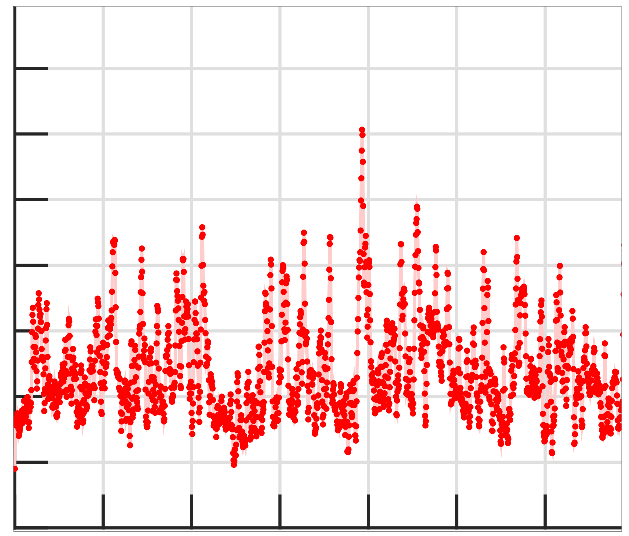
A much smaller data set, but still much more than we could effectively display so we continued to explore different downsampling rates based on the logic above.

1/15 seconds = 8640 data points
1/30 seconds = 4320 data points
1/60 seconds = 2160 data points

This is when we had an insight.

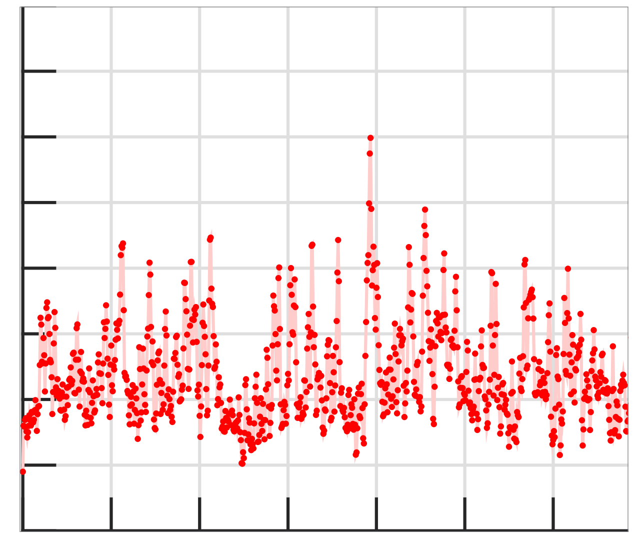
2160 = 2 × 1080.
1080 is a common screen resolution.

Heart rate sampled once every 15 seconds



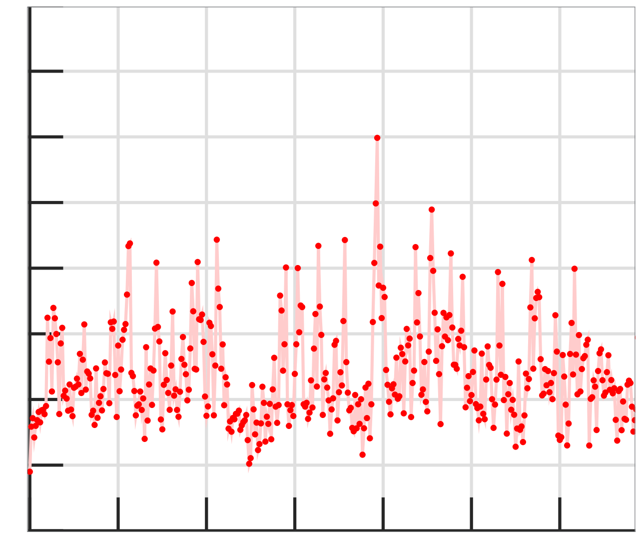
When we sampled once every 15 seconds, the results gave us 8640 data points for the 36 hours of the study.

Heart rate sampled once every 30 seconds



When we sampled once every 30 seconds, the results gave us 4320 data points for the 36 hours of the study.

Heart rate sampled once every 60 seconds



When we sampled once every 60 seconds, the results gave us 2160 data points for the 36 hours of the study.

Presentation

While deciding an optimal downsampling rate, we had to be mindful of the final presentation as they have a relationship and effect each other.

The diagrams would be the heart of a detailed written report. The report would be formatted as a letter sheet (8.5-by-11 inches) in portrait format because it's a standard for papers. We knew that the diagrams would not present well at small sizes so we decided early on that we would create a larger format appendix in which to present the diagrams. We would optimize the diagrams for a tabloid sheet (11-by-17 inches) in landscape format, because most of the data was in a time series.

Let's do some calculations about the width of the diagrams in relation to the number of data points.

17 inches wide at 72 dpi = 1224 points

But of course, we can't use the entire 17 inches.
We need to account for margins, axis labels, header, footer, etc. If we allow 1 inch margins around all edges we end up with a usable area of 15-by-9 inches.

Let's see how many points we can fit into a 15 inch wide area.

15 inches wide at 72 dpi = 1080 points
15 inches wide at 144 dpi = 2160 points
15 inches wide at 288 dpi = 4320 points

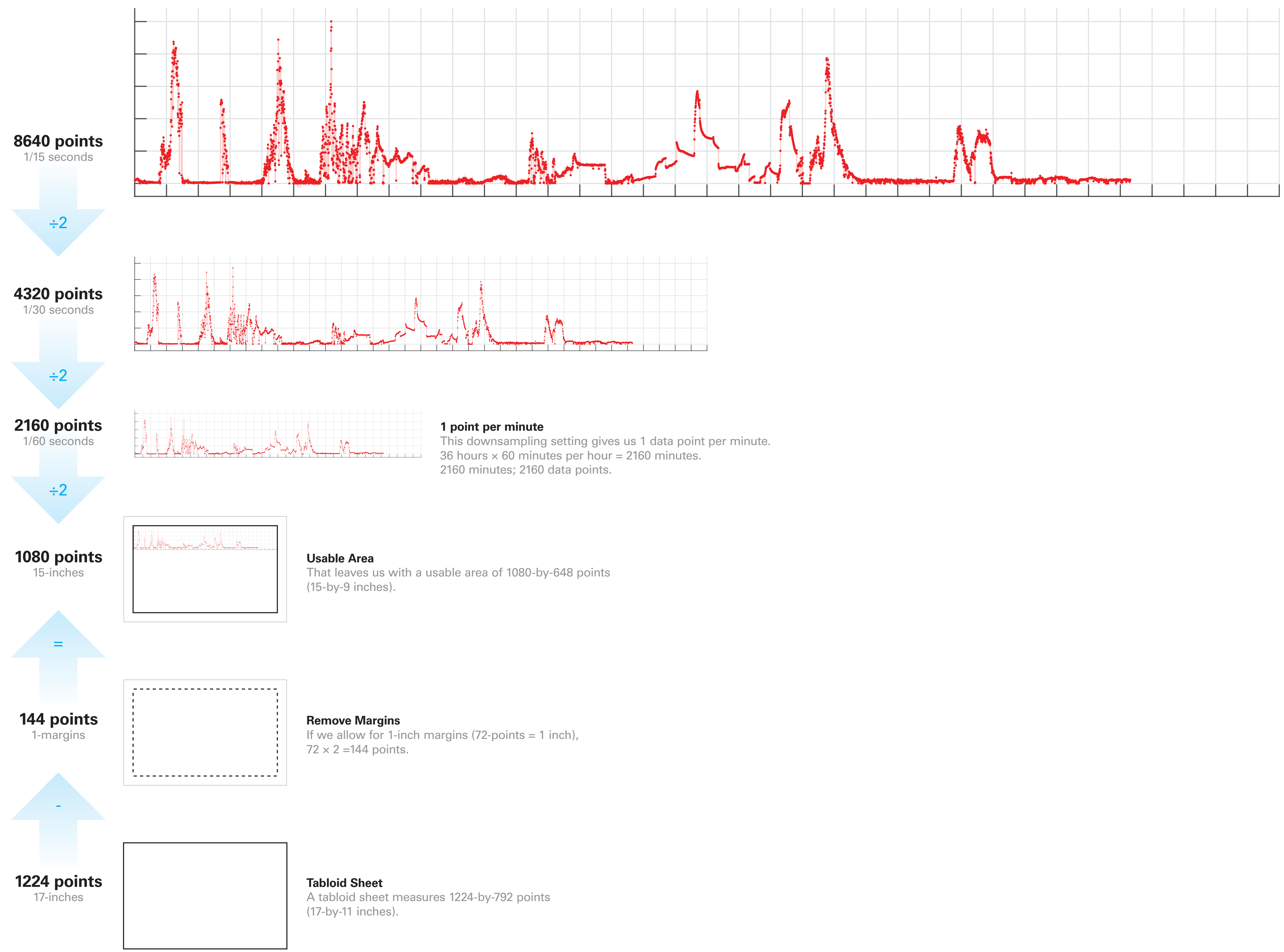
One can see the relationship between the downsampling rates on the previous page, and the presentation numbers above. Based on this relationship, we decided to downsample the data at 1/60 (once a minute). This gave us a much smaller file, but still enabled us to display many data points.

2160 is 2 × 1080; very orderly.

So we took our data output and scaled it down 50% to fit perfectly in our 1080 usable area.

With these settings,
1 point in the diagram = 2 minutes in the data;
and 30 points = 1 hour;
which is very understandable.

Convergence of downsampling with designing the presentation area



Filtering

A key finding is that identifying individual moments of arousal (as measured by changes in skin conductance associated with sweating, which is associated with stress and other forms of arousal) can be difficult with today's technology. While the process can be useful in short sessions managed in controlled settings, it is far more challenging in long sessions that are unmanaged, and in which the data is being analyzed post hoc, rather than being monitored in real time. What's more, correlations with video and researcher observations are difficult in longer sessions (plus we had no video or audio with which to correlate).

One source of difficulty is the confounding of factors. For example a person moving may cause sweat, without correlation to stress. Likewise, a person sleeping under heavy blankets may simply get hot and start sweating. In some cases, sophisticated signal processing techniques may be used to tease out instances of stress from background "noise". For example, skin conductance might be compared to data on movement to suggest that a spike was related to activity. Or the rate of change (how fast a signal spikes) might also be used to suggest becoming hot rather than becoming stressed. These examples suggest the potential for significant future research. And the need for easy-to-use signal processing tools.

Case Study: Stress or Sweat?
The image at top right shows a plot of the raw, unfiltered electrodermal activity (EDA) data. Remember, EDA is arousal, which might be due to skin sweatiness, which is the best proxy we have for stress.

The yellow bar indicates the time spent sleeping. Viewers might not expect to see big EDA spikes during the middle of the night, and yet we can clearly see one just before 4 AM. Was the participant stressed in a lucid dream? Or did they simply get too hot under their blankets?

Our PI decided that this was sweat and not stress and the spike misrepresented what was really happening. The spike would have to be removed, but how?

Our PI found a plug-in to MatLab called Ledalab which performed the filtering he was looking for, but unfortunately, couldn't output the vector graphics we required. We called on Patch Kessler, our MatLab expert to see if he could understand what Ledalab's algorithm or output the data.

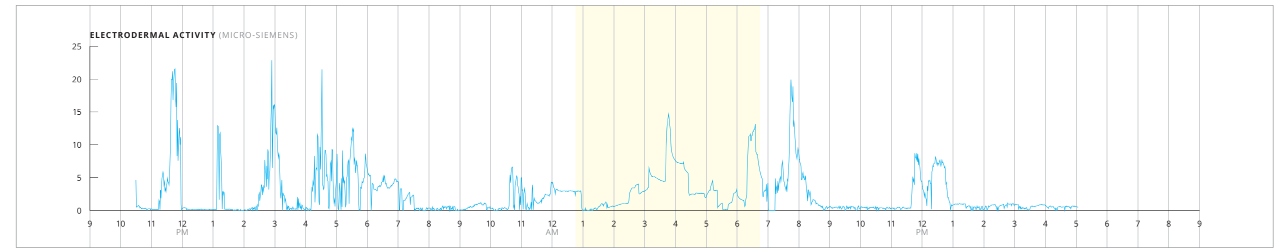
In the context of EDA data, "slow" (tonic) and "fast" (phasic) seem to refer to a slow moving baseline, and then fluctuations about that baseline. The image at middle right shows an experiment filtering the EDA data. The tonic and phasic data were computed using a basic IIR filter (Infinite Impulse Response) on the original data. The image shows the original EDA data (blue), and a moving "tonic" baseline (red). Learn more about IIR filters at: https://en.wikipedia.org/wiki/Infinite_impulse_response

After detailed inspection, we realized that this method of filtering was not showing the same results as the Ledalab plug-in. So rather than re-invent the wheel, Patch spent some time getting Ledalabs to work and output vector graphics.

The image at bottom right shows a final EDA data plot. The original data is shown in gray and the filtered data is shown in blue.

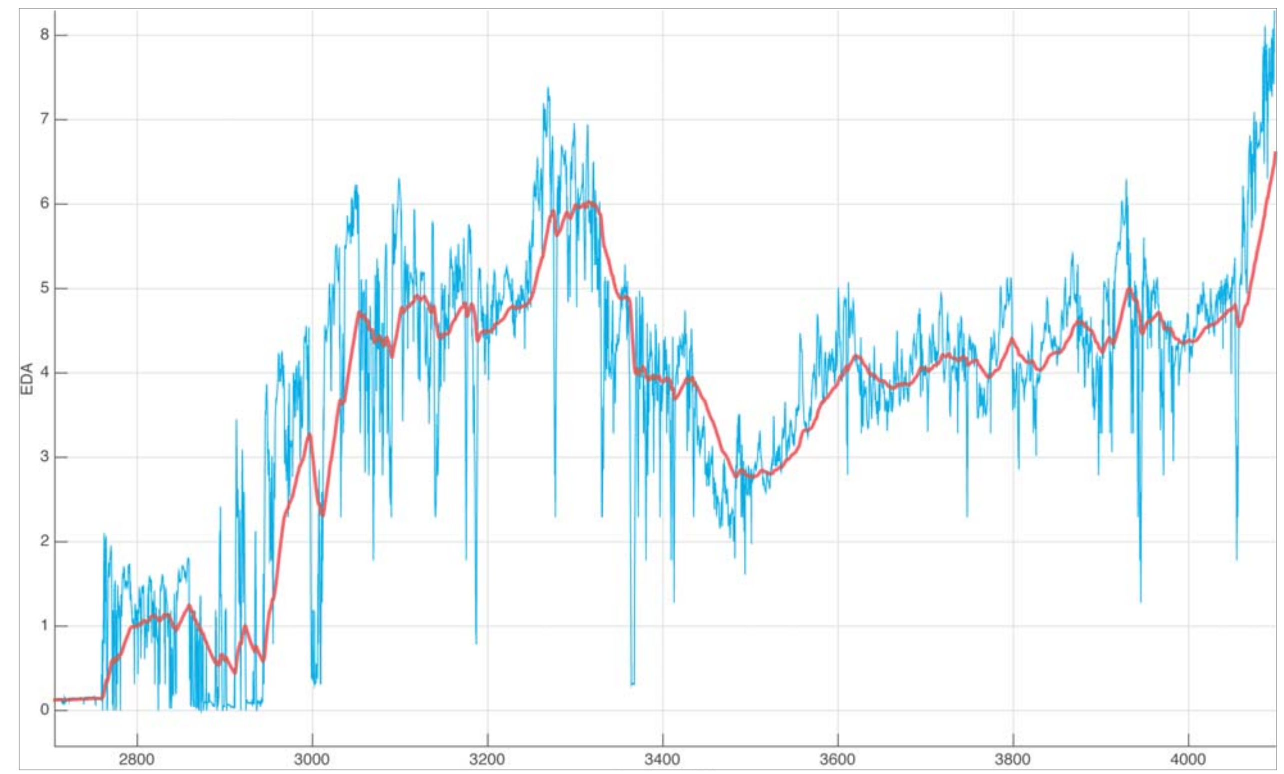
Learn More about Ledalabs
The algorithm appears to be an attempt to undo some physiological dynamics by deconvolving the raw data with some standard response functions. The analysis details are available in following paper by the authors of the software: http://www.sciencedirect.com/science/article/pii/S0165027010002335 or visit http://www.Ledalab.de

Unfiltered Data



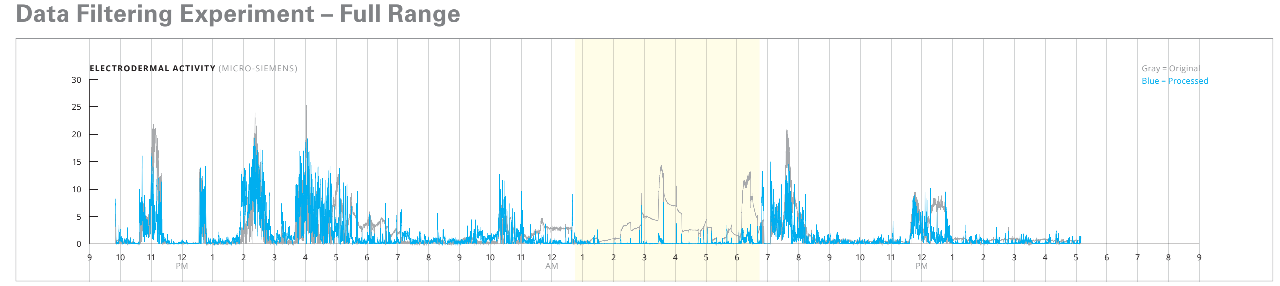
Full 36 hours of EDA, with sleep indicated in colored bars. Note spike during sleep. What's the cause? Stress or simply getting hot?

Data Filtering Experiment – Detail



Blue = original data Red = moving "tonic" baseline

Data Filtering Experiment – Full Range

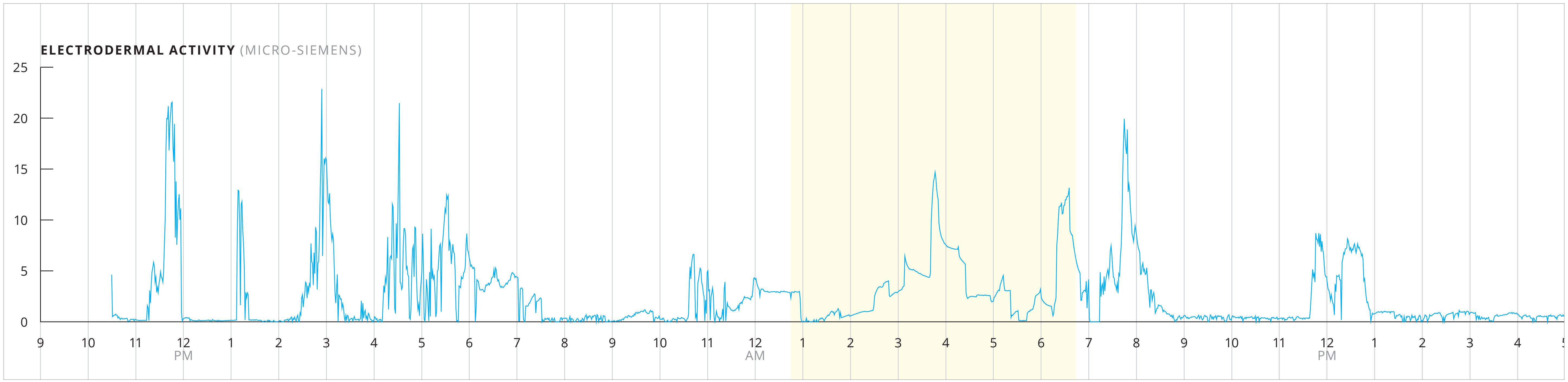


Gray = original data Blue = processed data

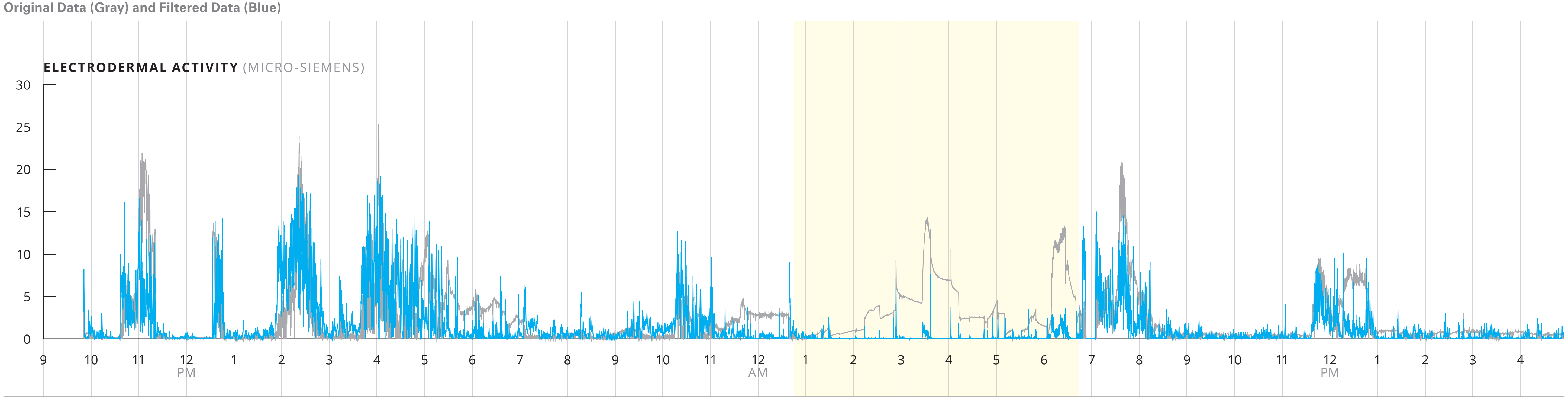
Filtering Detail

Here's a zoomed in view of the unfiltered and filtered data for comparison.

Unfiltered Data



Original Data (Gray) and Filtered Data (Blue)



Design

Principles

These principles provided guidance when it came time to design the diagrams:

- The x-axis is always a consistent width of 36 hours; from 9 AM the first day, to 9 PM the second day.

- The y-axes within a diagram type must remain consistent so that viewers have a basis for comparison. Varying the axis scales is bad practice.

- The y-axes must start at zero. Starting at a non-zero value is misleading and is bad practice.

- The values of the y-axes must be derived from the data itself. The highest value observed becomes the uppermost unit. The range from highest to lowest value is divided into logical, round numbers.

- We would cut extra data before the start and after the stop of the study, but we would not go into the middle of the data and delete or modify data.

- Make data "pop out". The project was all about data, so we wanted the data to come first. We made cyan the default color for data as it relates back to the project branding. We also made anything that wasn't data black or gray.

- Don't combine data plots. We wanted the data to be presented simply and stand alone. Some sensors collect multiple types of data—the Netatmo weather station for example senses CO2 and Noise—but each has a different Y axis scale. If we combined the plots, we'd have two Y scales which gets complicated.

We visualized the data collected and presented it in a consistent and non-arbitrary fashion. Others could achieve similar results by following our methods.

Visual Details

There were many visual details to consider when designing the diagrams. We performed an exhaustive exploration. The list below provides a sense of the space of possibilities and highlights the choices we made.

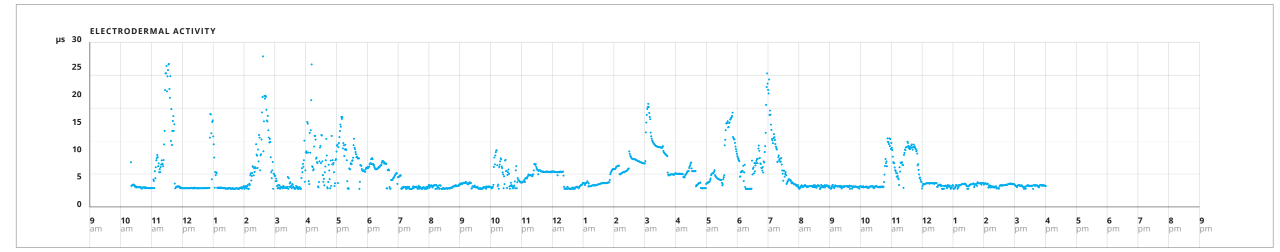
X-axis typography
- Left-aligned vs. center-aligned numerals
- Bold vs. roman numerals
- Black numerals vs. gray
- AM/PM designator for every entry vs. select times
- Black vs. gray time designators
- Small ticks vs. extended lines

Y-axis typography
- Baseline-aligned vs. baseline shifted numerals
- Bold vs. roman numerals
- Black numerals vs. gray
- Small ticks vs. extended lines
- Unit label combined with title vs. center-aligned vs. bottom-aligned vs. top-aligned
- Unit label bold vs. roman
- Unit label black vs. gray
- Unit label caps vs. large
- Unit label semi-colon vs. parens vs. em dash

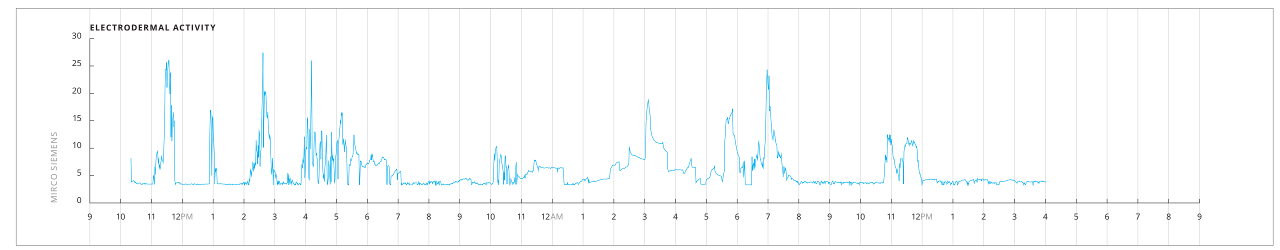
Chart area
- Vertical lines vs. horizontal lines vs. both
- Solid lines vs. dashed vs. dotted
- Thick vs. thin lines
- Black vs. gray lines
- White background vs. gray vs. other

Line style
- Fill vs. line
- Dots vs. lines vs. both
- 0.25 vs. 0.5 vs. 1.0 stroke weight
- 0.50 vs. 1.0 vs. 2.0 dot radius
- Butt cap vs. round vs. pointed

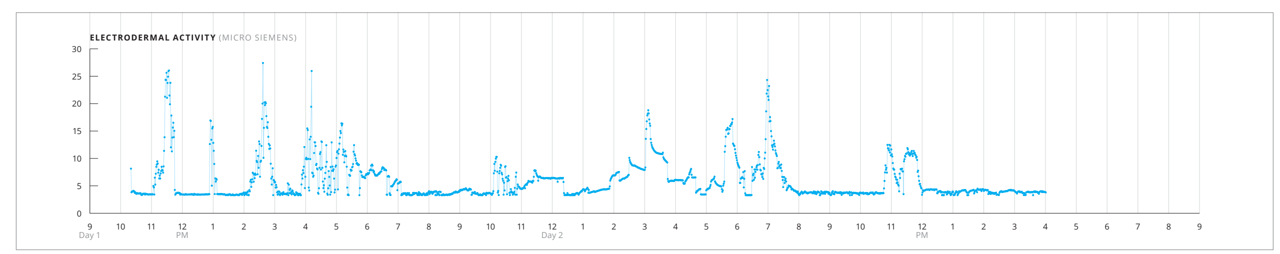
Sample Explorations



Dots for data, bold numerals, am/pm for each entry, extended lines for both X and Y axes, and units labeled at top.

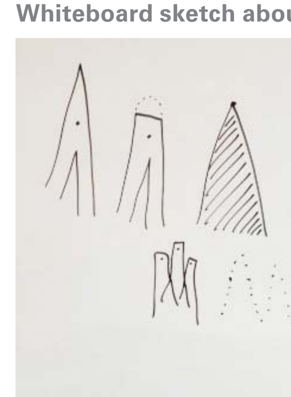


Solid data line, roman numerals, am/pm for selected entries, extended lines for X axis only, and units labeled at bottom.



Dots and lines for data, roman numerals, am/pm for selected entries, day 1 and 2 indicated, and units combined with title.

Whiteboard sketch about line caps



Adobe Illustrator features different types of line corners, or "caps".

The pointed style extends past the data point to create a sharp tip, but we feared the data would be perceived as taller than it actually is.

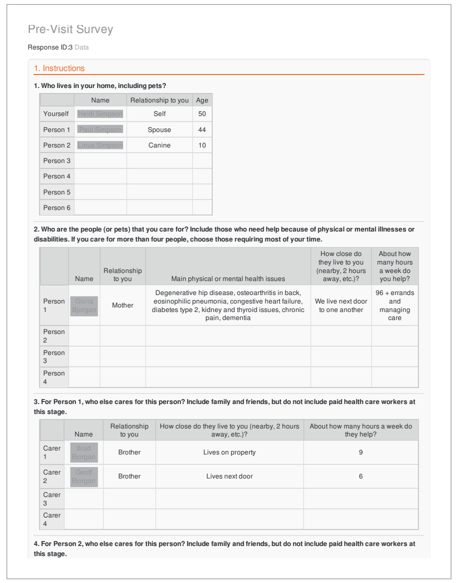
We explored using round caps, butt caps, and even using a fill instead of a line stroke.

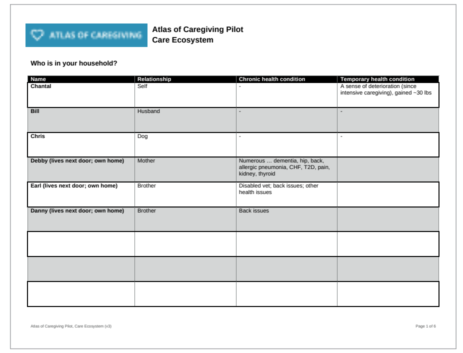
Line weight also becomes a concern when dealing with high density data, because it start to overlap with itself, obscuring the data.

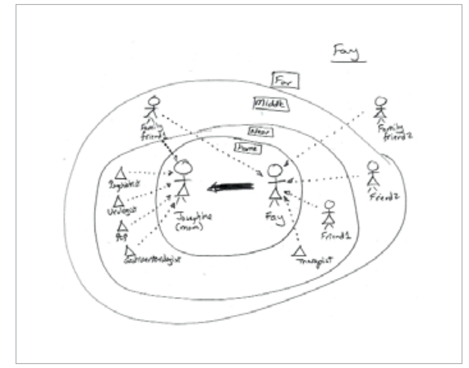
Part 4

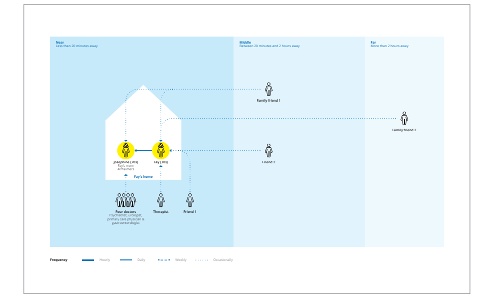
Data Processing to Create Visualizations

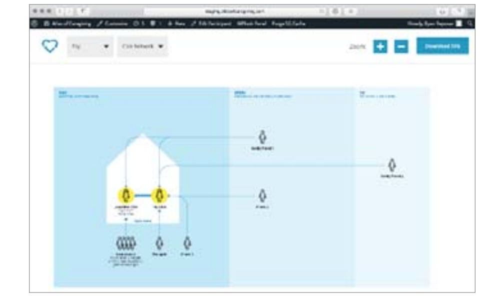
Care Network Diagram











Pre-Visit Survey
Participants are asked to fill out this form quickly, to provide top-of-mind answers rather than worry about accuracy or completeness.

Care Ecosystem
Information in the Pre-visit Survey (as well as information gathered in second interview) is transcribed into the Care Ecosystem document.

Names are aliased when entered into the document.

Rajiv Word

Care Network Sketch
Information in Care Ecosystem document is used to sketch a diagram of the caregiving relationships. See blog post showing and explaining how to draw and learn from care maps: http://atlasofcaregiving.com/put-your-family-caregiving-on-the-map/.

The form of the diagram and notations for indicating the relationships have evolved over the duration of the project. See document *Care Network Diagram Style Rationale* (AoC_Care_Network_Diagram Style_151223b.pdf) for more details on the design thinking.

Rajiv Pen + Paper

Care Network Diagram (Print)
Using information in Care Ecosystem document and Care Network Sketch, a final diagram is created in Adobe Illustrator using predefined rules and symbols.

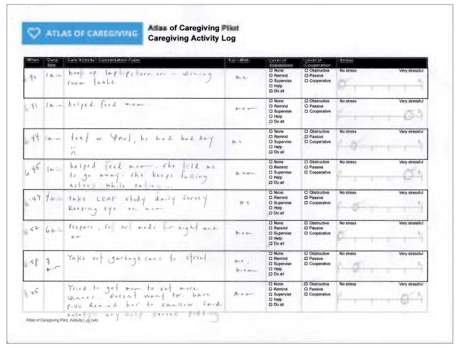
Result: 1 diagram × 14 households

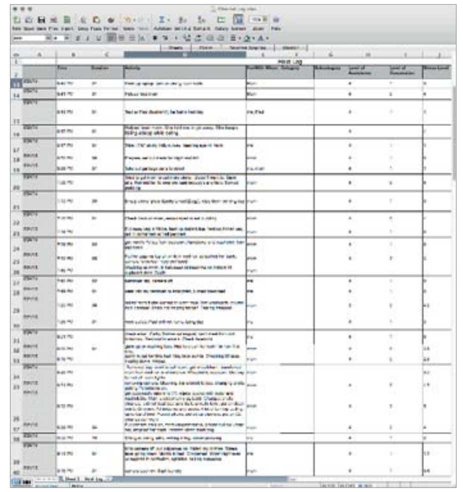
Knut Illustrator

Care Network Diagram (Web)
An SVG is saved from Illustrator and uploaded to the appropriate participant directory in WordPress. Website users can zoom in and out of the SVG to view more detail, and can download the SVG to their computer.

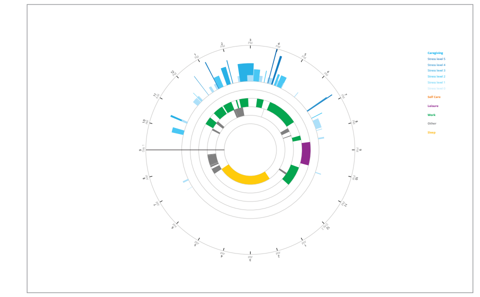
Ryan Wordpress

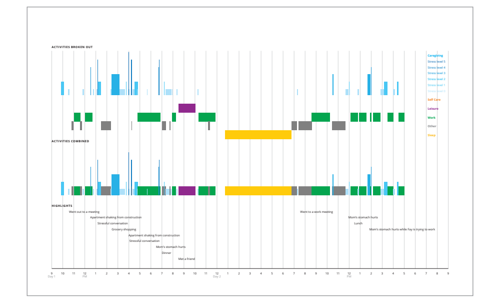
Activity Logs

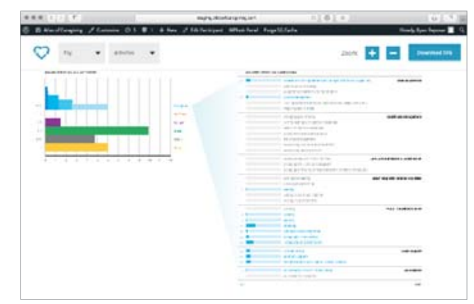


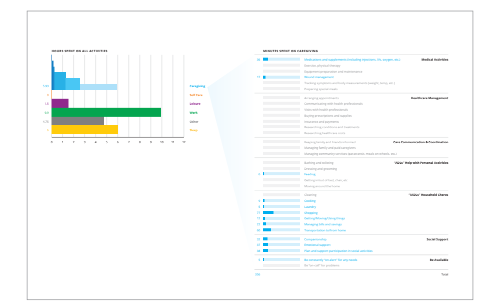


Program









Hand Written Caregiving Activity Log
Study participants fill in the log with time, duration, activity description, who care is for, level of assistance, level of cooperation, and level of stress.

Accuracy and the amount of data collected is dependent on each participant.

Transcribed Caregiving Activity Log
Activity log is transcribed into a spreadsheet (XLS). Activities are coded against a standardized and categorized list (a “controlled vocabulary”).

CSV File
The spreadsheet is saved as a CSV (comma separated value) file.

Date	- data not used
Time	- start of activity
Duration	- length of activity
Activity	- displayed in rollover
For/With Whom	- data not used
Category	- color code activity
Subcategory	- activities details
Sleep	- binary 0/1
Labels	- text for “highlights”
Level of Assistance	- data not used
Level of Cooperation	- data not used
Stress Level	- plotted in activity

Note: Cell formatting has to be followed strictly, as any mistakes (typos, etc.) will result in charting display errors. Activity rollover text was later removed due to privacy concerns.

Dawn Excel

Activity Logs (Web)
Data in the CSV file is input into a customized D3 charting program and Wordpress template to generate three diagram variations:
- 24-hour Log (radial)
- 36-hour Log (linear)
- Activities
  - adding up total hours spent on all activities
  - detailing total minutes spent on caregiving

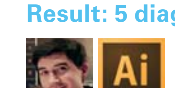
Note: 24 and 36 hour diagrams feature rollover text on the website.

Cody Program D3 Wordpress

Activity Logs (Print)
The final step was to download the SVGs from the website and place into Illustrator. Some special handling is required to ensure elements displayed accurately in Illustrator:
24-hour Log = scale down 80%
36-hour Log = place in 1280 × 900 canvas
Activities = place in 1220 × 900 canvas

Note: The 36-hour log was further customized inside of Illustrator: a “combined” version was made with all activities in a single line, and highlight text items were added manually from the “Labels” column in the log.csv files.

Result: 5 diagrams × 18 participants



Ryan Illustrator

Photo Log

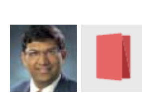
Narrative Clip
The Narrative Clip automatically takes one picture every 30 seconds throughout the day, storing the photos in its memory.

The camera is worn by participants as a clip or pendant. The direction the camera faces depends on where the participant is facing. Image can be somewhat random.

Narrative Uploader App
When installed on a Mac or PC, the Uploader app transfers the photos, via a USB cable, to the computer and uploads the photos to Narrative's cloud.

Note: If the option is selected the Narrative Uploader app will save photos to the connected computer with raw, unprocessed images.

We did not use these in production.



Preview in Narrative App and Website
After uploading to Narrative's cloud storage photos can be viewed in Narrative's mobile app and website.

The mobile app was used to review photos against the activity logs and in participant follow-up interviews to confirm activities.

Photos uploaded to the narrative cloud are processed—with auto image rotation, exposure adjustment, noise filtering, and saved with a name based on their time stamp.

Example File Name:

Photos Downloaded from Website
Processed photos are downloaded for production.

Photos Processed with Automator
With large file size (300-500k) and variations in orientation and image size of processed photos, custom scripts were created with Apple's Automator app to batch crop and resize photos to 64 × 64 pixels.

Batch process resized to shortest size first, then cropped off the long edges to make a square photo.

Photo Log (Print)
A custom program is used to select photos at 15-minute intervals, based on the time stamp in the file name, and place them into a photo grid layout.

The grid can display a full 36 hours. Since not all studies started at 9:00 AM empty white squares were placed until the start time.

Sometimes a participant turned off the Narrative Clip for privacy, or to sleep. This is represented with gray squares.

If a full face is visible, the photo is replaced with a black square to ensure participant privacy.

The photo grid is saved as SVG format with images embedded as binary data, and placed inside of Illustrator.

The Illustrator file was then placed into the final report, which was built in InDesign.

Result: 1 diagram × 20 participants

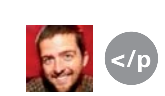
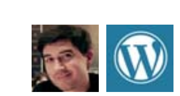
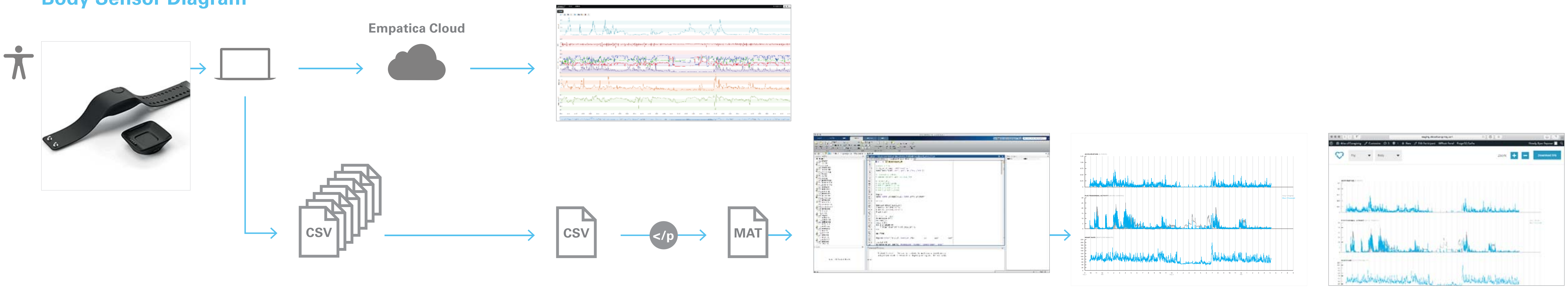


Photo Log (Web)
An SVG is taken from the code output and then uploaded to the appropriate participant directory in WordPress. Website users can zoom in and out of the SVG to view more detail, and can download the SVG to their computer.



Body Sensor Diagram



Empatica E4
The E4 monitors and records data from multiple sensors for blood volume pulse (BVP), heart rate inter-beat interval, motion, temperature, and electrodermal activity.

Sensors and Components:
- Photoplethysmography (BVP)
- EDA electrodes
- Thermometer
- 3-axis accelerometer

Empatica Manager App
When installed on a Mac or PC, the Manager app transfers the raw data, via a USB cable, to the computer and uploads the data to Empatica's cloud.

Raw Data Transfered as CSV Files
TEMP.csv— measured in 4Hz (not used)
Data from temperature sensor expressed degrees on the Celsius (°C) scale.

EDA.csv— measured in 4Hz (plotted)
Data from the electrodermal activity sensor expressed as microsiemens (μS).

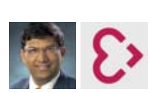
BVP.csv — measured in 64Hz (input)
Data from photoplethysmograph.

ACC.csv— measured in 32Hz (input)
Data from 3-axis accelerometer sensor. The accelerometer is configured to measure acceleration in the range [-2g, 2g].

IBI.csv—calculated from BVP (input)
Time between individuals heart beats extracted from the BVP signal.

HR.csv—calculated from BVP (plotted)
Average heart rate extracted from the BVP signal.

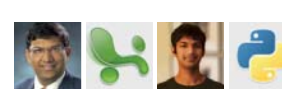
tags.csv (not used)
Event mark times (user created).



Preview
Data from Emaptica E4 is stored in the cloud and is processed and displayed in Empatica's Connect website.

Data Combined into One CSV File
Data from all CSV files were combined using Python scripts.

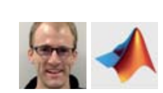
Timestamp— measured in 64Hz
EDA— measured in 4Hz
BVP— measured in 64Hz
HR—calculated every 10 seconds
ACC - X— measured in 32Hz
ACC - Y— measured in 32Hz
ACC - Z— measured in 32Hz
ACC moving average—calculated at 1Hz



Convert to MatLab File
A custom script was created in MatLab to delete unwanted rows and sample all data to once per minute. This process took about one hour per participant. With 20 participants the process took over 20 hours.

Process in MatLab
A custom script was created in MatLab to generate graphs using the data sampled to once per minute. A Ledalab filter was used to normalize EDA data, original data was retained and displayed in gray.

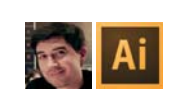
All graphs were programed to output a pixel perfect layout.



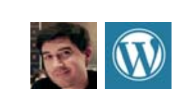
Body Sensor Diagram (Print)
An EPS is exported from MatLab. The EPS file is modified in Adobe Illustrator, removing labels and key—to be replaced with labels consistent with all other diagrams.

Data may be cropped for start and end times but is otherwise untouched. All graph families use the same scale and begin at 0 so as not to distort the data.

Result: 3 diagrams × 19 participants



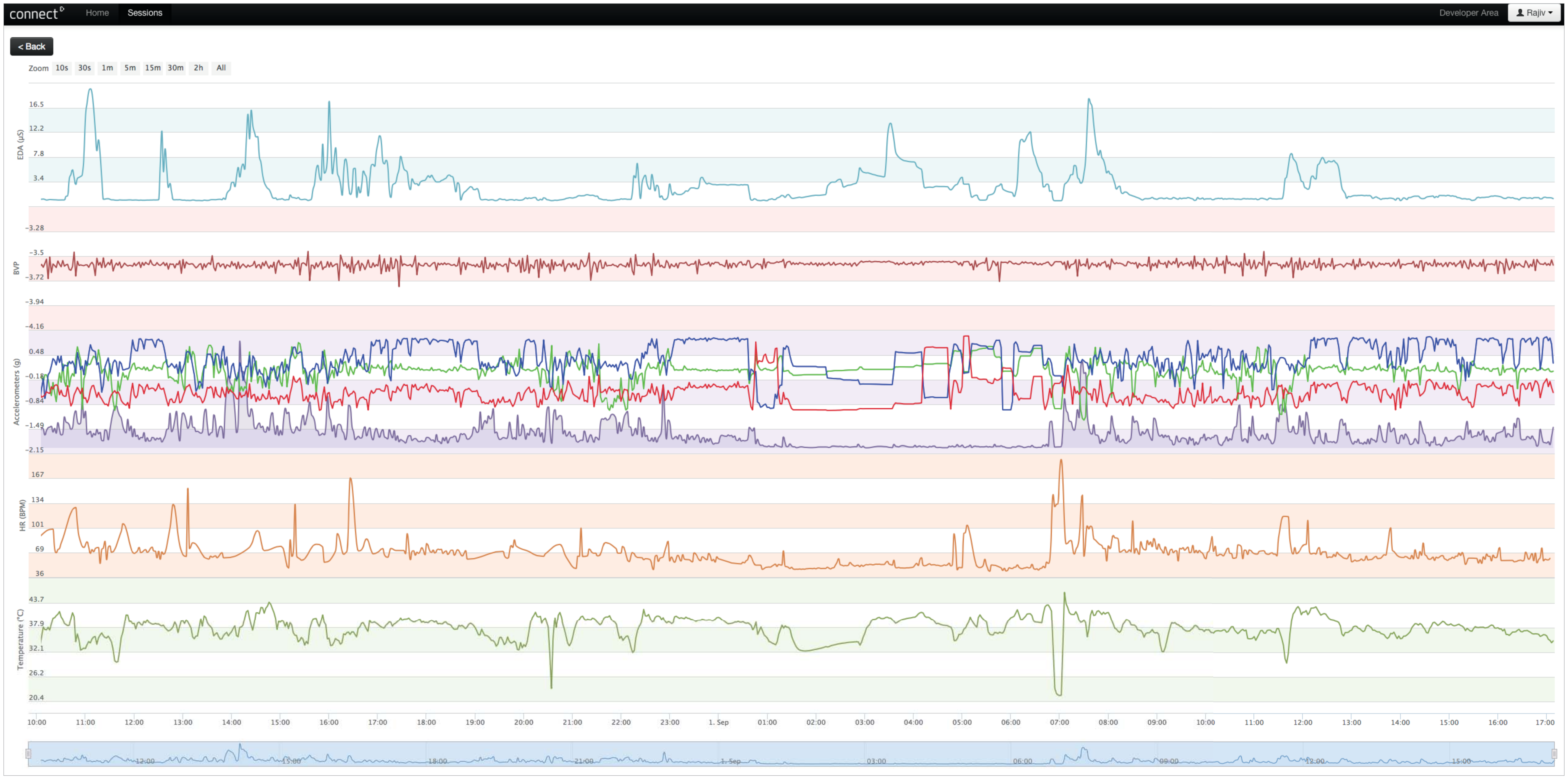
Body Sensor Diagram (Web)
An SVG is saved from Illustrator and uploaded to the appropriate participant directory in WordPress. Website users can zoom in and out of the SVG to view more detail, and can download the SVG to their computer.



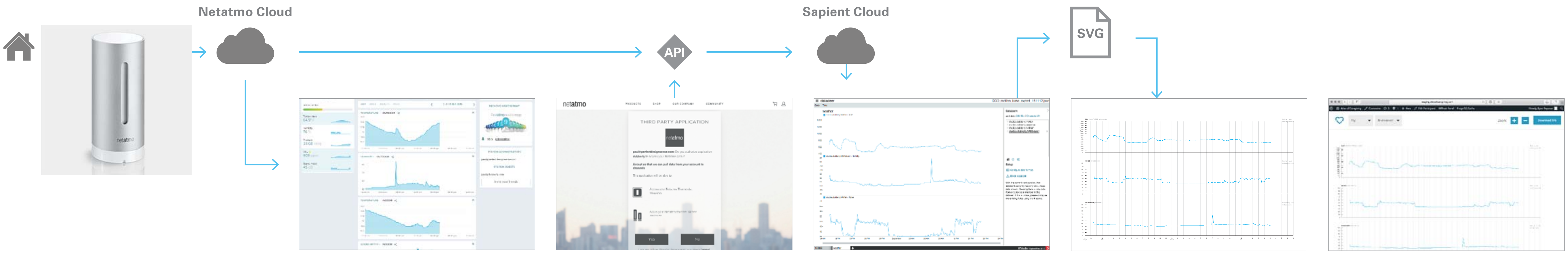
Body Sensor Data Detail

Empatica's Connect website stores and displays all the data in a single chart.

All Empatica Data



Environment Diagram



Netatmo Weather Station
Data from one indoor module, placed in the living room, and one outdoor module, placed in the bathroom, collects data and sends it to Netatmo's cloud.

Sensors and measurements:
- Temperature (indoor):
  Ranges from: 32°F to 112°F
- Temperature (outdoor):
  Ranges from: -40°F to 150°F
- Humidity (indoor and outdoor):
  Ranges from 0 to 100%
- Barometer:
  Ranges from: 260 to 1160 mbar
- CO2 meter (indoor):
  Ranges from: 0 to 5000 ppm
- Sound meter:
  Ranges from: 35 dB to 120 dB

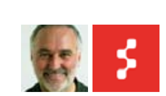
Note: gray text indicates data which was not used in the diagrams.

Preview
Plots of all Weather Station data can be viewed in real-time in Netatmo's web app.

Authorize Sapient API
Using Sapient's connection to Netatmo's API the Netatmo devices are authorized to pass their data to the Sapient Datadeer app.

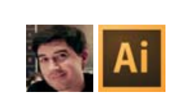
Visualize in Sapient Datadeer App
Data collected into Sapient's cloud storage is imported into Datadeer and temperature and noise data from the main module, and CO² data from the bathroom module is plotted over a 36 hour period.

We used data from these sensors:
Main Module (in living room)
- CO2 meter:
  Ranges from: 0 to 5000 ppm
- Sound meter:
  Ranges from: 35 dB to 120 dB
Outdoor Module (in bathroom)
- Humidity:
  Ranges from: 0 to 100%

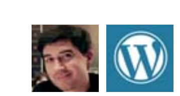


Environment Diagram (Print)
A large format SVG is generated and exported from Datadeer. The SVG file is modified in Adobe Illustrator, removing labels and key—to be replaced with labels consistent with all other diagrams. The data was not altered.

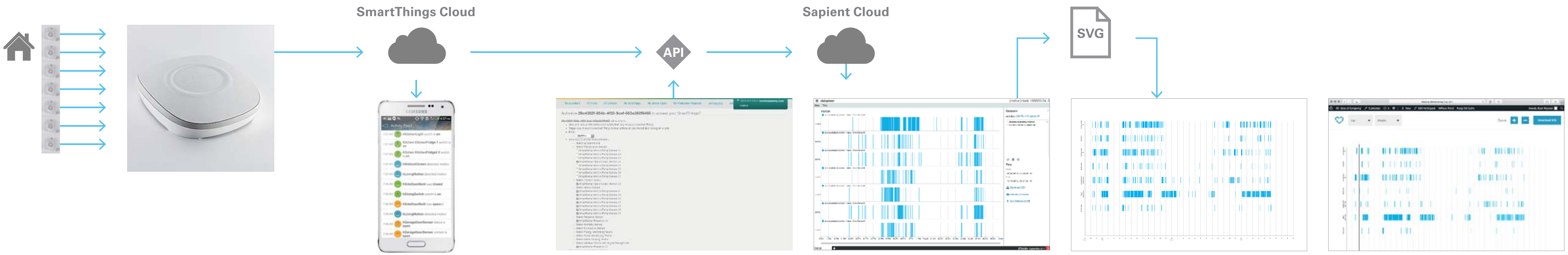
Result: 3 diagrams × 8 households



Environment Diagram (Web)
An SVG is saved from Illustrator and uploaded to the appropriate participant directory in WordPress. Website users can zoom in and out of the SVG to view more detail, and can download the SVG to their computer.



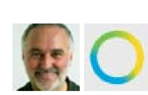
Motion Diagram



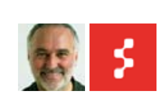
SmartThings Hub
Seven individual motion sensors send data to the SmartThings Hub. The Hub uploads the data to SmartThings cloud.

Preview
Activity from the motion sensors can be viewed in real-time in the SmartThings mobile app.

Authorize Sapient API
Using Sapient's connection to SmartThing's API, the SmartThings devices are authorized to pass their data to the Sapient Datadeer app.

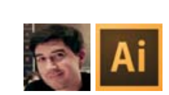

Paul SmartThings

Visualize in Sapient Datadeer App
Data collected into Sapient's cloud storage is imported into Datadeer and activity is plotted from each motion sensor over a 36 hour period.


Paul Datadeer

Motion Diagram (Print)
A large format SVG is generated and exported from Datadeer. The SVG file is modified in Adobe Illustrator, removing labels and key—to be replaced with labels consistent with all other diagrams.

Result: 1 diagram × 7 households


Ryan Illustrator

Motion Diagram (Web)
An SVG is saved from Illustrator and uploaded to the appropriate participant in WordPress. Website users can zoom in and out of the SVG to view more detail, and can download the SVG to their computer.


Ryan Wordpress

Floorplan Diagram

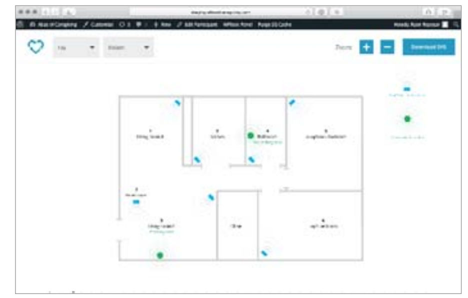
Floorplan Sketch
A hand-drawn sketch of the home, indicating placement of all sensors, is created during the initial interview.

Rajiv

Floorplan Diagram (Print)
The floor plan is redrawn in Adobe Illustrator and saved as an SVG file.

Result: 1 diagram × 7 households

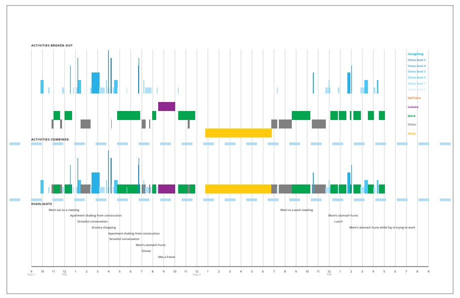
Knut Illustrator



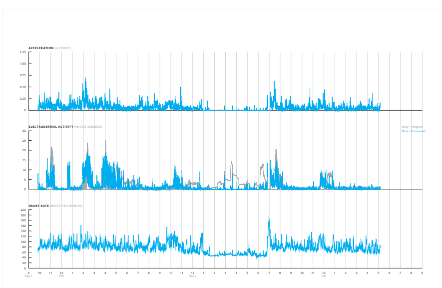
Floorplan Diagram (Web)
An SVG is saved from Illustrator and uploaded to the appropriate participant in WordPress. Website users can zoom in and out of the SVG to view more detail, and can download the SVG to their computer.

Ryan Wordpress

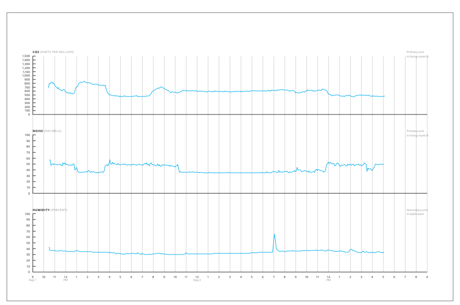
Summary Diagram



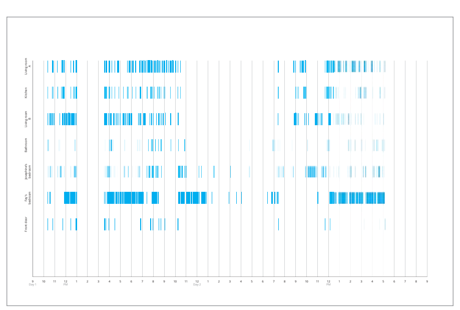
36-hour Log (activities combined)



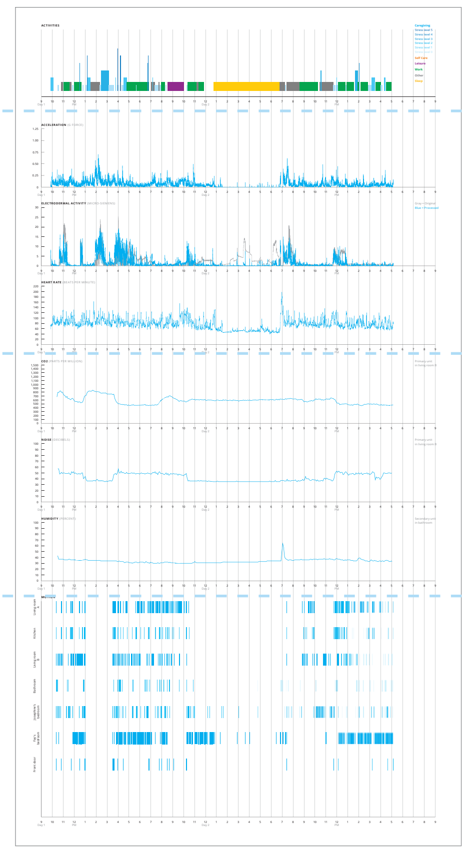
Body



Environment



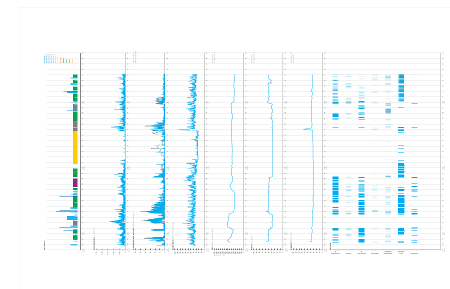
Motion



Summary Diagram (Web)
The 36-hour log (activities combined), body, environment, and motion diagrams are stacked on top of each other in a single diagram to enable easy comparison between sensor data. Valuable insights resulted from these composite diagrams.

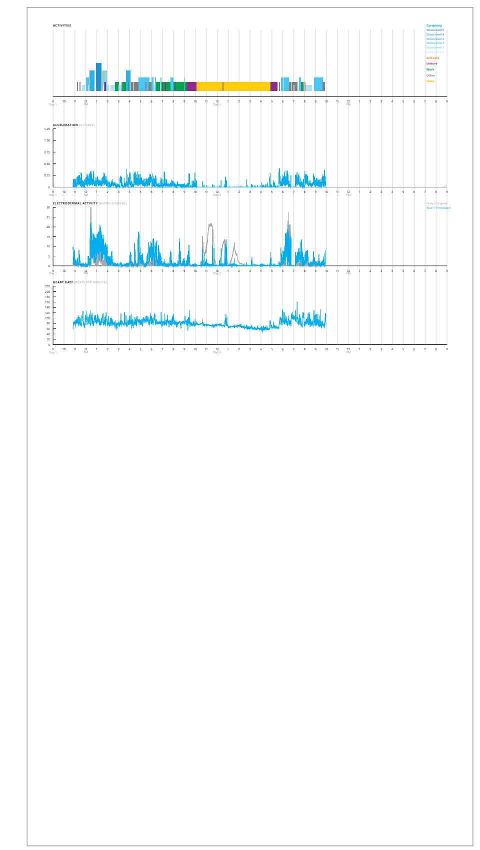
Result: 1 diagram × 18 participants

Ryan Illustrator Wordpress



Summary Diagram (Print)
The tall summary diagram—which is well suited for scrolling on the web—isn't so well suited for our 17-by-11 inch printed book, so we scaled it down by 50% and rotated it 90°.

Ryan Illustrator



Note:
Some participants didn't have data collected for all four diagrams, so their summaries we're shorter. This example shows a participant with just the 36-hour Log and Body sensor data (they didn't have environment or motion sensors in their home).

Thank you

People
Design projects increasingly involve teams.
This project included help from fifteen people.

John Cain
Sapient, Sensor and data expert

Peter Binggeser
Sapient, Sensor and data expert

Aniket Bhatnagar
Sapient, Sensor and data expert

Pasindu Wewegama
Sapient, Sensor and data expert

Elliott Hedman
mPath, Sensor expert

Dawn Nafus
Intel, Ethnography

Rajiviv Mehta
Atlas of Caregiving, Principal Investigator

Shalin Mehta
Coding and production

Hugh Dubberly
Dubberly Design Office, Co-Principal Investigator

Patch Kessler
MatLab wizard

Robin Bahr
Project management

Ryan Reposar
10 weeks, design, wrangling, and production

Paul Souza
6+ weeks, exploration, wrangling, and design.

Cody Wackerman
5 weeks, design and production

Knut Snystad
4 weeks, design and production

Tools
The team used a wide range of tools.

Microsoft Word

Microsoft Excel

Tableau

MatLab

Python

Intel Dataviz Tool

Adobe Illustrator

Adobe InDesign

WordPress

Data Driven Documents (D3)

Automator

Sapient Datadeer