Design, deployment, and algorithmic optimization of zoomorphic, interactive robot companions

Emoters, Inc.

Overview

Pets provide companionship, stress reduction, and even performance enhancement [28]. However, pets cannot be adapted sufficiently to many people's needs, such as frequent travel, living in residences where pets are not permitted, and physical or cognitive disability. Through zoomorphic, nonverbal robots, Emoters seeks to make those benefits more widely available and less costly, including to those who cannot have pets: the disabled elderly, young professionals with little time and much stress, remotely stationed military personnel, and many others.

This proposal seeks funding to design (1) simple yet believably lifelike robots; (2) a puppeteer platform that serves as the robots' eyes, ears, and brains; and (3) a cloud-based system that connects to the platform and permits the rapid improvement of these robots through data collection and experimentation, swiftly increasing their value to users. To achieve these objectives, the Emoters team will apply their expertise in machine learning for human-robot interaction, computer vision, mechatronics engineering, and secure data collection and analytics.

Key Words

human-robot interaction, artificial intelligence, machine learning, learning from demonstration, sequential decision-making, computer vision, cloud robotics, robotics hardware, consumer electronics, electronic toys

Subtopic Name

Electronic Hardware, Robotics and Wireless Technologies (EW) – RH1. Learning, Intelligence and Motion

Intellectual Merit

This Small Business Innovation Research Phase I project proposes (a) to develop first-generation robot pets that will be ready to place in thousands of customers' hands and (b) to situate Emoters to grow to deliver millions of sophisticated robot pets across the world, including to the many people who cannot have pets. Towards these outcomes, the following innovations will be pursued in Phase I: (1) a puppeteer platform that wirelessly controls robot characters; (2) specification of a mobile, social robot character through machine learning; (3) perception of robots and their environment; (4) reliable autonomous recharging; and (5) a simple cloud-based infrastructure for gathering usage data and conducting field experiments on versions of robot characters.

Broader/Commercial Impact

This proposal would have commercial impact in the near-term and over many years of future development. Over both phases, this proposal covers R&D to create a product with the potential to sell millions of units in the U.S. toy industry. The proposal also supports the development of R&D infrastructure that will be a critical component of the expansion in subsequent years to the U.S. pet industry (as a robot companion), which is larger and has less direct competition for robotic entrants. While the R&D of Phase I is conducted, Emoters will separately prepare for a successful crowdfunding campaign to bring these robots to market. Emoters' first-generation robot product, a result of the Phase I project, will support STEM education and robot hobbyists of all ages by facilitating user-friendly modification of its hardware and software as well as the creation of users' own robots and behavioral programs. Given the interactive nature of these modifiable robots, they are likely to have a strong appeal to females, who are underrepresented in STEM fields. The project's ultimate goal is to develop and market interactive robots that can improve the quality of life for anyone through companionship.

Project Description

1 Elevator Pitch

In his 20s, the PI deeply desired his own canine companion. But frequent moves and rented residences kept him from doing so until he was 28 and owned his home. Adopting Azula marked a lasting upswing in his quality of life. When he finished his Ph.D. at UT Austin and took a job as a postdoctoral researcher at MIT, available lodging forbade having dogs. Azula was left in Austin for 6 months—an apparent source of stress and loneliness for both dog and owner—until a suitable dog-friendly apartment became available.

Pets provide companionship, stress reduction, and even performance enhancement [28]. Unfortunately, many people are less fortunate than PI Knox in resolving challenges to pet ownership. In the United States, where the pet industry has $60 billion per year in revenue [8], 41 million US households do not have pets (35% of US households) [11]. These non-owners include the disabled elderly, those allergic to pets, frequent travelers, children whose parents are hesitant to hand them a live animal, people with limited time for caretaking, stationed military personnel, and those in residences where pets are not allowed. With aging populations in developing nations and increasing urban density, these pet-preventing limitations will worsen.

Emoters believes small, life-like, interactive robots—built to be nonverbal and more zoomorphic than anthropomorphic—will address these needs, bestowing the rewards of non-human companiionship on many who cannot own an organic pet, as well as those who do but may also benefit from further companionship. Current offerings of interactive robots are limited by their cost, agency, inorganic and repititve behavior, and the high level of human intervention required to maintain their active presence. Many such ”robots” are simply toys designed for children, with the trappings of a pet—such as obeying commands and playing fetch—but lacking the ability to create long-term relationships with typical adults. When robots *can* deliver value in rich, authentic social relationships, hundreds of millions of people will benefit greatly.

Across two phases, funding from NSF would support both Emoters' initial development of a highly successful consumer product and create the R&D infrastructure to conduct human-robot interaction field experiments at great scale and efficiency. With this infrastructure, Emoters will algorithmically optimize towards interactive, zoomorphic robots of increasing value to their users. This proposal details a unique strategy for leveraging the commercialization of life-like robots for their further development. Emoters has the technical expertise to complete this proposal: human-robot interaction, machine learning, artificial intelligence, computer vision, mechatronics design for manufacturing, and secure data collection and analytics. Additionally, this project will build upon the PI's NSF-funded research on machine learning for human-robot interaction.

The initial consumer product consists of interactive robot characters and a cloud-connected platform that acts as the eyes, ears, and brains of the robots, acting wirelessly as their puppeteer. Preliminary robot and "puppeteer platform" designs can be seen in Figures 1 and 2, respectively. **The following innovations are expected by the end of Phase II** (see Phase I objectives in Section 5.2):

1. Design of a *puppeteer platform* that wirelessly controls robot characters and enables them to be both *comparably low-cost and intelligent*.
2. *Specification of a mobile, social robot character through machine learning from demonstration data created during teleoperated interaction*, capturing organic and playful styles.
3. *Robust perception* of robot poses, poses of objects of interest, human activity (gestures, utterances, etc.), and generic obstacles.
4. *Reliable autonomous recharging*, permitting months of character persistence without user intervention.
5. *Cloud-based infrastructure for gathering usage data and conducting field experiments on versions of robot characters*.
6. *Algorithms for large-scale optimization of robot characters*.

With the support of NSF, Emoters expects that an initial consumer product would be available before the end of Phase II. This product would not only bring immediate value to users but would be a major accomplishment in long-term social human-robot interaction, situating Emoters to make swift progress in delivering increasingly valuable robot companions to the public.

Figure 1: *Candidate designs for first-generation non-verbal robot characters.*

2 The Commercial Opportunity

The first generation of Emoters' robots will compete within the toy industry, but the product will subsequently be developed—through the R&D infrastructure that is an innovation of this proposal—to be marketed as a robot that provides legitimate non-verbal companionship, akin to but distinct from existing pets.

Pet animals provide immense value to many people. But the ability to adapt them to fit people's needs is limited, leaving many without pets or with less pets than they would prefer. Common reasons for not owning pets include cost of ownership, cleaning up after the pets, lack of time, potential damage to property, frequent travel, allergies, and that pets are not allowed in their residence. Additionally, the likelihood of considering the adoption of a pet reduces with age [7], reflecting the exacerbation of the above problems by age. Life-like interactive robots provide a potential solution to all of these issues, since they lack common allergens, rarely damage people's property, and can be built to fit various needs. They can be designed to travel well, be affordable, and to require levels of caretaking that scale with their purchaser's desire and availability.

2.1 Markets sizes and customers

The U.S. pet market is predicted to be $60 billion in 2015 [8], three times as large as the $21 billion U.S. toy market [4]. Emoters estimates the worldwide pet market to be more than $110 billion, which would grow to $200 billion if robot pets were readily available and provided as much value as existing pets. This estimate is based on the proportion of households without pets, average pet expenditures per household, reasons for non-ownership, and other relevant data.

Figure 2: *An early conceptual design of the Emoters puppeteer platform. The platform performs most of the perception and decision-making tasks for the robots, wirelessly controlling them at a low level (e.g., sending motor velocity commands). The first-generation version will likely not have the platform-specific surface and will use metal contacts for autonomous robot recharging. See Section 5.1.3 for further description.*

Emoters will release initial products in multiple generations to different target markets. The first generation will target early adopters in the U.S. who want a *hackable* robot creature. This group includes robot hobbyists, students and educators focused on STEM, robotics researchers, software developers, and other technically adept people. An estimate of this serviceable addressable market (SAM) is 1 million users.

The second generation will target young professionals in the United States. This group includes many who live in places that do not allow pets, may not have enough time for an organic pet, or travel too often to care for a pet. This tech-savvy target group is also relatively similar in tastes and lifestyle to early adopters of technology, making this group an appealing bridge from robot hobbyists to other large market groups. This group is estimated to yield a serviceable addressable market of 11 million users.

Subsequent generations will expand to accommodate different market groups. Groups of particular interest are: the elderly who could benefit from extra companionship without the financial or caretaking requirements of a pet and families with children with that (a) exhibit mild to strong interest in robotics and understanding how robotic systems inter-operate, (b) exhibit social disabilities that may benefit from increased interactions and exposure to unique environments, and (c) want a pet but have never cared for one before. Additional plans include marketing the interactive robots globally. An estimate for the serviceable addressable global market of would-be pet owners is more than 400 million users in developed countries alone.

2.2 Market validation

This section discusses validation of interactive robots and digital creatures, as well as validation of the specific product Emoters is designing, initially marketed as a robotic toy. Evidence suggests that people will buy these products and derive personal value from them. Further, the features of Emoters' robots will promote widespread and long-term usage and an illusion of life at a level that currently available products do not (see market analysis presented in Figure 1).

2.2.1 Validation of our initial product, from user testing

Emoters researchers have conducted user testing with over 50 people, with each session including approximately an hour of robot interaction and subsequent questions. Of those, the 35 most recent are almost all young professional adults. These participants interacted with an early version of the robot that is secretly teleoperated in a so-called Wizard-of-Oz scenario [23]. This robot is a prototype and not particularly attractive, having been designed before our team included professional designers and fabricated with an inexpensive 3D printer. Nonetheless, after interaction with the robot, **35% said they would purchase it and its platform as-is** for $100 or more. Given the prospect of an improved robot with technologically-feasible features added, **82% of these participants said they would purchase the improved version** for $100 or more. A montage of human-robot interaction during these sessions can be viewed at `https://youtu.be/ArecH7uwwgw`. The robots' transition from initial prototype to this feature-rich robot pet is an important part of the innovation in this proposal, and is discussed in Sections 3 and 5.

2.2.2 Validation of initial product, from comparison

Robot toys in general are a growing sector of the toy industry—a sector that is getting credit for improvements to the industry's overall revenue [21]. The BB-8 robot, by startup Sphero, is a remote-controlled character from the 2015 Star Wars film. At a similar $150 price point to our planned starter package of robot and puppeteer platform—the BB-8 was predicted to sell one million units this past holiday season [9].

Pet-like toys, both robotic and virtual, have been highly successful. Tamagotchi and Furby are probably the two best-known offerings. More than 79 million Tamagotchi units have been sold ($900M in gross sales), and Hasbro's Furby has been a top-selling toy twice: after their initial release in 1999 and after their 2012 re-release. The first generation Furby sold 32 million units in the first 3 years [1]. Its spin-off, Furby Boom, was also named the best-selling toy of 2013 in the UK [19], with a new edition of Furby also in the top 10.

After Tamagotchi's release 19 years ago, virtual pets have continued to be a successful sector of the toy and gaming industry. Released in 2005, Nintendogs sold more than 24 million units and has been the second best selling series for the popular Nintendo DS. In 2012, online virtual pet ecosystem Moshi Monsters had 70 million registered users, $73 million in revenue, and $16 million in profit. At the time of writing, according to ThinkGaming.com more than 5% of the top 300 highest grossing iOS apps involve virtual pets [2]. People who raise virtual pets have a particularly strong emotional attachment. For example, when usage of Facebook games had declined, Electronic Arts shut down several of its Facebook games, including Pet Society, which was created by a company EA had acquired. At its closure, Pet Society had 500,000 users who played at least once a day. Users protested and organized a boycott of the company, eventually succeeding at securing a non-commercial license to operate the game. One protester at the time said, "Shutting down Pet Society is like someone coming into our house and shooting one of our real cats. But wasn't that the point of Pet Society, to get you emotionally invested?"

In comparison to virtual pets, life-like interactive robots are poised to have even greater impact. A series of studies examined the different effects of interacting with a virtual character versus a physical character (that is largely identical beyond its embodiment); these studies consistently found that the physical character resulted in more enjoyment and in the human treating the character more like a social entity (e.g. improved cooperation and turn-taking, more interaction, and judgments of the character being credible and perceptive) [13, 27, 18, 10].

2.2.3 Validation of robotic pet companions

Despite the fact that robots as pets may seem far-fetched to casual observers of human-robot interaction, there already exists surprising validation of their potential to create attachment. For example, some users who had especially strong bonds to their robots, and funerals were held in Japan to mourn the passing of these "family pets" [24]. These robots were 10–16 years old.

Product	Weakness vs. Emoters	Features supporting long-term use & illusion of life	IoT (reporting & updates)	<$150	Market includes cognitively normal adults
Emoters' robots	-	++++++	+	+	+
virtual pets	no physical interaction	+++	+	+	+
Spin Master Zoomer	children's toy	++	-	+	-
WowWee CHIP	children's toy	+++	-	-	-
Sony AIBO	discontinued	++	-	-	+
Innvo Labs Pleo	innovation halted & $350	++	-	-	+
Paro	nursing home only & $6000	++	-	-	-
Hasbro Joy For All	nursing home only & shallow experience	+	-	+	-

Table 1: Competing products and comparison to Emoters' proposed product. The features column includes a count of the following features that encourage long-term usage and illusion of life: sets manageable expectations by morphology and behavior, shares the physical world with the user, exhibits lifelike speed, has its own interests and is active when user is away, and persists without user effort (e.g., through autonomous recharging), and personality directly emulates a live human puppeteer. Partial fulfillment of a feature is counted as a half increment, and counts are rounded to the nearest whole number. This feature count, though subjective, nonetheless provides valuable information for the evaluation of this proposal.

2.3 Business model

The robots will operate in conjunction with a *puppeteer platform*, discussed in detail in Section 5.1.3. The platform performs perception (including computer vision) and decision-making computation for the robots, and therefore the robots are simply unusually feature-rich remote-control toys controlled by the puppeteer platform, not a person. The platform is central to Emoters' monetization strategy, which early on focuses on hardware sales:

- **Starter package** – Consists of the puppeteer platform and a single robot. (Estimated price: $120)
- **Additional robots** – For users who own the platform already, robots can be inexpensively added. Robots will be designed to be more entertaining as a group and to be diverse enough to be considered collectible. (Estimated price per robot: $45)
- **Data-driven accessories** – With data collection, accessories are designed in response to strong data-driven demand. For example, if many users decorate their robots with hats, through videos that users share on social media through an Emoters app, data would indicate this trend, and hats might then be developed for purchase.

Emoters will also explore the potential for software-based revenue. Such revenue might come from subscriptions to software updates, "in-app" purchases (e.g., of new robot skills) for small amounts of money, and licensing the puppetry platform to give autonomous personality to third-party robots and RC toys.

2.4 Competition

Emoters' competition includes companies that make virtual pets, expensive robot pets, pet-like robot toys, and remote-control robots. In addition to products already mentioned in this section, notable offerings include Hasbro's Joy of Life line of robot pets for seniors and WowWee's robotic dog CHiP. Both were announced in November 2015, signaling that robotic pets are gaining increasing attention from the toy industry. Although competition is expanding and could come from several directions, Emoters' robots are highly differentiated from existing and announced options. The market for robot companionship can support multiple successful contenders, given the distinct and varying tastes of people for both pets and electronic characters. Emoters also has a team with deep experience in technical areas needed to execute this strategy: robotics, electrical and mechanical engineering, computer vision, artificial intelligence, and machine learning. Most robotic toy companies lack some of these skills. Lastly,, our product design has several critical advantages over all known competitors, detailed below and summarized in Figure 1.

1. Unlike Emoters' robots, none of the competitive robotic offerings are socially interactive, autonomous, *and* cloud-connected. Therefore, none are set up for the large-scale character optimization that we propose. Instead they will be slowed by standard and soon-outdated development practices.
2. Emoters' robots will find their recharging base with near-perfect consistency. Consequently, our robots can "rest" and wake themselves for months without any user interaction, permitting a much wider range of user interaction levels, which on the low end is unlike what electronic toys have previously allowed.
3. Competitive offerings have limited interaction with the physical world. Those those that do interact with it either see all objects as undifferentiated obstacles or, at best, have a small and predetermined set of

objects they are programmed to recognize and interact with. The puppeteer platform will use computer vision for object detection, employ a set of models that can be quite large and customized for each user, allowing differentiated interaction between the robots and various objects.

4. Most other offerings are not autonomously mobile and therefore have limited agency—an important component of seeming to have a mind [12]. That is, they need to wait for human interaction to do anything interesting. Emoters robots conversely will act upon their environment to fulfill their own goals.
5. Much of the competition's robots set high and unobtainable expectations by resembling dogs or cats. Emoters' robots will be designed to be abstract and avoid strong association with any specific organic animal to limit initial expectations on the behavior and intelligence of the robot.
6. Emoters' platform model reduces cost considerably, enabling our robots to arguably be the most intelligent option in the under-$500 price range.

2.5 Market risks

Beyond the typical risks of a hardware startup (execution speed, funding, branding/marketing, building something that people want, etc.), several points pose risks to Emoters' success.

Currently, a stigma exists against adults establishing authentic companionship with a machine. Nonetheless, early adopters will experience and profess the benefits of interactive robots, eventually leading to mainstream openness. Similarly, popular media like the movie *Her* have powerfully placed audiences in scenarios in which they viscerally envision the potential for human-machine companionship.

For the initial product, one assumption is that the robots will be able to sustain long-term interaction with people. Many design factors support this assumption, including autonomous recharging, the expectation-setting simplicity of the robots, and the ability to optimize their behavior as they are used by consumers. Another assumption is adequately addressing people's privacy concerns with having an internet-connected camera and microphone always on in their homes, offices, or wherever they place the robots' platform. Smartphones and connected security cameras already set precedent that will help assuage fears, but the system will be designed to gather only data that has an explicit purpose, to protect users' data, and to allow users to view and selectively limit what data is gathered. However, gathering *some* usage data is integral to the business strategy, and it will design privacy options to balance business needs with those of users. See Sections 5.1.3 and 5.1.6 for technical discussions of privacy.

With these market conditions in sight, Emoters has engaged a strategic set of mentors and equity-based advisors, described in Section 4, bridging all stages of development and relevant product markets.

2.6 Commercialization strategy

As discussed in Section 2.1, Emoters plans to develop and release its initial product in stages according to the tentative timeline below.

Stage I: Robot hobbyists version This version will be **sold through crowdfunding** and ship with an autonomous robot creature (not marketed as a "pet") that is goal-driven and interactive. The robot hardware, the perceptual software, and the platform-to-wireless communication protocol will be released open source and may include support for the child-friendly programming language Scratch [22]. Roughly six months after SBIR funding, investment for manufacturing will be sought via crowdfunding, with first versions ready to ship around one year from SBIR funding. Based on comparable crowdfunding campaigns and heuristics for the number of pre-orders to expect after a campaign, sales of 2500 starter kits and 1250 additional robots are forecasted, yielding revenue of $430,150 and a margin of $146,035. Crowdfunding has been chosen for the following reasons: to reach early-adopting customers who purchase robotic and STEM-education products, a highly present demographic in crowdrunding; to facilitate a high-profile launch; because of the large quantity of online, publicly available advice and tactics for crowdfunding; and because Emoters' mentors include numerous people who have run unusually successful crowdfunding campaigns (see Section 4). **As the R&D efforts of this proposal are conducted, Emoters will also undergo crowdfunding preparation.** This preparation will include building an email list and Facebook page (which already together number 1000 people); hiring a PR consultant, crafting our media narratives, designing a press-friendly product demo, and reaching out to select journalists; developing a compelling video; testing the unit economics of online ads and securing lines of credit for ad spend; and engaging with niche influencers who would promote our product. Emoters plans to ship the robots and puppeteer platforms of this stage approximately 7 months after the end of Phase I. The outcome of this stage will be the ability to observe interactions with the robots, getting metrics from thousands of users. Emoters also aims to support open-source contributions from the community, which will amplify its own development efforts.

Stage II: Young professionals version This version will ship with an autonomous robot that has improved capabilities for interpreting human activity, emotionally expressive behavior, and a unique personality that develops over time. Phase II funding would support the extensive R&D needed to create these improvements. It would launch in February 2018 for pre-order and ship in August 2018, at which time we would begin selling units direct and through online retail. For 2018, estimated sales of 23,000 starter kits and 28,750 additional robots yield revenue of $4,693,513 and a margin of $2,351,859. Approximately $4,000,000 in additional cash will be needed to manufacture projected units without pre-orders, which will be sought through a Series A round of investment, raised based on the traction shown through the crowdfunding campaign and usage data from shipped units. With this timeline, profitability is reached in late 2018. This stage will continue in 2019, with increased marketing expense and expansion into boutique brick-and-mortar stores like Brookstone, yielding forecasted revenue of $22,767,324. The aforementioned figures and timeline are a projection; many other paths to profitability exist, and the investment predicted affects the speed and extent of Emoters' success but is not necessary for Emoters to grow and advance the vision of this proposal.

Stage III: Optimizing companionship From 2020 onwards, Emoters' will continue to improve the robots, adding empathy through emotional inference algorithms and optimizing towards companionship. Product lines will also diversify. For example, the cloud-based API may be able to subsume the functions of the puppeteer platform, permitting the robots to roam free. In this stage, Emoters will raise investment to fund rapid expansion of its market reach and the quality-of-life improvement the robots grant their users.

3 The Innovation

Emoters was founded to introduce interactive robots to the marketplace. One goal is to provide a high level of intelligence and personality on a low-cost robot, which can be accomplished by off-loading the robot's perception and intelligence onto a computer embedded in the robot's environment. Another aim is to build "believably alive" robots that develop individual relationships with their users, achieving a suspension of users' disbelief through months or even years of interaction. To bridge from the toy robot market to selling robots as pets, the product described above will be developed steadily towards robots that provide value on par with dogs and cats. To achieve this second component, Emotes will conduct algorithmic optimization across the tens of thousands of robots already sold, at a (experimental) scale and efficiency that no current organization can achieve. Emoters believes the question of how to create high-value interactive robots will be best solved empirically, and this two-part strategy situates Emoters to be multiple steps ahead of competitors.

3.1 Description of the innovation

Emoters seeks NSF SBIR funding to transition human-robot interaction (HRI) research into a low-cost robot pet system. To create the system, numerous existing technologies will be integrated: some that exist and others that need to be developed and thus constitute innovations on their own. These innovative components are elaborated upon in Section 5, and pertain either to human-robot interaction or to specific technical challenges that improve robotics in the home. Put briefly, the technical challenges include perception, machine-learning-based character design, building the cloud-based infrastructure for gathering data and testing different versions of the robot software, and algorithms for optimization of the robot characters based on the results of that testing.

3.2 Stage of development

In terms of the vision for this proposal, none of the proposed innovations have been fully achieved; however, considerable progress has been made towards the first generation of the product.

Emoters has conducted multiple iterations on teleoperated prototypes of the platform-and-robot system. The robots are currently controlled by the puppeteer platform's radio-frequency messages containing pulse-width-modulation values for the motors, pitch sequences for the robots to convert into square-wave nonverbal utterances, or color and intensity sequences for LEDs inside the robot's semi-translucent body. Figure 3 shows the internal electromechanical components and chassis of the most recent version of this robot.

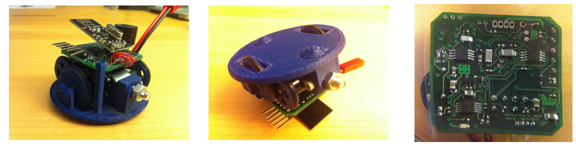

Figure 3: *Electromechanical components and chassis of the prototype robot currently employed for user testing.*

Emoters has implemented a version of the puppeteer platform on two types of computers, linux desktops and Raspberry Pi, which will likely be used as the in-home platform computer for the first generation product. Software components on the platform are interoperable on both computer types and run as nodes in the Robot Operating System (ROS) [3]. Because secret teleoperation (i.e., Wizard of Oz) is a critical part of Emoters' strategy for designing robot behavior (discussed in detail in Section 5.1.5), teleoperation through a gamepad and a browser-based interface was implemented. The browser-based interface will allow teleoperation of robots outside of the local network, and some high-level decision making may be implemented on it, which could form the beginnings of the cloud-based component of the system. Elements of perception and autonomous decision-making have been implemented and/or integrated as well. These include a fiducial marker library and object recognition test harnesses, an early version of a vision-based hand tracker, and autonomous path planning and following. As described in Section 3.4, modifications of open source projects have been fed back into the relevant community.

Emoters has hosted more than 50 potential users for in-office robot testing so far. The testing focuses on robot designs (including mechanical and audio aspects), the proposed feature set, how we describe the product and its usage, what objects are provided in its environment, and the teleoperation style and strategy.

In March 2016, Emoters hired **Pushstart Creative**, an Austin-based design firm focused on Internet-of-Things and software consumer products, for a comprehensive first-run design strategy. Emoters is partnering with Pushstart to determine its branding, positioning, and design theme, as well as specific designs and high-fidelity prototypes of the robots and the puppetry platform. Guiding this design is a combination of feedback from user testing, anonymized market testing (with demographics), and engineering viability sessions with Emoters. This design will be the first investor- and customer-visible product that Emoters uses for production.

3.3 Key technical challenges and risks

The proposed system is technically complex, consisting of a three-layer robot perception and behavior architecture (robot, puppeteer platform, and cloud), as well as an infrastructure for experimentation and data-collection. That complexity poses a technical risk by itself. Manufacturing, if counted as a technical challenge, is also a risk at which many hardware startups falter. This proposal includes several technical innovations that pose substantial challenges (all described in detail in Section 5).

1. Design of a *puppeteer platform* that wirelessly controls robot characters and enables them to be both *comparably low-cost and intelligent*. In Phase I, the platform will be comprised of an off-the-shelf single-board computer, a wide-angle video camera, a microphone, and a robot charging base. The computer, camera, and microphone will likely be off the shelf. In Phase II, these components will be designed by Emoters to reduce the platform's cost of goods sold.
2. *Specification of a mobile, social robot character through machine learning from demonstration data created during teleoperated interaction, capturing an organic and playful style.* The PI's NSF-funded research pioneered this approach (discussed in Section 5.1.5), in which supervised learning is employed to create a model of the operator's socially interactive *demonstrations* [15, 6]. This model is then used in lieu of the teleoperator. However, prior work will need to be adapted in Phase I to this domain. In Phase II, we will design algorithms that can learn behavior from both teleoperation data and each robot's individual experience, so that robot personalities and capabilities reflect their experience.
3. *Robust perception of robot poses, poses of objects of interest, human activity (gestures, utterances, etc.), and generic obstacles.* In Phase I, the tracking of robots, a few objects of interest, and generic obstacles will be completed. Limited support for inference of human activity, such as detecting a vocal utterance (but not necessarily its content) will also be completed. Phase II will focus on improving human activity detection and expanding the accuracy and diversity of object recognition.
4. *Reliable autonomous recharging permits months of character persistence without user intervention.*

This challenge will be completed in Phase I and validated in testing with potential users. By Phase II, Emoters will have shipped to customers, who will provide more exhaustive stress testing and may reveal unforeseen challenges.

5. *Cloud-based infrastructure for gathering real-time usage data and conducting field experiments on versions of robot characters.* In Phase I, a basic version of this infrastructure will be completed that allows small-scale, manually designed experimentation. Phase II will address the challenges of scaling to large numbers of users.

Emoters is particularly well-suited to overcome the risks of the above innovations—already containing the in-house expertise needed for Phase I (see Section 4)—and fulfill its vision to deliver robot pet companions.

3.4 Intellectual property and defensibility

In September 2015, Emoters filed a provisional patent application that covered system architecture, hardware design for the robots and puppeteer platform, cloud connectivity, and various plans for algorithmic solutions to computer vision and decision-making. Emoters is now consulting with patent law firms that focus on hardware startups to determine what to patent from the provisional application and more recent work. Importantly, if MIT decides to patent existing IP from PI Knox's prior employment (a foundation for the automatic learning of robotic characters), a royalty-free license will be granted to Emoters for Knox's role as a creating author.

Emoters has already and will continue its policy of contributing back to open source projects like ROS [26] camera drivers and Small Size League Vision (SSLVision [29]) wherever possible. Beyond existing libraries, some components of the Emoters system will be open sourced for the first generation product.

In the current IP landscape, patents are not considered an effective tool for startups to prevent copying by large competitors, since startups lack the cash to entertain expensive lawsuits. Instead, *defense from competitor copycatting comes from branding, an exceptional team that can execute quickly, and a value proposition that is reliant on proprietary technology that is challenging and time-consuming to reverse engineer*. Most of the core innovations of this proposal will be proprietary: the cloud infrastructure, software modules that execute high-level and mid-level decision making, and software for *development* of the robot characters. These components will not be exposed to customers or competitors, making them particularly defensible. Nonetheless, we will pursue patents as our budget permits, to increase Emoters' value and to provide leverage in unlikely case a competitor sues Emoters. (As one data point, 0 of approximately 100 clients of Schox Patent Group were sued by competitors over the firm's more than 10 years of operation.)

3.5 NSF lineage

Emoters proposes to design robot behavior by applying machine learning to logs from secret teleoperated interaction sessions (described in detail in Section 5.1.5). This method will be built upon PI Knox's postdoctoral research at MIT from 2012–2014 under Professor Cynthia Breazeal, who is also the co-founder of VC-backed personal-robot startup Jibo. This research was funded by the NSF Expeditions in Computing *Grant Socially Assistive Robotics* (Award Number: IIS-1138986). Emoters would be granted a free, perpetual license to use this patent-pending technique.

This project is also related, though indirectly, to the PI's doctoral dissertation research, entitled *Learning from Human-Generated Reward* [14], which sought to answer how the approach of reinforcement learning needs to be adapted to be able to learn from live positive and negative feedback from technically naïve human trainers. This research could be informally described as animal training algorithms for robots. There are no immediate plans to apply this research to an Emoters product, but opportunities may be identified. PI Knox's doctoral research was supported by an NSF Graduate Research Fellowship from 2008–2011.

Additionally, Emoters' teleoperation system is central to its strategy for robot character development. This system is built upon Robot Web Tools for ROS [26], the origin of which was supported by NSF (Award Numbers: IIS-1149876, IIS-1208497, and IIS-0844486).

4 The Company/Team

The key personnel for this project include the following members of Emoters.

- **William Knox, Ph.D. (PI)** – Dr. Knox has been working full-time on Emoters since late January 2015 and would be able to focus half of his time on this project, spending the other half on business-related activities for Emoters. Dr. Knox's relevant research background includes robotics, machine learning, artificial intelligence, psychology, and experimental methodology. His NSF-funded research has won best paper awards at relevant conferences [16, 17], he was awarded the award for best dissertation from UT Austin's Department of Computer Science in 2012, and he was named to IEEE *Intelligent System*'s AI's 10 to Watch in 2013, compiled every two years. He conducted post-doctoral research at MIT's Media Lab with Prof. Cynthia Breazeal (of personal robot startup Jibo), where he developed and tested an algorithm for learning socially interactive robot behavior from teleoperated control. Dr. Knox will oversee this project, conduct user testing, create the algorithms for character design from teleoperation and cloud-based character optimization, and contribute to the software engineering effort.
- **Eric Zavesky, Ph.D.** – Dr. Zavesky has contributed approximately 25 hours of effort per week to Emoters since May 2015. He will be responsible for the design and implementation of computer vision and tracking algorithms as well as the extensive systems integration. In 2009, Dr. Zavesky received a Ph.D. in Electrical Engineering from Columbia University, focusing on computer vision and human-computer interaction, where he was later Adjunct Faculty. Dr. Zavesky's industrial and academic experience is broad: 12+ years in computer vision and video processing, 8+ years in audio processing, human-computer interaction, and machine learning, and more than 20 years in general software development. He has 13 issued and more than 34 pending patents in the video and telecommunications fields.

  [content redacted]

These personnel provide the expertise to accomplish Phase I objectives (see Section 5.2. Phase II will involve further challenges, such as moving from a simple hosted database of usage data to a scaleable cloud-based system, and personnel will be expanded as required to meet these challenges.

Advisors and mentors A number of advisors and mentors are part of the company's strategic decision making process. Key parties who provide frequent consultation are listed below, with **bold** typeface indicating signed letters of support for this SBIR proposal.

- **Pushstart Creative** - Austin-based IoT and software product design firm and *partner for first-generation robots*

  [content redacted]

Other relevant mentors, with whom Emoters has repeatedly consulted, are listed here: *[content redacted]*.

Company vision Emoters will grow to employ hundreds of people. Emoters' products will provide delight, companionship, and institutional know-how for effectively developing social behavior in robots. Emoters' five-year vision involves having made significant steps towards this goal, with 30 to 100 employees and 100,000 to 500,000 customers. The robots would be developed to the point that our customers would be deriving many of the benefits of organic pets, including companionship and reduction in stress [28]. By this time, Emoters would have developed appropriate APIs for third-party use (or deployment) to study the social benefits of our robots' interactions with users. Additionally, evaluations of the benefits of STEM-based interfaces (e.g. Scratch) and therapeutic applications (e.g., in elderly homes, children's hospitals, and military posts) may be pursued. More detailed information about our timeline over the following five years can be found in Section 2.6.

Existing operations and revenue history All efforts at Emoters are preparing for this project, the fruits of which the company will be based upon for years afterwards. However, funding will allow Emoters to hire the project's key personnel full time, which is critical to speed progress sufficiently to develop and market this product and begin to make an impact on people's lives. Emoters has no revenue history.

5 Technical Discussion and R&D Plan

5.1 Detailed description of the proposed project and innovations

Here we describe the proposed project, detailing its important components and how they function as together as a system, with emphasis on the points of innovation.

5.1.1 System architecture

Figure 4 shows the overall system architecture, which can be divided into two functions: a robot behavior system (Section 5.1.5) and a system for data collection and experimentation (Section 5.1.6) that modifies the robot behavior system. The robot behavior system is itself a three-layer architecture: quick and reactive control on the robot itself, perception and high- and mid-level decision-making on the puppeteer platform, and cloud-based function calls that the platform can make for heavy computation.

5.1.2 Robot design

This section describes the design of the robot characters from two perspectives, features desired for successful human-robot interaction and the electromechanical design to implement those features.

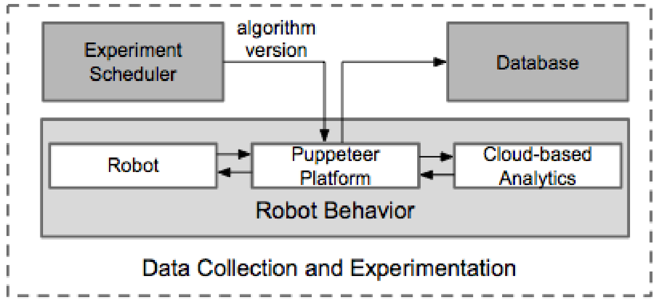

Figure 4: High-level system architecture.

Human-robot interaction The visual aesthetics, tactile feel, and auditory output of the robot are each critical to the user's experience. This section discusses specific design considerations below with Emoters' philosophy, which is combined with feedback from user testing to guide the robot's design.

Visually, we are focusing on both the robot's form and movement. The form is being designed to be small, abstract, sleek, cute, and not easily associated with any specific organic animal. Figure 1 shows some candidate designs at the time of writing. It is expected that the robot's small size to create associations with simpler categories of animals, and it will also make the surface the robot lives upon larger proportionally to its body, increasing the space of possible behavior. To increase the robot's expressivity and illusion of life, we are design the robot's movement to be highly variable in speed and to be relatively precise at high velocity. The robot will also have expressivity through control of the color and intensity of LED lighting, which will illuminate a selected part of the robot's body.

The exterior body of the robots will be made sufficiently rugged to withstand long-term usage and moderate physical interactions. Its outer layer, however, will be designed to invite and reward touch, possibly through very short fur, silicone, or soft-touch paint. The robot will also have on-board touch sensing.

The robot will create sound voluntarily through nonverbal utterances and involuntarily, mostly through motor noise. Like the small size, the utterances are nonverbal to avoid setting expectations of intelligence that the robot cannot meet (which would lead to the same disappointment that many competitor's offerings create). Motor noise will be reduced through the choice of motors and gears, as well as potentially using the robot's programmable audio output to mask motor noise.

Also critical to the user experience is that the robots act to achieve goals that are not reliant on human interaction. To that end, the robots are being designed with a slightly concave front section, which will ease the pushing of various objects. In later generations beyond this proposal, more dextrous forms of manipulation will be considered.

Lastly, the robot's management of its energy needs is an innovation of this project. In particular, through the overhead camera on the puppeteer platform, *the robot will recharge itself with unprecedented reliability for a mobile robot*. As a consequence, the robots will be able to persist as part of their users' physical environment for months without any user intervention. In contrast, free-roaming robots that self-recharge, like iRobot Roomba models, often cannot find the charging base and users to search for it and then manually recharge it. Two large benefits arise from this innovation: long-term human-robot interaction will be supported by the negligible cost to the users' time to keep the robots running, and the robots can support a wider range of user interaction levels than previous autonomously mobile robots and electronic toys could. *Some users may want to interact minimally with the robots, creating a qualitatively novel form of long-term interaction with an electronic character that can be nonetheless rewarding, like periodically observing fish in a tank or dogs at play.* The robots will also be designed to be able to run for at least 30 minutes at a time (though 1.5 hours will be targeted) before needing to recharge while the robot "sleeps". The battery will be chosen to support this run time and also give several months of recharging before it needs replacing.

Robot hardware The robots' mechanics can be abstracted as a two-wheeled cart, giving it two degrees of freedom. The wheels, however, will be hidden under the robot's body, invisible without lifting the robot or lowering one's eyes to the surface on which the robot sits. Optical encoders will give accurate measure of each wheel's angular rotation, with high granularity. With this data, the robots will run a closed-loop controller (e.g., PID) to achieve more accurate velocities. Wheel rotation data will also be sent to the puppeteer platform, which will use this data and delayed computer vision inference to estimate the robot's translation and rotation. The robot will also sense its battery voltage and current. And it will sense touch, likely through capacitive sensing, data from which will also be sent to the platform. The robot will play audio and light sequences as communicated by the platform. The robot and platform will conduct two-way wireless communication via radio frequency communication (likely Bluetooth Low Energy). Lastly, the robot's chassis will connect interchangeably to different exterior shells, allowing variety in the robot's appearance without changing the electromechanical design.

5.1.3 Puppeteer platform hardware design

The platform's main components are a computer, a video camera, a microphone, and a robot charging base. See Figure 2. In the first generation, beta version of the product that we plan to ship in 2017, the computer will be an inexpensive single-board computer (SBC) such as the Raspberry Pi 3 ($35 retail). Emoters plans to design custom SBC from the second generation onward. This SBC will connect to a wide-angle camera that provides a birds-eye view of the surface underneath, where the robot will reside, as well as nearby people's legs and hands. Balancing vision-based detection of human activity—for example, to permit the robot to look at a person who is walking by—with privacy concerns will be critical to the successful and ethical deployment of these and future interactive robots. The SBC will also connect to a microphone. The platform will communicate wirelessly with both the robots and to the internet, and it will contain a battery charging station to which the robots can dock, appearing to sleep. The platform will not have a screen but will instead present a user interface (e.g., for initial setup and further purchases) through users' smartphones, with which the platform will communicate through Emoters' servers, a local network, and/or Bluetooth.

The use of a platform comes at one considerable disadvantage: the robots cannot be free roaming. However, its advantages easily outweigh this drawback. The external camera and the limited distance the robots will travel from it allows the robots to autonomously recharge with near-perfect reliability. With a fixed camera position and the possibility to add artificial lighting if needed, the environment presents less variety to the computer vision models, easing their task somewhat. Performing perception and decision-making largely in the puppeteer platform—instead of on the robot, as is the norm in consumer robotics—has several advantages: sensing and computation needs of the robots are lowered dramatically, permitting much lower per unit cost for the robots; perception can be performed once and used by all robots; the platform's sensors and computer are less expensive than they would be on the robot, where they would need to be smaller and more robust to rough handling; the robot can be smaller and lighter, lessening load on batteries; and third-party remote-control robots and toys can be given socially interactive, autonomous behavior through the platform. The platform also supports reliable internet connectivity, since it does not move on its own. Lastly, the platform will help Emoters address privacy concerns. A free-roaming robot with computer vision would leave users uncertain at times about whether they are in the camera's view, whereas our platform is stationary; with its downward facing camera, our platform should also capture less sensitive human activity than a forward-facing camera. (For users' privacy, an optical cap can be provided and placed over the wide-angle camera, creating blinders that allow it to only see directly below where the robots are.)

5.1.4 Perception in platform

Perception of state relevant to the robot comes from several sensors. On the platform is a overhead camera and a microphone. On the robots are two optical encoders, a touch sensor, and a sensor for battery voltage and current. Information from robot sensors will used directly by the robot for simple, reactive behavior in case the robot is carried away from the platform, but it will also be communicated to the platform.

The camera will provide much of the perceptual information used to provide context for the robot's behavior. Using computer vision to accurately detect the robots' poses and velocities is critical to their control. Also, the platform will detect various rigid objects of interest and generic obstacles, which will the robot will typically try to avoid. As the robot character's capability to interact with people is developed, we will implement detection of human state and activity, such as hand locations, gesture recognition (e.g. pointing for the robot to go somewhere), detection of people's general location when not immediately adjacent to the platform, and detecting facial expressions of people who have brought their faces down to the robots' level.

From the microphone, the platform will detect aspects of human speech—perhaps understanding a few simple words or merely that *an utterance* was vocalized. The platform may also detect music, which the robots may react to and even learn songs to sing later. Data from the touch sensor, likely in the form of a binary determination of whether the robot is currently being touched, will also provide context to aid autonomous human-robot interaction. The battery readings will determine whether the robot has successfully docked with the charging station and detect certain scenarios in which the robot is not working properly and needs repair or replacement. The readings from the optical encoders will improve the platform's estimate of the robot's pose and velocity—which would be problematically delayed if inferred from only the camera stream—increasing its ability to execute high-speed, precise movement with the robot. The pose estimate will be maintained through a tracking filter such as a Kalman filter.

Computer vision algorithms will be implemented for foreground separation (to detect the objects in the platform area), hand detection (to understand direct human interaction), object recognition (to recognize a specific set of objects via color, shape, and derived local-feature models), and object tracking (to reduce the effects of errors in the aforementioned algorithms' inferences). Object detection may be accomplished through careful tuning of an existing fiducial marker library (e.g., ArUco markers in opencv). Location, orientation, and identification metadata will be generated for each object visible on the platform. In later generations beyond this proposal, Emoters plans to infer further metadata such as facial expression and person recognition (to disambiguate different users of the same system). Field data and experimental results will inform the improvement of the perceptual system's speed and accuracy in different lighting and usage scenarios. Like the rest of the platform's software that operates the robots, the inference models used for computer vision and possibly other parts of perception can be updated through Emoters' servers.

5.1.5 Robot decision-making

On the puppeteer platform, each robot will be controlled by a separate artificial intelligence process. The precise artificial intelligence algorithm will be determined through machine learning applied to demonstration data from past sessions of secret teleoperation, as we detail below.

Research-based foundations *Wizard of Oz* is a technique used in human-computer interaction prototyping and research. A Wizard-of-Oz interaction involves a human user interacting with an interface that *appears* to be autonomous but is secretly being controlled by another person. Wizard-of-Oz interactions are common in research for human-robot interaction. During the PI's postdoctoral appointment at MIT, he led the design of *a system employing Wizard-of-Oz interaction with a socially interactive robot to gather teleoperator demonstration data*. In its raw form, this time series data includes full sensory history and teleoperator control history for each interaction of interest. The data is then discretized into time steps (100ms each in this past work), providing samples for supervised learning. A sample from a time step contains—as is the norm for supervised learning—a set of feature values and a label. Here, the label is the teleoperator action (including a potential no-action action), and the feature values are descriptive statistics from sensory and action history. From these samples, *a probabilistic model of the teleoperator is learned, which can be used in lieu of the teleoperator, yielding autonomous social behavior that emulates the teleoperator.* This technique was termed "**learning from the wizard**" (LfW). Previously in this proposal, LfW has been referenced by more general descriptors similar to "machine learning from demonstration data created during secret teleoperation".

In the experiments above, social robot interaction was between a child (4–8 years old) and a small plush robot, who share control of an educational app on a tablet between them. The learned autonomous behavior was evaluated against Wizard-of-Oz teleoperation and tablet-only interaction (with no robot present) in a randomized experiment of 85 participants. The experiment found the teleoperated robot and the autonomous robot programmed by LfW elicit similar behavior from their human interaction partners. Additionally, when children were asked whether the robot was human-controlled or autonomous, approximately half in each condition thought it was human-controlled. These results demonstrated that a robot programmed via LfW can successfully engage in highly social, interactive tasks at the level of a human-teleoperated robot.

Designing LfW for Emoters' robot pet domain Emoters will employ LfW to specify autonomous robot behavior, allowing robots to be modeled after highly organic behavior by a puppeteer who is herself immersed in the interaction with the user. LfW is expected to result in not only better behavior than that available in competitive offerings, but a qualitatively different one. The predominant method for determining the behavior of an electronic toy is to program behavior or behavioral models directly, for instance via if-then rules or finite state machines. Direct programming results in a greater gap between the person specifying the behavior and its real-time effects; closing this gap through learning from demonstration promises considerable improvements in the authenticity and variety of Emoters robots' behavior.

However, the application of LfW to the Emoters domain will require significant innovation, since it extends the technique from a simpler problem. The action space in previous work was discrete and one-dimensional: a set of buttons that triggered robot actions. The robot in past work also had limited physical interaction with the world; its physical movement was used for expression only and did not result in locomotion or manipulation of physical objects. In contrast, the action space for Emoters' robots over a time step includes two motor commands, LED RGB values (or possibly a sequence of them), and audio waveforms. This action space presents two challenges, with its higher dimensionality that includes continuous variables. First, the complexity of the model must grow, which risks overfitting. Second, the timing and choice of each robot's movement, audio, and lighting will be interrelated. This interrelation and the complexity of each robot's interaction with the physical environment motivate a hierarchical representation of a robot's internal state and behavior. This hierarchy would likely include its emotional state, current goal, outcomes of recent attempts to achieve goals, and other information that would not be captured in typical demonstration data. Three potential solutions are to manually define the hierarchy and add controls to the teleoperator interface for indicating internal state; to manually define the hierarchy and algorithmically infer latent state; or to algorithmically infer the hierarchy as a function of observable history. There is promise and risk in each of these options. For example, the first option increases interface complexity for the teleoperator, possibly beyond his or her capabilities. One way to reduce teleoperator load would be automating parts of the hierarchy. For instance, the teleoperator could designate only the high-level goals, communicative actions, and emotional state. Alternatively, this information could be given to a teleoperator, who then controls robot movement, sound, and audio to fit the high-level information.

Solving the problem above would leave ample follow-up work for Phase II. In particular, Emoters would add the ability for robots to learn from their individual experiences. The most obvious choice to add this capability is to incorporate some form of reinforcement learning [25]. However, it is unclear how to elegantly combine LfW with reinforcement learning while retaining the organic feel of behavior learned from demonstration. Additionally, the choice of reward function would be challenging; a robot that acts optimally with respect to a reward function may be an awful interaction partner. Emoters may address this challenge by employing inverse reinforcement learning [20, 5, 30] to learn a reward function that, when optimized for, results in behavior that emulates the teleoperator. The robot could then learn a forward model from its experience that it uses to accrue reward according to the learned reward function. In one example, demonstration data may contain a pattern of the robot going towards "food" objects and avoiding rough handling by humans. Inverse reinforcement learning could output a reward function that encourages pursuit of food and discourages putting oneself in danger of rough handling. Consider resultant robots with two different histories presented with the opportunity to obtain food near a user's hand. A robot with a gentle and trustworthy owner would learn that abuse has a low probability, and so it would likely choose to get the food immediately. Another robot, with a history of being handled roughly, would predict abuse and avoid the hand, waiting for the user to leave before movement.

In Phase II, Emoters may also incorporate learning from human-delivered reward and punishment, making the robots trainable by their users, which was the focus of the PI's dissertation [14] and is a popular candidate feature in user surveys.

5.1.6 Data collection and experimentation

Data collections, storage, and privacy Emoters is dedicated to respecting its customers' privacy. To that end, Emoters will follow Internet-of-Things security best practices and be transparent about its use of customer data. Additionally, guided by standards enforced in the telecommunications industry, all user data is encrypted in transit and storage and its access will be restricted to a minimal set of staff and anonymized wherever possible. Specific details follow.

Potential customers will be provided a clear description of Emoters' data collection, storage, and privacy policies before they purchase the system. Emoters may also create an online interface that allows users to view and manually erase data collected from their platform. Data stored locally on the platform includes a definable expiration date so that only recent history is available. The platform itself will be firewalled and incoming local access (via serial, wireless, or keyboard) will be disabled. All data from users' platforms will be encrypted both when in transit to Emoters' servers and while in storage there. Though aggregate information across many users will be accessible to Emoters personnel, individual information stored by Emoters will be made accessible *only* on a need-to-know basis. Sensitive, personal information (SPI) that uniquely identifies users will be stored in a separate database with different access privileges that are also granted on a need-to-know basis. Entries in these two databases will be linked only with a randomly generated identifier code. Emoters does not plan to collect audio or video unless it is expected to yield significant benefit to users, and audio and video would only be collected for an individual platform if its user explicitly opts in to permit such collection. Emoters may allow users to opt out of collection for some categories of data, selected based on user feedback. Emoters will never sell to third parties data that is specific to individual users.

Data to collect Emoters will optimize its robots towards behavior that is increasingly valuable for users. To do so, Emoters will aggregate three categories of data: user interaction metrics, system health metrics, and contextual data. With the first generation product, user interaction data collected will include whether a person was interacting with the system at a given time, human-robot touch events, and events in which the user rewards or feeds the robot. Such data will allow Emoters to assess mission-critical outcomes, such as whether different versions of the robot behavior code change the amount of human-robot touch or affect the proportion of users that still interact with the robots after several months. In later generations of the product, collection will also include information regarding human affective state and interaction with an Emoters smartphone app that interfaces with the puppeteer platform (e.g., allowing users to edit and post their platform's video stream to social media).

To calculate system health metrics, Emoters will collect data regarding each robot's battery, instances in which a robot's execution of the platform's commands do not match with expectations (indicating potential electromechanical failure), and measurements of bandwidth and latency between the platform and Emoters' servers. Such information will permit Emoters to provide excellent customer support. For instance, Emoters will be able to automatically detect that a robot is broken and under warranty, at which point Emoters could then notify the user that a replacement is being shipped, without the user having to take a single action.

Emoters will also gather contextual information, such as the user's coarse geographical location and the robot software and hardware versions on the platform. Like the data described above, this contextual information will be employed to improve user experience. For instance, coarse geographical location information could be combined with measurements of platform-to-server latency to identify regions where Emoters should deploy servers, reducing the latency for nearby users.

Conducting field experiments With the collection of the data described above, an experiment will consist of running alternate versions of robot software on the platforms of randomly chosen user subsets. The marketing community calls this technique *A/B testing*. Differences in robot software could be in the perceptual subsystem, the behavioral subsystem, or other subsystems such as wireless communication with the robots. Emoters will implement an experimentation interface that permits Emoters personnel to easily schedule the experimental deployment of a version of code. A visualization of experimental results along key metrics will also be designed and implemented. This system, once implemented, will enable Emoters to rigorously and efficiently improve the quality and value of user interaction with its robots. Additionally, this system will be valuable to companies developing robots and smart toys, who could pay licensing fees for its use.

In Phase II, Emoters would focus on the design of optimization algorithms that automate the experimentation process. Instead of Emoters personnel manually choosing versions of robot software to deploy, it would be parameterized and searched by an optimization algorithm for the set of parameters resulting in the most desirable user experience. This parameterization is a generalization of the manual-experimentation system to be developed in Phase I. Automated optimization will boost the efficiency with which Emoters learns. For instance, if a software version in the experimentation space results in an undesirable user experience, the automated system—with more sophisticated estimates of true values—could recognize with only a few samples that it is highly improbable for this version to receive positive results, immediately remove that version from all users' platforms, and focus on alternate parameters. In Phase II, Emoters will also seek to reduce the time it takes to determine the quality of a software version by identifying short-term metrics that predict long-term interaction metrics. Ideally, interaction aspects in early few weeks—or even hours—may be predictive of interaction over years and the identification of these metrics (perhaps through machine learning) will speed the optimization of user experience. Lastly, in Phase II Emoters will adapt the data-collection and experimentation infrastructure built in Phase I to scale to millions of users.

5.2 Phase I key objectives

Emoters' key objectives for the 6 months of Phase I are listed below. These objectives represent feasible steps towards the broader goals of this overall proposal; however, completion of these objectives also situates Emoters with a *minimum* viable product that is ready to go to market. Associated criteria for technical and commercial feasibility of this proposal are also described. Feasibility will be demonstrated if at least 3 of the 5 objectives below are completed according the their feasibility criteria.

- **Robust tracking of the poses of robots, a few objects of interest, and generic obstacles** – *Feasibility criterion: On teleoperated data from robot interaction during user testing across a range of lighting conditions, new objects, and various surfaces that approximate the range of expected usage, tracking results in accuracy of object centroids within 2 cm for 80% of the frames. For objects of known shape, this criterion also includes accuracy of object orientation within 30 degrees for 80% of the frames.*
- **Secure infrastructure for data collection and field experimentation** – *Feasibility criterion: With in-home deployment of 5 or more systems for user testing, Emoters can collect the data types described in Section 5.1.6, store the data in a secure database, and remotely determine what version of robot software is being run in each deployed system. Results of variations in deployment software automatically populate a password-protected browser-based visualization. Additionally, 50% of participants in user testing state that the privacy characteristics would not prevent them from purchasing and using the system.*
- **Hardware design of the low-cost puppeteer platform and robot** – To create a manufacturable product for consumers, Emoters will require electrical- and mechanical-engineering development, in addition to work by Emoters' partner Pushstart Creative. *Feasibility criterion: These two hardware products have reached "looks-like, works-like" prototype stage and are ready to iterate with Emoters' manufacturing partners to design for manufacture. At a scale of 1000 units or less, the bill of materials is less than $20 for the robot and $80 for the puppeteer platform.*
- **Hardware and software design of highly reliable autonomous recharging** – Phase I's criterion focuses on recharging without potential user-caused problems, which will be investigated in Phase II. *Feasibility criterion: Without user interference, robots can recognize the need to recharge and autonomously dock with the recharging station before dying from lack of charge in 95% of instances.*
- **Character development using machine learning on teleoperation data** – Discussed in Section 5.1.5, Emoters will adapt PI Knox's prior NSF-funded research to Emoters' domain. *Feasibility criterion: Qualitatively, the character exhibits goal-achievement and communicative actions, conducted in different expressive styles that fit an appropriate mood or emotion for the robot given its past history.*

Biographical Sketches

[content redacted]

Facilities, Equipment, and Other Resources

1 Facilities

Emoters currently rents approximately 500 sqft of commercial real-estate as an office space in Austin, Texas. This facility houses a room for user evaluations, staff meeting, and two room for development and daily activities. All locally accessed equipment (e.g., printers), instrumentation, and computers are stored at this location.

2 Equipment

As a hybrid software and low-cost hardware company, Emoters' in-house equipment roster is short and low-cost itself.

- **development computers** Emoters currently has six development computers and each employee utilizes his or her own personal computer for development. Emoters' computing resources host software for project management, remote-access, and development prototypes of Emoters' products.
- **deployment prototypes** Emoters utilizes several (currently four) single board computers (SBC) as prototypes and initial deployment platforms for its product. These SBC's usually cost less than $40 and are purchased as-needed from a pre-allocated capital budget.
- **printers** On-site printing facilities include one office inkjet printer and one inexpensive 3D printer. Emoters currently employs distributed printing services for the manufacture of its printed circuit boards (PCBs) that serve as the electromechanical platform for its robots. As development progresses, Emoters will out-source printing jobs and eventually move to exclusive use of an injection-mold manufacturing facility.
- **cloud services** Emoters' employs cloud-hosting services for high-connectivity and high-bandwidth internet services. These services include domain name hosting, web site access, database storage, and a source code repository. Some services currently run on shared resources subleased personally by staff, but with full funding, Emoters will migrate these services to private and secured resources owned by the company.
- **electrical workstations** Electrical work stations used by staff at Emoters for robot and platform development, including ohmmeters, circuit testers, and microprocessor controllers are housed at Emoters' primary facilities. Staff also use personal workstations in off-hour sessions, but Emoters compensates staff for these resources.
- **ancillary software** All software used in development of Emoters' products are either (a) wholly owned by Emoters' through license purchases, (b) open source software, or (c) software owned individually by staff at Emoters and explicitly permitted for use by employment agreements.

Data Management Plan

Emoters is a software and hardware robot company that will collect usage data of its products and services for product improvements and greater user satisfaction. All data generated in this SBIR Phase I project is considered proprietary. The sections below, as well as Section 5.1.6 in the Project Description, detail Emoters' plan for data management.

1 Expected Data

Emoters will collect and store the following types of data that are specific to each user, but the most of the data will be usability and interaction data, as described in detail in Section 5.1.6 of the Project Description.

- **Interaction Metadata** Emoters may collect interaction data that describes how a human user, robot, and objects within a scene are utilized during play with the robot.
- **Interaction Content** Emoters may collect video and audio content, if users opt in to permit such collection, as part of its description of the interaction between robot and human user. By design this data should not include any recognizable personal biometrics (e.g. face, fingerprint, voice), but through the course of normal interaction by the user, these artifacts may temporarily become available. Emoters may also collect video and audio content *explicitly at the request of the user* such that the video and audio content may be shared on social media or in personal electronic correspondence.
- **Sensitive Personal Information** For the purpose of understanding and characterizing product use, personally identifiable information about a customer may be collected, including a user's name, home address, phone number, age, sex, household demographics, and other general demographic information. Other personal information regarding payment details may also be collected. Emoters may store these details with a financial services partner instead of managing the information in-house.

2 Period of Data Retention

With best efforts to adhere to industrial data retention, Emoters will scrub and store metadata about its products. Specifically, personally identifiable interaction metadata may be stored for up to one year. After one year, data will be scrubbed of personal information (e.g. any linkage to specific SPI) and may be stored as metadata that is aggregated anonymously by demographic or randomly generated user identifier. This aggregated data may be stored for immediate use for up to five years. After this time, the data will be parameterized as data models and/or archived into restricted-access, long-term storage exclusively utilized for long-term longitudinal studies.

3 Data Formats and Dissemination

Wherever possible, data will be stored in semi-textual (e.g. JSON, XML, etc.) but encrypted format. While a specific vendor has not been chosen for data management at Emoters, SPI and long-term data storage will employ 256-bit AES keys or better (as described by ISO 27001) and internet-channel communication and short-term data storage will employ version 2 of DSA or RSA keys, as commonly supported by the SSH protocol, version 2. Short-term data is defined as data that exists on a user platform, data in transit to Emoters, and data that is temporarily stored on disk. Online data, which exists in an actionable database at Emoters, will be dissociated from any SPI and will be referenced by a randomly generated moniker.

Although deferring to the data governance described above, dissemination policies for a few forms of user specific data are detailed below.

- **Sensitive Personal Information** Emoters will never release sensitive personal information (SPI) to any vendor, partner, or subsidiary without direct customer consent and explicit company imperative for operation (e.g. billing, collections, product delivery).
- **Marketing Data** Emoters may share conclusions regarding its data about the design of robots with contracted marketing firms to develop additional revisions of the robots. Emoters may also refer to external marketing firms for expertise to engage the appropriate advertising firm or marketing strategy.

4 Data storage and preservation of access

As described in the Section 3, data will be stored in an encrypted format. On-line data (that which exists in a database) will be dissociated from any SPI or other metadata that could uniquely identify a user on its own. Short-term (file, in-transit) and long-term (archive) data will be encrypted with 128-bit and 256-bit keys accordingly.

Access controls will be created within Emoters to appropriately limit access to SPI, user metadata, and other user content to designated staff. Within Emoters, functions associated with traditional privacy review boards (PRB) may be integrated with polices described by Emoters' IRB, described in Section 6. A typical Emoters employee will have access to only aggregated usage an interaction metadata to best assist in their design of robot and their corresponding interaction algorithms.

5 Scientific Publications and Products

As a company with strong roots in academic pursuits, Emoters will strongly consider opportunities for open access of aggregated, anonymized metadata and algorithms. A balance between peer-reviewed publications, patents, and interal trade secrets will be regularly evaluated. Emoters is currently fostering mentorships with its proximal institution, The University of Texas at Austin through exposure of some hardware design processes to graduates. Additionally, as part of Emoters' market and product strategy described in Sections 2.1 and 3.4, developer kits and design references will be released to the public.

6 Human Subjects and Vertebrate Animals

Emoters expects to be either exempt or expedited for IRB approval, which it will immediately seek if and when it receives informal notification of this proposal's approval. A letter of support from the IRB consultancy Quorum Review IRB is included as a supplementary document.

References Cited

[1] Furby - Wikipedia, the free encyclopedia. https://en.wikipedia.org/wiki/Furby. Accessed: 2015-10-20.

[2] Pets iOS Games - App rankings, revenue & ARPU estimates. https://thinkgaming.com/app-sales-data/tag/pets/. Accessed: 2015-09-12.

[3] ROS-Introduction. http://wiki.ros.org/ROS/Introduction. Accessed: 2015-10-20.

[4] Total toy market size by country 2009 / 2011. http://www.statista.com/statistics/194399/toy-market-size-by-country-since-2009/. Accessed: 2015-11-20.

[5] Pieter Abbeel and Andrew Y Ng. Apprenticeship learning via inverse reinforcement learning. In *Proceedings of the twenty-first international conference on Machine learning*, page 1. ACM, 2004.

[6] Brenna D Argall, Sonia Chernova, Manuela Veloso, and Brett Browning. A survey of robot learning from demonstration. *Robotics and autonomous systems*, 57(5):469–483, 2009.

[7] American Human Association. Keeping pets (dogs and cats) in homes: A three-phase retention study. phase i: Reasons for not owning a dog or cat, 2012.

[8] American Pet Products Association. 2015-2016 APPA national pet owners survey, 2015. Accessed: 2015-10-20.

[9] Greg Avery. 'Star Wars: The Force Awakens' toy droid BB-8 is transforming Boulder's Sphero. http://www.bizjournals.com/denver/blog/boosters_bits/2015/10/star-wars-the-force-awakens-toy-droid-bb-8-is.html. Accessed: 2015-12-01.

[10] Wilma A Bainbridge, Justin W Hart, Elizabeth S Kim, and Brian Scassellati. The benefits of interactions with physically present robots over video-displayed agents. *International Journal of Social Robotics*, 3(1):41–52, 2011.

[11] United States Census Bureau. Table 4. Estimates of the housing inventory. http://www.census.gov/housing/hvs/data/q315ind.html, 2015. Accessed: 2015-12-01.

[12] Heather M Gray, Kurt Gray, and Daniel M Wegner. Dimensions of mind perception. *Science*, 315(5812):619–619, 2007.

[13] Cory D Kidd and Cynthia Breazeal. Effect of a robot on user perceptions. In *Intelligent Robots and Systems, 2004.(IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on*, volume 4, pages 3559–3564. IEEE, 2004.

[14] W. Bradley Knox. *Learning from Human-Generated Reward*. PhD thesis, 2012. Dissertation page: http://web.media.mit.edu/ bradknox/Dissertation.html.

[15] W. Bradley Knox, Samuel Spaulding, and Cynthia Breazeal. Learning from the wizard: Programming social interaction through teleoperated demonstrations. In *Under review*.

[16] W. Bradley Knox and Peter Stone. Combining manual feedback with subsequent MDP reward signals for reinforcement learning. In *Proceedings of The 9th Annual International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, May 2010.

[17] W Bradley Knox, Peter Stone, and Cynthia Breazeal. Training a robot via human feedback: A case study. In *Social Robotics*, pages 460–470. Springer, 2013.

[18] Hatice Kose-Bagci, Ester Ferrari, Kerstin Dautenhahn, Dag Sverre Syrdal, and Chrystopher L Nehaniv. Effects of embodiment and gestures on social interaction in drumming games with a humanoid robot. *Advanced Robotics*, 23(14):1951–1996, 2009.

[19] Billy Langsworthy. Furby Boom named best selling toy of 2013. http://www.toynews-online.biz/news/read/furby-boom-named-best-selling-toy-of-2013/041659. Accessed: 2015-12-01.

[20] Andrew Y Ng, Stuart J Russell, et al. Algorithms for inverse reinforcement learning. In *Icml*, pages 663–670, 2000.

[21] The Associated Press. Toy industry makes big sales strides. http://triblive.com/business/headlines/9415884-74/toys-toy-sales. Accessed: 2015-12-01.

[22] Mitchel Resnick, John Maloney, Andrés Monroy-Hernández, Natalie Rusk, Evelyn Eastmond, Karen Brennan, Amon Millner, Eric Rosenbaum, Jay Silver, Brian Silverman, and Yasmin Kafai. Scratch: Programming for all. *Commun. ACM*, 52(11):60–67, November 2009.

[23] Laurel D Riek. Wizard of oz studies in hri: a systematic review and new reporting guidelines. *Journal of Human-Robot Interaction*, 1(1), 2012.

[24] Jonathan Soble. A robotic dogs mortality, June 2015. [Online; posted 17-June-2015].

[25] Richard S. Sutton and Andrew G. Barto. *Introduction to Reinforcement Learning*. MIT Press, Cambridge, MA, USA, 1st edition, 1998.

[26] Russell Toris, Julius Kammerl, David V Lu, Jihoon Lee, Odest Chadwicke Jenkins, Sarah Osentoski, Mitchell Wills, and Sonia Chernova. Robot web tools: Efficient messaging for cloud robotics. 2015.

[27] Joshua Wainer, David J Feil-Seifer, Dylan Shell, Maja J Mataric, et al. Embodiment and human-robot interaction: A task-based perspective. In *Robot and Human interactive Communication, 2007. RO-MAN 2007. The 16th IEEE International Symposium on*, pages 872–877. IEEE, 2007.

[28] Daisy Yuhas. Pets: Why do we have them? *Scientific American Mind*, 26(3):28–33, 2015.

[29] S. Zickler, T. Laue, O. Birbach, M. Wongphati, and M. Veloso. SSL-Vision: The Shared Vision System for the RoboCup Small Size League. *RoboCup 2009: Robot Soccer World Cup XIII*, pages 425–436, 2009.

[30] Brian D Ziebart, Andrew L Maas, J Andrew Bagnell, and Anind K Dey. Maximum entropy inverse reinforcement learning. In *AAAI*, pages 1433–1438, 2008.

7 Supplemental details on encryption technique, security risk and privacy protection

7.1 Data encryption

As described in the Section 3, data within Emoters' datacenters is segregated into SPI (sensitive personal information) and other usage data required for product maintenance. SPI data (user name, customer address, image data outside of the immediate play area, customer communications, payment information, etc.) will be online, in a segregated database as PGP archives where only a unique, randomly generated BAN (billing account number) identifies a customer. Non-SPI data (product usage, robot names, robot interaction data, usage profiles, etc.) are also on-line and uniquely identified by the BAN. At the source (user's remote platform), non-SPI data is stored with BANs only and any larger SPI data (ancillary image or audio data) is retained locally as PGP encrypted archives before upload and erasure. Smaller SPI artifacts (payments, address information, etc.) are not stored on end-user platforms at any point.

7.2 Security Risks

Security risks from on-line or physical data breaches are low with Emoters' platform. On-line SPI is stored in a segregated, privileged database as encrypted PGP archives. Unique PGP keys will be generated quarterly such a data loss effects only customers that joined within a quarter and share a PGP key. PGP keys themselves will be stored on encrypted storage and accessible only by privileged automated billing systems through pre-specified business functions (billing, etc). Data storage on the user platforms will be specific to that user and free of SPI. Regular, automated data transfers to Emoters' servers will upload usage data over authenticated SSH2 data channels, add the data to Emoters' datacenter archives, and then immediately destroy local copies.

7.3 Privacy Protection

Emoters is sensitive to user's privacy and the need for its protection. While users may not opt-out of usage data uploads, they may restrict the data to not contain any SPI content (image data outside of the play area, audio tracks, etc). Emoters has no immediate plans for marketing campaigns with user data, but if engaged, users will be provided similar opportunities to opt-out and restrict usage to security- and product-only contact.

7.4 Demographics for user testing

Of the most recent 35 user testing participants, 54.2% were male and 45.8% were female. 0% were under 18 years of age, 14.3% were 19–25, 34.3% were 26–32, 37.1% were 33–40, and 14.3% were 41 or older. We also asked the 29 most recent (of those 35) their highest education level. 31.0% had "some college", 34.5% had a college degree, 27.6% had a graduage degree, and 6.9% had a doctoral degree. Some of the nine participants who had "some college" performed professional work despite not completing college; two were engineers, one worked in video production, and one had been a video game animator.