Quori: A Community-Informed Design of a Socially Interactive Humanoid Robot

Andrew Specian\textsuperscript{1}, Ross Mead\textsuperscript{2}, Simon Kim\textsuperscript{3}, Maja Matarić\textsuperscript{4} and Mark Yim\textsuperscript{1}

Abstract—Hardware platforms for socially interactive robotics can be limited by cost or lack of functionality. This paper presents the overall system—design, hardware, and software—for Quori, a novel, affordable, socially interactive humanoid robot platform for facilitating non-contact human-robot interaction (HRI) research. The design of the system is motivated by feedback sampled from the HRI research community. The overall design maintains a balance of affordability and functionality. Initial Quori testing and a six-month deployment are presented. Ten Quori platforms have been awarded to a diverse group of researchers from across the United States to facilitate HRI research to build a community database from a common platform.

Index Terms—Social Human-Robot Interaction; Product Design, Development and Prototyping; Humanoid Robots; affordable hardware; non-contact HRI; socially interactive robots.

I. INTRODUCTION

This paper presents Quori (Fig. 1), a novel, socially interactive robot for facilitating non-contact human-robot interaction (HRI) research, both in the lab and in the wild. Quori aims to provide an affordable, high-quality platform that allows HRI researchers to conduct meaningful user studies by deploying systems in the real world. The Quori platform includes both hardware and software to help facilitate HRI. Quori was designed and produced with support from the National Science Foundation Computing Research Infrastructure grant, which included the support for ten Quori platforms to be distributed to researchers in a competitive project proposal process.

This paper introduces the approach to designing Quori and the details of the complete system. In Section II, we provide some background on HRI. Then, in Section III, we discuss our engagement with the HRI research community that helped inform the most important hardware and software capabilities for a socially interactive robot for HRI research; the data collected from this diverse group of researchers in the broader HRI community directed the design decisions for robot hardware and software—this quorum of researchers inspired the name “Quori”. The rest of the paper focuses on describing the hardware (Sections IV-V) and software (Section VI). Section V-A presents our approach to achieving an affordable design. Section VII describes testing of the system and the four-month deployment of Quori at the Philadelphia Museum of Art.

II. BACKGROUND

HRI research focuses on rich multimodal interaction that occurs between humans and machines [1]. Two subsegments of this field include contact-based HRI and non-contact HRI.

Contact-based HRI intersects with medical robotics, haptics, rehabilitation robotics, and is related to manipulation research and to human-robot collaboration.

Non-contact HRI grew out of social robotics [2] and socially assistive robotics [3], and focuses on the perceptual and computational aspects of HRI that involve no physical manipulation of the environment or intentional tactile interactions with people [4], thereby complementing contact-based HRI. Robot hardware and software necessary for pursuing the challenges of non-contact social HRI have many unique requirements. These platforms must be capable of recognizing multimodal social and behavioral signals, reasoning over those signals and behaviors, and generating appropriate affective, expressive, and communicative behaviors in response [2], [4]. The range of relevant socially interactive behaviors includes eye gaze (i.e., where/when to look), use of space (e.g., where to be, how large to gesture), timing behavior (i.e., turn taking), expressive behaviors (e.g., how to gesture, how to move, how to communicate), body language (e.g., how to move the body to express a personality/character), and non-verbal and verbal communication (e.g., what sounds to use and when, speech processing, natural language understanding, and dialog management), for all of which the problems of behavior recognition, understanding, selection, and control must be solved [5].
The production of certain social behaviors (e.g., facial expressions and hand/arm gestures) necessitate expressive degrees of freedom (DoFs) on HRI platforms that other robots might not require; for instance, the human face provides rich signaling of affect for human understanding [6], so robots with expressive faces can exploit this social communication channel to convey affect and intent. Expressiveness goes well beyond the face: affect, demeanor, intent, personality, and character are all expressed subtly and effectively through body language. Functional gestures (e.g., deictic pointing gestures) are most readily produced with arms. Therefore, social robots often have arms or other DoFs not used for object manipulation.

The HRI design features include the robot head, arms, torso, and mobility hardware and software and their associated costs.

A. Robot Heads for HRI

Robot faces can aid in gaze definition, lip readability, and emotional expression [2] among other purposes that facilitate HRI. Faces can be molded and static (e.g., SoftBank’s Pepper robot [7]), can be mechanically actuated for expressivity (e.g., Kismet [2] and EMYS [8]), can be human-like with few DoFs (e.g., Bandit [9]) or many DoFs (e.g., Sophia from Hanson Robotics, which use synthetic skin and facial muscle actuators [10]). Another design solution uses displays for affordable and versatile faces (e.g., Kiwi [11], which uses Facial Action Units and visemes). Finally, another approach is to use rendered faces projected onto the head, as was done on Quori, as with screens, projected faces have an inherent flexibility for variability and testing, and they also add the potentially more natural head-shaped form, enabling research into how such heads and faces are perceived and accepted by users.

To maximize flexibility of Quori’s rendered face, we exploited the availability of affordable portable projectors in a retro-projected animated face (RAF) capable of projecting an image 360° around the head. Smaller section RAFs have been shown to be highly expressive, such as in Lighthead [12], Mask-Bot [13], Furhat [14], and Engineered Arts’ Sociobot [15]. The RAF allows for expressive and customizable faces for HRI research [16] with potential benefits in eye gaze detection over flat screens [14].

B. Robot Arms for HRI

Non-contact HRI robots can use arms for expressivity without touching objects or people. Common functional robotic arm gestures include deictic pointing (e.g., [17]), task demonstration (e.g., in promoting seated exercises for the elderly [9]), and imitation (e.g., in autism movement training [18]). The human arm can be modeled as four DoFs when considering the shoulder and the elbow as described by the model in [19]. Since robot cost and complexity increase with the number of DoFs, it is useful to consider functionality with a subset of those DoFs. One DoF can perform simple pointing and express limited affect [20]. Two DoFs in the shoulder allow for more accurate and natural-appearing pointing and expressive motion [21]. Higher DoF arms have been used in rehabilitation and related tasks mentioned above [22]. Quori’s shoulders have two DoFs each, and can modified to add additional DoFs; for details see Section V-C.

Robot arms are an inherent safety risk, due to possible collisions. Safety can be characterized by the Head Injury Criterion (HIC) [23], a combination of arm inertia and stiffness. To ensure a safe low-inertia system, low-mass mechanisms with gravity compensation can be used [24]. Quori uses lightweight arms with a friction clutch and low-power motors to minimize damage from accidental collision.

C. Torsos for HRI

Nearly all robots used in HRI research have no flexibility beneath the torso—they lack the ability to bend forward. A waist joint that moves the torso can increase expressivity in a novel way; however, designing a waist joint can be challenging. The weight and motion of components that are not near the axis of rotation require a stiff structure and significant torque to control large torso inertia. Some robot designs avoid the challenges of a waist joint and instead use vertical movement of the torso with a linear actuator (e.g., the PR2 [25]). Designs for humanoid robot waists are diverse and range from the hip joint of a legged humanoid robot (e.g., the NAO [26]) to commercial wheeled mobile robots (e.g., Pepper [7] or Car-O-Bot 4 [27]). Two- and three-DoF waist designs use gravity compensation methods, such as the pulley and tendon system [28] [29]. A single DoF design allows for simpler gravity compensation; while less movement can be produced by a single DoF, it can produce sagittal motion and still enhance expressivity [30]. Quori’s one-DoF torso is gravity-compensated, fits within the shell of the robot, and requires no linkages or pulleys.

D. Robot Mobility for HRI

Some socially interactive robots are non-mobile, often operating in tabletop configurations (e.g., Kiwi [11], LuxAI QT [31], Furhat [14], EMYS [8], Kaspar [32], and Keepon [33]). Most mobile socially interactive robots capable use wheels (e.g., Bandit [9], [34], Mayfield Kuri [35], and Pepper [7]), as they are lower cost, safer, and more easily controlled than legged robots (e.g., Robothespian [36], Hubo [37], and Nao [26]). For systems requiring only simple mobility, affordable wheeled systems have employed the Kobuki base, used for the Turtlebot 2 [38], which features a differential drive and a zero-degree turning radius. Alternatively, omnidirectional designs improve the mobility of the system and simplify navigation by removing path-planning constraints [39], [40]. An affordable holonomic mobile base can be achieved with a differential drive mechanism by adding a turret in which its axis of rotation is offset from the midpoint of the two drive wheels [41] (Figure 8). This configuration is also called the dual-wheel caster-drive mechanism [42] and is used in the Human Support Robot [43], but is otherwise uncommon. Quori’s holonomic mobile base uses this dual-wheel caster-drive mechanism and is optimized for mobility using the design tools from [41].
E. Affordability for HRI

The space of robot platforms for HRI research is growing; however, we surveyed the HRI community and found that no HRI robot platform met all the features researchers identified as important: modularity, openness, appearance, actuation, sensing, behavior range, and affordability. The full HRI community survey results are found in Section III. The PR2 is an open design, modular, and general purpose robot platform that has been highly successful in enabling mobile manipulation robotics research; however, it is not designed for social HRI: it lacks social expressiveness and DoFs useful for body language, its hefty size (180 kg) also makes it intimidating for many real-world users, and its cost (US $400,000) is not affordable to many researchers. The NAO from Softbank Robotics has been a popular choice among HRI researchers. It is comparatively affordable (US $12,000 for one with arms and mobility), and has an appealing design; however, as a closed hardware platform, it is not modular, and it also lacks facial expressiveness. The robot platform that comes closest to meeting the main requirements of the HRI research community is Pepper [7] (also from Softbank Robotics), which costs of approximately US $20,000 [44] (with an annual subscription fee), making it more affordable than the PR2, though more expensive than many tabletop HRI platforms; also, like the NAO, Pepper lacks an expressive face, and features closed hardware and semi-closed software (i.e., closed-source with some open APIs available to interface third-party software) that is robot-specific (i.e., does not work with non-SoftBank robots).

Tabletop platforms offer inherently safe interaction and lower cost, though often in exchange for reduced DoFs and expressivity: some are very simple, such as Keepon [33] and Maki (an affordable (US $5,000 fully assembled) 3D-printed tabletop robot head with mechanical eyes); some are more complex, such as Kiwi (using an adapted Stewart platform for expressive motion and a display for the face [11]); and some are humanoid, such as Kasper. Quori’s affordability is discussed in Section III-C and Section V-A.

F. Software for HRI

Robot control for HRI typically relies on lower-level software to provide standard robotics capabilities, including basic perception and navigation. Robot Operating System (ROS)\(^1\) is the most commonly used middleware software platform in general robotics research and is also frequently employed in HRI research. HRI software is then added to handle the various aspects related to the interaction between the robot, the user(s), and the social context. This includes supporting human and activity recognition, speech and natural language understanding, and robot speech, dialog, body language, and facial expression generation. Intuitive tools can aid content creators (who are often non-programmers, including writers and animators) to create and deploy multimodal conversational content for human-robot interactions [45]. HRI systems use standard software tools for voice-based interactions also used in human-computer interaction, including Amazon Alexa, Dialogflow\(^2\), and Voiceflow\(^3\). Analogously, HRI systems use standard software tools from graphics and gaming for animating digital characters, including Maya\(^4\) and Blender\(^5\). Because these tools were originally developed for other domains, they typically only offer unimodal capabilities (e.g., dialog vs. movement), causing fragmentation in HRI system development. Our goal for Quori is to provide a more holistic and unified multimodal socially interactive robot software infrastructure capable of supporting face-to-face HRI “out of the box” while also being extensible to facilitate HRI research [45].

III. DESIGN METHODOLOGY

HRI is a large and rapidly growing field of research and development, involving a very wide range of researcher interests and needs, presenting an exceptional design challenge for a general-purpose socially interactive robot platform. We employed an iterative community-driven design process to inform the design, hardware, software, and cost of the robot platform. We engaged the broader community of interest through online surveys, conference workshops, and symposia.

A. Research Community Surveys

To inform Quori’s design, hardware, software, and cost, we distributed two online surveys to the HRI community\(^6\). Surveys #1 and #2 were sent out in Fall 2014 and Fall 2015, respectively, and yielded responses from 37 and 50 survey participants, respectively; nearly all responses were received within the first 48 hours, reflecting community interest and engagement. The surveys elicited considerations regarding (1) appearance and actuation, (2) sensing and behaviors, and (3) cost. The survey results constituted the foundation for the Quori platform design, and are summarized below. Relevant results for Surveys #1 and #2 are shown in Appendices A and B, respectively; demographics are shown in Appendix A, Table VII and Appendix B, Table X, respectively\(^7\).

1) Robot Appearance and Actuation Considerations: An interactive cartoon-like character with a hard-shell outer covering was preferred over more human-like, biomimetic, or animal-like appearance (Appendix A, Table IV, Prompts 1-2). Survey respondents indicated that the robot should be separated into two independent parts (Appendix B, Table VIII, Prompt 4): an expressive upper body, and a mobile base (Appendix A, Table IV, Prompts 6-8 and 9, respectively). The expressive upper body should include actuation of the neck (nodding and shaking), face (eyes, eyelids, eyebrows, and lips), two arms, and if possible a spine and shoulder actuation (leaning forward and backward, and shrugging) (Appendix A, Table IV, Prompts 6-8 and Appendix B, Table VIII, Prompt 2). The mobile base should ideally be omni-directional (Appendix A, Table IV, Prompt 9). Human-robot dialog (through robot speech recognition and production capabilities) is the preferred communication interface (Appendix A, Table V,

\(^1\)https://www.ros.org/
\(^2\)https://dialogflow.com
\(^3\)https://www.voiceflow.com
\(^4\)https://www.autodesk.com/products/maya/overview
\(^5\)https://www.blender.org
\(^6\)Via mailing lists, such as HRI-Announcement and robotics-worldwide.
\(^7\)Most respondents identified as young and White/Caucasian.
Prompts 1-2). The overall robot should be 0.71–1.48 meters in height with the expressive upper body atop the mobile base (Appendix A, Table IV, Prompt 3). Respondents requested the robot be gender-neutral and offered options for establishing a gender identity (Appendix A, Table IV, Prompt 4).

Survey respondents requested “cartoonish” physical and social characteristics (Appendix A, Table IV, Prompt 1). In the second survey, respondents did not believe that the mobile base, arms, hands, or chest played a significant role in creating a cartoonish character; instead, they indicated that the use of a retro-projected face, vocal characteristics (e.g., speech), and visual behavior (e.g., expressive face, arm, and body gestures) would be sufficient for customizing a cartoonish character (Appendix B, Table VIII, Prompt 1).

2) Robot Sensing and Behavioral Considerations: According to the survey responses, the robot should support both automated perception and control of abilities that are commonly used in face-to-face social interactions (Appendix A, Table V, Prompts 1-2). The platform should include color (RGB) and depth cameras (RGB+D) for person and object tracking, as well as a microphone array for speech recognition (Appendix A, Table V, Prompt 3). Survey respondents were not consistent with regard to the mobility requirements for human-robot interactions (e.g., proxemics), as some researchers preferred that the robot have the ability to move around the environment and others preferred a static tabletop platform (Appendix B, Table VIII, Prompt 4), reinforcing our choice to separate the upper body and the mobile base.

3) Robot Cost Considerations: Survey respondents were asked what they would expect to pay and what they would be willing to pay for a socially interactive robot platform. Those who requested a mobile platform expected to pay $25,000–$50,000, while those who requested a tabletop platform expected to pay $2,500–$10,000 (Appendix A, Table VI, Prompt 1). However, there was high variability in how much researchers were willing to pay (Appendix A, Table VI, Prompt 2): the maximum was $100,000, selected by only 17% of respondents; 84% of respondents were willing to pay $5,000, which we used as the upper bound on the basic Quori hardware platform cost, ensuring that our implementation would meet the needs and budgets of the research community.

B. Community Engagement Meetings

We presented and discussed Quori prototypes at four research workshops between 2016 and 2018. We hosted two of those workshops: 1) the AAAI 2016 Spring Symposium on “Enabling Computing Research in Socially Intelligent Human-Robot Interaction: A Community-Driven Modular Research Platform”8, and 2) the Robotics: Science and Systems (RSS) 2016 workshop on “A Community-Driven Modular Research Platform for Sociable Human-Robot Interaction”9; the other two workshops were: 3) the AAAI 2017 Fall Symposium on “Artificial Intelligence for Human-Robot Interaction”10, and 4) the 2018 Human-Robot Interaction Conference workshop on “Social Robotics in the Wild”11.

These workshops served to collect community feedback and seek consensus toward Quori’s design, with discussions and insights from attendees that complemented the quantitative data we collected with web surveys. At the workshops, we presented a feasibility analysis of Quori’s modules [46] and progress on each module. Attendees of the Spring Symposium were involved in breakout sessions in which they discussed Quori hardware and software design [47], with key feedback from the 2016 workshops strongly informing decisions about the head size and priorities of the arm DoFs. As an indication of active participation, the Symposium included 20 paper presentations5. The 2017 and 2018 workshops provided input that helped to finalize the panel design (see next section), as well as determine camera placement. The 2018 workshop provided a means of disseminating Quori’s progress [48] before announcing a call for competitive proposals from researchers interested in participating in the Quori Beta Program; the Call for Proposals was an unexpected source of feedback, in which awardees’ desire for a controllable waist DoF justified the additional cost of adding that DoF. The rest of this paper discusses how Quori was designed to address the surveyed and expressed needs of the HRI research community (Fig. 3).

C. Comparing Quori to Relevant Platforms

In this section, we compare relevant existing robot platforms in relation to the requirements from the HRI research community presented in Section III. Table I uses a competitive matrix to highlight the degree to which existing platforms meet the needs we identified in Section III. For this project, we define open and modular hardware and software as the ability for the system to be fully observed, modified, and reconfigured (e.g., via adding or exchanging modules) by a researcher; for example, Quori’s spine is designed to work with custom arm modules or a head module, and all of Quori’s software modules are built using open-source ROS wrappers and communicate via audible ROS interfaces, so they can be readily exchanged with alternative software implementations. The open-source software, documentation, and hardware designs will be released on the project website (https://www.quori.org). The key observation of Table I is that Quori’s design meets the hardware, software, and cost requirements identified in Section III; other relevant HRI platforms either lack affordability or are not as open or modular.

IV. Physical Appearance

Physical appearance is a key attribute for a robot designed for social interaction. We used designer expertise on the team to create an aesthetic that fits the needs of the HRI community as indicated by the community input, allowing modular appearance accessories while retaining stylistic consistency when appropriate. Quori is shown in Fig. 1.

8http://www.quori.org/community-input-meetings (Online proceedings: https://www.aaai.org/Library/Symposia/Spring/ss16-03.php)
9http://www.quori.org/community-input-meetings/#rss-16-1
10http://ai-hri.github.io/2017
11http://socialrobotsinthewild.org
A. Physical Appearance Design

Quori has an “envelope” design. The underlying robot and mechanical systems are clad with a panelized torso, base, and arms. This allows for the physical mechanisms of the robot and the body shell to remain separated. The design includes ranges of motion for each rigid body part with guarantees of no self-collision. Each panel incorporates design considerations for appearance, fit, and finish, as well as ease of disassembly for maintenance and repair. The base panel curvatures are designed to increase the distance between the user and the robot for safety and social proxemics [34].

Quori’s holistic appearance required intentional design. The torso, arms, and base are an identifiable, self-consistent whole: color, seaming, and surface curvature are continuous among the parts. These features address specific community-identified HRI issues of gender, the Uncanny Valley [50], and acceptance. Gender identity is dampened without being generic, the size and appearance are slightly abstract to not mimic human physiognomy and therefore avoid the Uncanny Valley, and the geometry of the robot is meant to be recognizable friendly—we avoided sharp corners and threatening musculature in favor of softly curved surfaces and eased edges that facilitate acceptance. The overall form has large parts with consistent features—the spherical head connects to the rounded torso by a stalk. The out-sized and softly curved forearms meet the torso, arms, and base are an identifiable, self-consistent whole: color, seaming, and surface curvature are continuous among the parts. These features address specific community-identified HRI issues of gender, the Uncanny Valley [50], and acceptance. Gender identity is dampened without being generic, the size and appearance are slightly abstract to not mimic human physiognomy and therefore avoid the Uncanny Valley, and the geometry of the robot is meant to be recognizable friendly—we avoided sharp corners and threatening musculature in favor of softly curved surfaces and eased edges that facilitate acceptance. The overall form has large parts with consistent features—the spherical head connects to the rounded torso by a stalk. The out-sized and softly curved forearms meet the torso by a slender femoral shaft. A geometrically simple waist supports the upper body. This yields a perception of a network of discrete, soft spheroids connected by simple masts, rather than a body that is blob-like or mechanical across its surface.

B. Manufacturing and Mechanical Features

As mentioned in Section IV-A, the panels need to be easily disassembled for repair and robot recharging. There are four panels on the chest (Fig. 2), four constituting the helmet, two for the lower torso (along with two service panel sheets), and one base cover. The front and back chest panels are removable to allow access to the head projector, the speakers, and the arm controllers, as well as to allow for service or inspection of the upper torso. The helmet parts are removable to allow access to the microphone array and RGB+D camera. The two black service panels (cut from 0.8mm haircell ABS sheets for flexibility, style, and durability) on the lower torso allow for quick access to the main power switch, the battery for charging, and the main computer and its peripheral connections. 3D-printed panels form much of the enclosure. They are interchangeable with different colors or materials, and can be easily removed or replaced via magnetic and mechanical alignment and securing features (Fig. 2, right), avoiding visible mechanical buttons or fasteners on the surface. Panel-mounted neodymium 5mm cylinder magnets with rated 1kg pull force provide enough strength to prevent the panels from shaking apart or falling off while allowing them to be easily pulled off by hand.

A significant amount of labor was required for post-processing. The 3D printed panels required approximately 60 person-hours of in-house labor to improve the finish of the parts, remove printing lines and artifacts, and apply a final color and sealing coat. Parts printed with ABS material required light sanding with 400 grit sandpaper to remove printing lines and then painting with a spray primer, followed by white spray-on acrylic gesso, and finally a spray coat of clear varnish. The PLA parts were processed in the same way except for an epoxy coating to fill in the printing layer lines and other artifacts such as glue-seams, since PLA is more difficult to sand than ABS. While ABS was easier and faster to process, we preferred PLA parts, as they were easier to print on our printers and significantly less expensive, in some cases nearly half the cost of ABS parts.

V. Module AND HARDWARE Design

Quori is 1.35m tall, consisting of an expressive upper body attached to an omnidirectional mobile base (Fig. 3), thus, within the desired 0.71–1.48 meters in total height (Appendix A, Table IV, Prompt 3).

Our hardware design consists of three key aspects: (1) validated utility through iterations with the HRI community...
for desired features, (2) affordability and targeted feature inclusion, and (3) longevity of impact through development of modular interface standards. The four hardware modules—head, arms, torso, and base—are described in the following sections, along with their power and sensor systems.

A. Cost, Manufacturing, and Design Analysis

Our primary mechanism for maintaining low cost was through elimination of features. By working with the HRI community to identify the most important hardware capabilities for a socially interactive robot, we maximized the value for HRI research while being cost-considerate. The community provided input via online surveys, hosted workshops, and conference presentations (Section III), the feedback for which guided DoF and feature reduction. Three properties not explicitly requested, but kept at high value, were (1) low audible noise from actuators, (2) fluidity of motion, and (3) physical appearance. Manufacturing processes were chosen appropriate for the prototyping quantity: laser cutting, 3D-printing, and water-jet cutting. Parts costs for each module are shown in Table II and sum to US $6,300. The highlights of the affordable design feature decisions are:

- **Head**: Elimination of mechanical DoF in the head/neck, using the projected face vs. conventional actuators.
- **Arms**: Reduction to two DoF and lightweight arms while enabling modularity and expandability for more DoF.
- **Torso**: Single DoF with gravity compensation.
- **Base**: Minimal DoFs for omnidirectional motion with a cost-effective holonomic mobile base drive design.

Details about each module design are discussed below.

B. Head Design

Quori’s head module uses a retro-projected animated face (RAF) system. It consists of a small projector (115mm x 46mm x 105mm) and a domed mirror to map a projected image onto the inside of specially-coated thin sphere (Fig. 4). These components fit within a compact space approximately the size of an adult human head (200mm diameter) and weigh about 975 grams (excluding the RGB+D camera and microphone array).

The projector is an AXAA P5 (US $290) with key properties: it is rated to provide 300 lumens, last 20,000 hours, and have native 1280x720 HD resolution; however, only ∼132 lumens are available to the spherical surface since the image reflected on the spherical head is a circle inscribed inside the projection rectangle. The projector is affordable (at US$290), has intermediate brightness, and short throw (20cm creates a focused 7.5cm x 12.7cm image). It creates a color image that is visible in most illuminated indoor environments where there is no sunlight saturation (Fig. 4, top).

The mapping of projected images onto the sphere’s surface is not uniform—the resolution is dense near the top of the head and sparse near the neck. The least dense equatorial line is approximately 200 pixels, compared the highest ring, which has over 2,000 pixels; thus, creating expressive faces to be displayed on the spherical surface via a projected image is not trivial. Our mapping algorithm transforms pixels on the sphere to pixels in a 2D image to be sent to the projector. Details of the design and mapping used to project images to Quori’s head can be found in our previous work [51].

---

**TABLE II**

| Subsystem        | Item                        | Qty | Cost [$] | Subtotal [$] |
|------------------|-----------------------------|-----|----------|--------------|
| Arms [2]         | Motor modules               | 4   | 105      | 420          |
|                  | Transmission                | 2   | 170      | 340          |
|                  | Joint Sensors               | 4   | 15       | 60           |
|                  | Structure                   | 2   | 25       | 50           |
| Shoulder Joint   | Joint                       | 2   | 67       | 134          |
|                  | 3D printed gears            | 2   | 30       | 60           |
| Torso            | Motor+Driver+CPU            | 1   | 120      | 120          |
|                  | Structure                   | 1   | 110      | 110          |
| Mobile Base      | Structure                   | 1   | 200      | 200          |
|                  | Motor+Driver+CPU            | 3   | 129      | 387          |
|                  | Electronics                 | 1   | 208      | 208          |
|                  | Laser scanner               | 1   | 100      | 100          |
| Social Sensors   | Camera                      | 1   | 235      | 235          |
|                  | Microphone Array            | 1   | 64       | 64           |
|                  | Speaker system              | 1   | 18       | 18           |
| Head             | Structure                   | 1   | 22       | 22           |
|                  | Mirror                      | 1   | 5        | 5            |
|                  | Painted Globe               | 1   | 13       | 13           |
|                  | Projector                   | 1   | 290      | 290          |
| Panels           | 3D Printed                  | 1   | 2000     | 2000         |
|                  | Access panels               | 2   | 2        | 4            |
| Electronics      | Battery                     | 1   | 75       | 75           |
|                  | Misc components             | 1   | 305      | 305          |
|                  | Onboard computer            | 1   | 1100     | 1100         |
| **Total [$]**    |                             |     | 6320     |              |

---

12Screen Coating: https://store.gooscreen.com/Rear-Projection_p_27.html
The illusion of motion (e.g., head shaking, nodding, and gaze directing) can be produced through projection. Since the robot’s head is a rotationally symmetric sphere with no protruding features (e.g., no nose or ears), head rotation can be simulated by projecting the image of a face rotating on the sphere without requiring additional motors or neck DoFs. Gaze direction can be simulated by coupling animation of the eyes with horizontal rotation of the whole upper torso ($M_r$ in Fig. 8). The waist joint (Fig. 6) may also be useful in supplementing gaze, especially for interactions below or above the neutral gaze of the robot (e.g., for users who are shorter or taller than the robot) or for objects very near or far away. Sensors in the head can be replaced via fasteners on the sensor mounting plate. The camera field of view is discussed in Section V-F and shown in Fig. 12.

C. Arm Design

Gestures are a key part of natural communication in social interaction. Quori’s arm design is affordable, modular, safe, and expandable. The shoulder module (Fig. 5) has two DoFs based on a design by [24]; however, our design differs in the use of 3D-printed bevel gears instead of a capstan cable drive. In addition, to save costs and complexity, we chose to not gravity-compensate the arm, thus, enabling the elbow and arm modules to be changed. The arm is driven by brushless DC motors through a transmission consisting of a friction wheel pair and a timing belt speed reduction (Fig. 5, left). The entire arm module mounts to the spine with fasteners.

Notable features of the arm design include resolution of the joint positions, drive motor abilities, and general safety considerations. The approximate resolutions of the joint position sensors are $0.022^\circ$ and $0.075^\circ$ for the shoulder joint (through the use of magnetic encoders on the output shafts, Fig. 5) and the drive motors, respectively. Access to both motor and shoulder positions allows the system to check for slippage between the friction wheel pair or timing belt stages, as well as perform automatic calibration upon boot-up of the system. The arm motors can produce approximately $0.15\text{Nm}$ and are able to rotate at approximately 16 revolutions per second, resulting in shoulder joint speeds up to $\sim1.2\text{ radians per second}$\textsuperscript{13}. The abduction/adduction DoF has a range of $\pm70^\circ$, and the circumduction DoF is continuous. The arm design is expandable; we designed access for power and/or communication for further joints in the arms (e.g., an elbow), while allowing the arm to rotate continuously. We achieved this via a shoulder joint slip ring with six available wires (rated to 2 A) (Fig. 5, right). As an example, two servo motors can be added as they only use three wires each. More DoF may be added through multiplexing.

We used the following operational safety measures: a torque limit on the drive motors; a low-mass, low-inertia arm mechanism and structure that is safe according to the Head Injury Criterion [23]; and a friction wheel designed to slip in case the motor generates too much torque or the arm is back-driven.

Our primary goals in arm design were to ensure safe and precise yet fluid motion for expressivity, while maintaining affordability. Manipulation (i.e., carrying some payload or applying forces to the environment) was explicitly not the goal of our design; thus, we used light-weight limbs and IQ Control’s position controlled, direct drive, and brushless servo motors\textsuperscript{14}. Arms that would be expected to lift, push, or pull would need structural stability that typically leads to heavier and more expensive designs. Furthermore, heavier arms require larger and thus more expensive motors to move. Lower-cost motors or servos could be used at the expense of precision for the case of brushless DC motors [52].

D. Torso and Waist Design

Quori’s torso module not only supports the arms and head (Fig. 6, left), but also has one DoF to lean forward and backward (Fig. 6, right). This design also minimizes acceleration-induced swaying generated during the motion of the mobile base, leading to fluid, natural, and appealing tunable motion. The batteries and onboard computer are stored in the torso.

The spine allows for easy attachment of additional custom hardware, such as arms or a head. A new head module can be attached to the spine using the provided mounting holes. The arms have similar mounting possibilities—shelves/ledges can be added to the spine for additional accessories, such as sensors, tablets, trays, container mountings, etc.

Considerable space is allotted for batteries and a computer\textsuperscript{15}: 17 cm x 15 cm x 21 cm and 20 cm x 20 cm x 7 cm, respectively. Currently, the battery bay space fits a 40-ampere-hour sealed lead acid battery that powers the whole robot. While many options exist for small form factor computers, we have ensured sufficient space for a computer with compu-

\textsuperscript{13}This is based on the motor properties at 12 Volts operation.

\textsuperscript{14}http://iq-control.com

\textsuperscript{15}Quori currently ships with a nuc8i7hv: Intel\textsuperscript{©} Core\textsuperscript{™} i7-8809G Processor with Radeon\textsuperscript{TM} RX Vega M GH graphics (8M Cache, up to 4.20 GHz)
Fig. 5. Left: CAD model of the arm module. Center: A sectional view of the compact differential transmission. Right: Section view of the arm differential highlighting how the torque is transferred while allowing 12 wires to be available to the arm with continual shoulder rotation.

Fig. 6. Left: Upper torso and waist hardware overview. Right: Extreme positions the robot achieves by bowing forward 30 degrees or leaning back 15 degrees. Mechanical limitations on the positions prevent self-collision.

Affordable actuation of the waist can be achieved with a counterbalance metronome design (Fig. 7, left). This design leverages the mass of the robot’s battery to offset the moment of the upper body of the torso, head, sensors, and arms. The moment that needs to be balanced changes, as the balance depends on the position of the arms $\phi_a$ which may be moving. Fig. 7 shows the torque required to hold the torso at its max bow position as the arms rotate in the plane. The effect of the extra counter-mass, the battery, and an ideal tuned counterbalance design is shown in Fig. 7. In its most difficult bowing position, the waist experiences a 16-Nm moment without counterbalancing (Fig. 7, purple line). It is very challenging to find a motor with this capability, that is also small enough to fit in the required space and is affordable. Instead, with proper counterbalancing, the battery and an extra 6 kg (Fig. 7, blue line), the peak torque requires less than 2-Nm. The major drawback to this counterbalance design is the increased inertia of the torso. The waist does not need to move very fast—less than $1 \text{rad s}^{-1}$—nor accelerate faster than $1 \text{rad s}^{-2}$, which leads to a max accelerating torque of about 2.5 Nm; for reference, the max required static holding torque of the final design is approximately 2.5-Nm.

To realize the counterbalance design, we used the model as a starting point (Fig. 7, blue line), and manually tuned the final counterbalance configuration during construction. The battery bay structure (made from steel) provides both a stiff structure to support the battery and contributes about half of the needed 6 kg counter-mass. Steel plates and bars underneath the battery (Fig. 6, left) allow for high-resolution calibration of the counter mass. Proper calibration shows gauge values below 3.0 Nm, a torque achievable by our lower-cost (i.e., less than US $100), low-profile actuator\(^\text{18}\).

2) Waist Transmission Design: A non-backdrivable transmission was chosen to minimize the energy required to bow so holding positions requires zero energy. An optional locking pin feature allows for the torso motion to be locked for shipping or if waist actuation is not desired.

Next, we present our approach to designing the single DoF waist which involves gravity compensation. The transmission design is also discussed.

1) Gravity Compensation Design and Tuning of Waist Joint: Robot motion is often caricatured as jerky with overshoot; for example, a person pretending to be a robot might exaggerate leaning backwards as they start to walk forward, then sway forward and back just as they stop walking, as a cantilevered stick might do as a damped oscillator. Avoiding these types of motions typically requires expensive, strong motors and precise feedback. Alternatively, adding mass to shift the center of mass (CoM) can change this behavior. A CoM below the axis of rotation causes the torso to lean forward during acceleration (opposite to the prototypical robot caricature motion), while a CoM at the axis of rotation reduces the motion.

\(^{16}\)https://ark.intel.com/content/www/us/en/ark/products/126143/intel-nuc-kit-nuc8i7hvk.html

\(^{17}\)https://developer.nvidia.com/embedded/jetson-tx1-developer-kit

\(^{18}\)This actuator—a window motor—is also quiet, especially when compared to small but high speed motors with larger gearing.
Holding Torque as a Function of Flexion Arm Position

Fig. 7. Left: Model used for tuning the CoM of the torso and waist actuator torque. The masses are separated into the upper body mass $m_{UB}$ (head and arm transmission), the arm link mass $m_a$, and the lower body mass $m_{LB}$ (battery and counter masses). Right: Maximum waist holding torque curves used to select a starting counter mass (CM) configuration for the torso. The lines are produced by simulating the arms flexion, $\phi_a$, in order to produce the maximum waist torque, $\ddot{\theta}_{\text{waist}}$, to hold the most difficult position of bowing forward.

Fig. 8. Quori base’s 3-DoFs produce holonomic movement in the ground plane. The axes driven by the three actuators are highlighted. $M_1$ and $M_2$ are drive motors and are equivalent to a differential drive. $M_T$ is driven by the turret motor. The axis of $M_T$ is distance $a$ from the $M_1$ and $M_2$ axis.

Friction damper pads on each side of the battery bay (Fig. 6, left) add dampening to the waist motion. The pads consist of soft foam and a PTFE sheet fastened to the battery bay, which push against an ABS plastic sheet. This design compensates for gear backlash and compliance in the structure and actuator and greatly simplifies smooth control. The damper increases the torque required to rotate the torso, but this effect was measured empirically to bring the waist motor torque to no more than 3.0 Nm, which met our goal.

E. Mobile Base Design

Quori’s holonomic mobile base has three motors (Fig. 8). Two casters serve to support the weight of the robot and increase the support polygon along with two driven wheels. The torso provides electrical power to the base via a slip ring between the turret and differential drive base (Fig. 8, right). Communication and control between the base and torso electronics occurs via a USB connection through a second concentric slip ring (Fig. 8, right). Extra space and USB ports are available for a laser scanner or camera in the lower section of the base.

Fig. 9. The laser scanner’s FoV. The sensor, marked as a red dot, is offset 100mm to maximize coverage. Sensor blind spots are shaded in gray. The outer circle shows the sensor’s 8-meter radius about the robot, marked as a yellow circle.

Quori’s base measures 48cm in diameter and 20cm in height, and can traverse any indoor floor that complies with the 2010 Americans with Disabilities Act (ADA) Standards for Accessible Design. This includes traversing over 0.635cm bumps (ADA 303.2), 1.27cm floor gaps (ADA 302, 407.4.3), and 1:12 inclines (ADA 405.2). The base has max speeds of 0.6 m/s in a straight line and $\pi$ radians/s rotation of the turret. The design tool presented in [41] verified the base parameters will achieve the desired maximum rotational and translational velocities given the motor limits. The positioning of the laser ranging sensor near the perimeter of the base maximizes the laser field-of-view (FoV) (Fig. 9). Finally, the design allows the base to act as a standalone module independent of the upper-body humanoid torso, should the user desire applications with either half alone.

The choice of design for the holonomic base ensures notable cost-reduction over other options; for example, with three motors, our base uses fewer actuators than other designs that require four or more motors [39]. Other holonomic designs may involve using an omniwheel or additional motors; however, they often suffer from performance drawbacks, such as vibration or complexity [40]. The manufacturing of Quori’s base is made more affordable by using laser-cut parts from sheet ABS and commercial off-the-shelf parts for the majority of the components, requiring only two machined parts to mount the motor to the base and the motor shaft to the wheel (Section V-A).

F. Power and Electronic Design

1) Power System: Quori’s power system was designed to operate untethered with the use of a 12V 40AH battery, which is also used as a counter balance. The battery is a Sealed Lead Acid (SLA) Absorbent Glass Mat (AGM) chemistry battery that is affordable (compared to lithium-based batteries) and stable without losing charge over long periods and has minimal risk of fires or acid spill. Shipping is also simplified as the battery only requires the use of a sticker stating “non-spillable battery” instead of additional regulations or costs. A potential downside of SLA batteries is the low energy density and high mass; however, we take advantage of this as a counterbalance, as discussed in Section V-D.

Most of the robot’s subsystems are 12V-based; the only significant voltage switching occurs in a DC-to-AC power...
inverter, which allows for a main computer (e.g., a laptop) to be used on Quori without requiring the selection or design of an additional DC-to-DC voltage regulator. A simplified diagram of components is presented in Fig. 10. The robot can also run in a tethered mode when not mobile.

2) Electronics Overview: Each of the sensors and main components connect via standard connectors and communication interfaces for simplicity, modularity, and potential future reconfiguration. The main connection type is USB with all connections using USB 2.0, except for the USB 3.0 RGB+D camera (Quori’s default PC has multiple USB 3.0 ports for future upgrades); Fig. 11 shows the components and connection types. HDMI transmits the head image data, allowing for future modifications. A stereo audio cable from the PC 3.5-mm audio port transmits audio to the chest speakers. A USB and HDMI port are accessible from the back panel of the robot for programming and debugging (Fig. 2, left).

3) Sensors for Social Interaction: Stereo speakers mount to a shelf on the upper torso behind the chest panel allowing for ample volume (60 dB SPL at 3 meters). Slots in the helmet provide the illusion of sound being produced in the head.

A ReSpeaker 2.0 four microphone array mounts to the top panel of the helmet for sound localization and speech recognition (Fig. 4). To test the sensor placement effectiveness for speech recognition, we performed a word error rate (WER) experiment using ten prerecorded English phrases produced from a hardware speaker at three distances (0.1m, 0.5m, and 1.0m), yielding an average WER below 13%. Additional or replacement microphones can be mounted inside the helmet.

An RGB+D camera mounts atop the robot’s head to the sensor plate and fits inside the helmet (Fig. 4). The position of the camera gives the robot a $60^\circ \times 49.5^\circ$ FoV that follows the robot’s gaze direction. The camera is also plainly visible, which helps to set reasonable social expectations for what is in its FoV (Fig. 12); this FoV can be manually adjusted to $\pm 25^\circ$ (Fig. 12, right). The current camera is an Orbec Astra Mini$^{19}$ in a DuriPOD Case$^{20}$, which is 120mm x 37.5mm x 32.5mm, but can be easily replaced.

VI. SOFTWARE ARCHITECTURE

Quori has two main software categories: (1) low-to-mid-level, including core control of each module (actuation and sensing); and (2) high-level social interaction software (animation and dialog tools), as shown in Fig. 13.

---

$^{19}$http://shop.orbecute.com/Astra-Mini_p-40.html
$^{20}$http://shop.orbecute.com/DuriPOD_p-47.html
A. Low- and Mid-Level Software and Networking

The low- and mid-level software, written for ROS [53], handles low-level actuator voltage commands, communication between microcontrollers and the main PC, and basic control and safety features. Low-level control is handled by microcontrollers, and middle-level control is handled by the onboard PC. The microcontrollers run low-level control independent of the PC. This means safety features (e.g., timers and position limits) are not affected by potential software errors or PC issues. The head module runs at the high level and is not discussed in this section. Commands can be sent to the robot and the status can be monitored wirelessly at 2.4 or 5GHz. Quori’s low- and mid-level software capabilities adhere to the ROS developer’s guidelines to provide an idiomatic experience for the HRI research community.

B. High-Level Social Robotics Software

In collaboration with Semio, we integrated high-level software packages that provide a set of socially interactive behavior APIs and developer tools that can be used by HRI researchers in a platform-agnostic way, analogous to those already in use in commercial products. This software aims to facilitate exploration of advanced topics in HRI without the technical burden of developing and maintaining social behavior primitives on an institutional or individual level. In particular, we provide software packages that enable verbal HRI (via speech recognition and generation), as well as nonverbal HRI (via pointing gesture recognition, and attention recognition and generation); in addition, we provide intuitive software tools to enable HRI research teams to rapidly create and deploy multimodal conversational content on Quori.

1) Speech Generation and Recognition: Speech generation and speech recognition were the most desired software modules by HRI researchers in our first survey. While speech generation and recognition are not novel technologies, they are essential in much of social HRI and are therefore key for a standardized platform for HRI research.

2) Attention Generation and Recognition: The attention generation software module uses a mixture of three models of human visual attention to produce robot eye gaze behaviors for face-to-face HRI. These models include (1) a neurobiological model of the human visual attention system [54], (2) a conversational gaze model for multi-party interactions [55], and (3) a functional gaze model governed by the need for the robot to track body features of the human user to conduct a successful interaction [56].

The attention recognition software module (Fig. 14) provides a probabilistic estimate of targets visually attended to by a human user based on a toe-to-head anatomically correct estimate of human attention (i.e., from human body range of motion to the distribution of photoreceptors in the human eye); distant eye gaze tracking might be pursued in future implementations of this module. The mean error in head pose estimation was determined to be 2.8°–6.0°, depending on the data set and environmental conditions. Our solution reduces the overall image search space for faces by 99% and subsequent attention recognition achieves a frame-rate of 100+ frames per second, which was a limitation of the camera hardware being used (i.e., the highest framerate camera we had available). Furthermore, our software file size is compact, only 110 MB (or up to 160 MB with examples).

3) Pointing Gesture Recognition: The static referential pointing gesture recognition software module considers the human kinematics of reaching and visual servoing (the relationship between arm kinematics and “simulated perception” of visual attention toward an object). The implemented solution yields a mean error in pointing target estimation of 0.17m; it is worth noting that this error is a function of the error rate in human joint tracking from the sensor suite used during testing and not a result of the algorithm.

4) Animation and Dialog Tools: During the community-driven design process, we learned that HRI research groups are often composed of both programmers and non-programmers. Teams interested in using robots for HRI experiments should be able to animate socially interactive robot behaviors and synchronize those behaviors with speech without needing to program, thereby lowering the development barrier for non-programmers. SoftBank Robotics offers similar software tools in their Choregraphe Suite and Animation/Chat.

21http://wiki.ros.org/DevelopersGuide
22https://semio.ai
23https://developer.softbankrobotics.com/nao6/naoqi-developer-guide/choregraphe-suite
24https://developer.softbankrobotics.com/pepper-qisdk/tools/animation-editor
25https://developer.softbankrobotics.com/pepper-qisdk/tools/chat-editor
Editors: however, those tools are specific to their NAO and Pepper robot platforms, and are not extensible to other robots. To support the Quori community, we developed two web-based tools—one for keyframe animation of expressive robot movement (similar to Maya or Blender; Fig. 15) and one for authoring human-robot dialog (similar to Voiceflow or Dialogflow)—for both programmer and non-programmer HRI researchers, to enable them to rapidly create and deploy socially interactive robot applications that rely on human-robot speech and body language. These tools operate within a web browser and do not require content creators to be familiar with or use Linux or ROS, lowering the barrier to entry for socially interactive robot application development. These tools are cross-platform, allowing the resulting applications to be integrated with, deployed to, and executed on a very broad range of platforms in addition to Quori. These tools and open-source wrappers around their APIs will be open-sourced for non-commercial applications via the web portal on the Quori website. All created content will be available for public use, viewing, copying, and modifying (similar to Wikipedia, but for conversational content) to support replicability in research.

VII. TESTING AND ROBUSTNESS

Simplicity of the mechanical design and transmissions are key to prevent failures and reduce low-level testing needs. To ensure robustness, we tested the basic function of the robot in stages, as individual modules and also as a fully integrated system in a laboratory setting. We examined performance metrics for each module, as well as some life-cycle tests. The culminating platform life-cycle test was a deployment in a public setting running typically for 14 hours a day, 7 days a week, for 6 months as part of an exhibit at the Philadelphia Museum of Art.

Table III lists the relevant specifications that are verified for each module before a robot is shipped. One of Quori’s unique mechanical designs is its arm transmission, which features a friction wheel pair that acts as a clutch and a speed reduction (Fig. 5). This part required dedicated testing and a redesign. We describe the process to inform future use of this design.

The material choices for the friction pair were originally MDF and urethane following the design in [24]; however, after life cycle testing, the longevity of the friction pair proved too short and the urethane roller failed via wear, being unable to transfer sufficient torque (less than 60% of its original potential). A urethane roller and aluminum wheel were the final materials chosen; this combination reduced wear on the urethane roller and maintained sufficient torque transfer ability after 70 hours of tested motion.

B. Deployment at Philadelphia Museum of Art

Quori was installed at the “Designs for Different Futures” exhibit at the Philadelphia Museum of Art from October 2019 to March 2020, and attracted over 183,000 visitors (Fig. 16). The curators chose Quori as an example of “a robot of the future”. Quori autonomously interacted with visitors and reacted with animated facial expressions, arm gestures, torso bowing, and tracking guests by rotating the base while staying fixed on the platform. An external monitor showed visitors a sample of what Quori could see along with a kinematic overlay of identified humans’ limbs in its FoV. This is the first application of Quori as a socially interactive robot platform. More applications and evaluations will follow as the Quori platforms are distributed to researchers.

Highlights of Quori’s reactions included waving “hello” to visitors who entered its FoV, dancing to gain attention, and bowing to greet visitors. The robot attempted to stay engaged with the closest visitor by tracking them with the base turret actuator. Quori then attempted to mirror the visitor’s arm movements (Fig. 16, right). If no one was in Quori’s FoV, it returned to an “sleep” position by rotating its torso to a center position and leaning over with its arms hanging and the face switching to a loading animation. The interaction with visitors was based on a finite state machine with action cool-downs and was designed to feel spontaneous without being repetitive.

Quori’s hardware performed well overall. Quori was active 7 days a week, from 8am–10pm. A power supply and power strip powered the robot, and its system turned itself on and off each...
day to reduce strain. Museum staff reported any issues daily, which mostly consisted of synchronization with the external monitor. Specifically, the museum required a total of 11 visits to address system issues; of those issues, 6 were related to motor drive code failure, 2 to other hardware issues, and 3 were software timing issues due to syncing multiple computers and monitors. The motor driver firmware was updated to prevent further failures. A weekly maintenance visit included those repairs as well as inspections of the whole system.

These visits led to a few design changes that allow for quicker and easier access to the robot, including: (1) replacing key fasteners on the torso panels with thumbscrews instead of socket heads, and (2) selecting a thinner material for the base service panel, which allowed for simpler removal and mounting. Quori’s hardware and software were able to perform well over a long-term installation in a public setting, and the experience enabled us to identify weaknesses and improve Quori’s hardware and software subsystems.

VIII. CONCLUSION AND FUTURE WORK

This paper presents Quori, an affordable socially interactive robot platform comprised of an upper-body humanoid with a rear-projection head, a one-DoF waist, and two gesturing arms, along with a holonomic mobile base. We describe the features and utility of the four modules and identify the decisions resulting in an affordable design. The modules were designed to produce an aesthetically pleasing whole that meets the requirements identified by the HRI community in surveys and workshops.

The ten Quori robots have been awarded to a diverse group of researchers across the United States with many multidisciplinary and cross-laboratory researchers. The Quori website includes information about the awardees, their research, and updates. Additional dissemination is planned through a Quori simulation using the Gazebo 3D robot simulator.

Further evaluation and assessment of the Quori platform will be possible after those ten research groups have received the robots and have had the time to develop new capabilities and perform user studies. In future work, we will perform a full-scale assessment of Quori’s effectiveness in meeting the needs of the HRI research community.

IX. ACKNOWLEDGEMENTS

This work is supported by the National Science Foundation under Grant No. CNS-1513275 and CNS-1513108.

REFERENCES

[1] M. A. Goodrich and A. C. Schultz, *Human-robot interaction: a survey*. Now Publishers Inc, 2008.
[2] C. Breazeal, “Designing sociable robots, mit press,” Cambridge, MA, 2002.
[3] M. J. Mataric and B. Scassellati, “Socially assistive robotics,” in *Springer handbook of robotics*. Springer, 2016, pp. 1973–1994.
[4] D. Feil-Seifer and M. J. Mataric, “Defining socially assistive robotics,” in *9th International Conference on Rehabilitation Robotics*. 2005. ICORR 2005. IEEE, 2005, pp. 465–468.
[5] R. Mead, A. Atrash, E. Kaszubska, A. St. Clair, J. Grecek, C. Clabaugh, B. Kohan, and M. J. Mataric, “Building blocks of social intelligence: Enabling autonomy for socially intelligent and assistive robots,” in *Association for the Advancement of Artificial Intelligence Fall Symposium on Artificial Intelligence and Human-Robot Interaction*, ser. AI-HRI, Arlington, Virginia, 2014.
[6] J. Ekman, *Facial Expressions*. John Wiley & Sons, Ltd, 2005, ch. 16, pp. 301–320. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/0470013494.ch16
[7] A. K. Pandey and R. Gelin, “A mass-produced sociable humanoid robot: Pepper: The first machine of its kind,” *IEEE Robotics & Automation Magazine*, vol. 25, no. 3, pp. 40–48, 2018.
[8] J. Edzesierski, R. Muszyński, C. Zoll, A. Oleksy, and M. Frontkiewicz, “Ensys—emotive head of a social robot,” *International Journal of Social Robotics*, vol. 5, no. 2, pp. 237–249, 2013.
[9] J. Fasola and M. J. Mataric, “A socially assistive robot exercise coach for the elderly,” *Journal of Human-Robot Interaction*, vol. 2, no. 2, pp. 3–32, 2013.
[10] J.-H. Oh, D. Hanson, W.-S. Kim, Y. Han, J.-Y. Kim, and I.-W. Park, “Design of android type humanoid robot albert hubo,” in *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2006, pp. 1428–1433.
[11] E. Short, D. Short, Y. Fu, and M. J. Mataric, “Sprite: Stewart platform robot for interactive tabletop engagement. department of computer science, university of southern california,” *Tech Report*, 2017.
[12] F. Delaunay and T. Belpaeme, “Towards retro-projected robot faces: an alternative to mechatronic and android faces,” in *Robot and Human Interactive Communication, 2009. RO-MAN 2009. The 18th IEEE International Symposium on*. IEEE, 2009, pp. 306–311.
[13] B. Pierce, T. Kuratate, C. Vogl, and G. Cheng, ““mask-bot 2i”: An active customisable robotic head with interchangeable face,” in *Humanoid Robots (Humanoids), 2012 12th IEEE-RAS International Conference on*. IEEE, 2012, pp. 520–525.
[14] S. A. Moubayed, G. Skantze, and J. Beskow, “The furhat back-projected humanoid head—lip reading, gaze and multi-party interaction,” *International Journal of Humanoid Robotics*, vol. 10, no. 01, p. 1350005, 2013.
[15] “Socibot: The robot that can wear any face.” [Online]. Available: https://www.engineeredarts.co.uk/socibot/
[16] F. Delaunay and T. Belpaeme, “Refined human-robot interaction through retro-projected robotic heads,” in *2012 IEEE Workshop on Advanced Robotics and its Social Impacts (ARSO)*, 2012.
[17] A. S. Clair, R. Mead, and M. J. Mataric, “Investigating the effects of visual saliency on deictic gesture production by a humanoid robot,” in *2011 RO-MAN*. IEEE, 2011, pp. 210–216.
[18] J. Grecek, E. Kaszubska, A. Atrash, and M. Mataric, “Graded cuing feedback in robot-mediated imitation practice for children with autism spectrum disorders,” in *The 23rd IEEE international symposium on robot and human interactive communication*. IEEE, 2014, pp. 561–566.
[19] D. Tolani and N. I. Badler, “Emys—emotive head of a social robot,” *International Journal of Social Robotics*, vol. 10, no. 04, pp. 393–401, 1996.
[20] “Misty ii product page.” Apr 2020. [Online]. Available: https://www.mistyrobotics.com/products/misty-ii/
[21] A. Brisen, C. Safos, A. Lockerd, J. Vice, and C. Lathan, “The cosmobot system: Evaluating its usability in therapy sessions with children diagnosed with cerebral palsy;” *Retrieved on*, vol. 3, no. 25, p. 13, 2005.
[22] M. J. Sobrepersa, V. G. Lee, and M. J. Johnson, “The design of lil’flo, a socially assistive robot for upper extremity motor assessment and rehabilitation in the community via telepresence,” *medRXiv*, 2020.
Simon Kim is a licensed architect and researcher in applied sciences within the disciplines of architecture and urbanism. Director of the Immersive Kinematics Lab and Principal of Ibañez Kim, his projects are funded by the National Science Foundation (NSF), Pew Center for Arts and Humanities, Social Sciences and Humanities Research Council (SSHRC), Canadian Heritage Foundation. He is also supported by residencies and fellowships at Autodesk, RAIR Philadelphia, MIT, and the Seoul Biennale. His graduate courses have partnerships with Seoul National University, Opera Philadelphia, Tyler School of Art, with sponsorship from Heerim Architects and Planners. Simon has published and presented his work globally at conferences such as IEEE, AI-HRI, ACADIA, eCAADe, ACSA, SmartGeometry, and for clients such as the MIT Museum, the Philadelphia Museum of Art, Harvard University, the Smithsonian, MoMA, Slought Foundation.

Maja Matarić is Chan Soon-Shiong Professor of Computer Science, Neuroscience, and Pediatrics at USC, founding director of the Robotics and Autonomous Systems Center, and lead of the Viterbi K-12 STEM Center. Her PhD and MS are from MIT, and her BS is from Kansas University. She is Fellow of AAAS, IEEE, AAAI, and ACM, recipient of the Presidential Award for Excellence in Science, Mathematics & Engineering Mentoring, Anita Borg Institute Women of Vision for Innovation, NSF Career, MIT TR35 Innovation, and IEEE RAS Early Career Awards, is active in K-12 outreach, and authored “The Robotics Prime” with MIT Press. A pioneer of the field of socially assistive robotics, her research is developing human-robot interaction for convalescence, rehabilitation, training, and education for children with autism spectrum disorders, stroke and traumatic brain injury survivors, and individuals with Alzheimer’s Disease. She is also co-founder of Embodied, Inc.

Mark Yim is the Asa Whitney professor of Mechanical Engineering who joined the University of Pennsylvania in 2004. He is the director of the GRASP Lab, the oldest robotics research lab in the country established 1980. His research group designs and builds a variety of electromechanical hardware. Demonstrations range from a humanoid robot on display at the Philadelphia Museum of Art to transforming robots that can change their shape to the smallest self-powered flying robot in the world. Honors include the Lindback Award for Distinguished Teaching (UPenn’s highest teaching honor); induction to MIT’s TR100 in 1999; induction to the National Academy of Inventors in 2018. He has over 250 publications and patents. He has started two companies, one in robotics and one medical device company making a steerable needle.
APPENDIX A
SURVEY #1

The first web survey was sent to 37 HRI researchers in Fall 2014. Result ratios below 37 represent the prompt includes skipped response(s). Tables IV-VII present the prompts, responses, and results from the first survey that are related to Quori’s design. Table IV presents the survey data for robot appearance and actuation considerations in support of Section III-A1. Table V presents the survey data for robot sensing and behavior considerations in support of Section III-A2. Table VI presents the survey data for robot cost considerations in support of Section III-A3. Table VII presents demographic data of Survey #1.

### TABLE IV
SURVEY #1: ROBOT APPEARANCE AND ACTUATION CONSIDERATIONS

| Prompt                                                                 | Responses   | Result       |
|------------------------------------------------------------------------|-------------|--------------|
| (1) Please select a preferred appearance of the HRI robot platform.    | Mechanical  | 33% (12/36)  |
|                                                                         | Cartoon     | 31% (11/36)  |
|                                                                         | Other       | 14% (5/36)   |
|                                                                         | Creature    | 11% (4/36)   |
|                                                                         | Human       | 6% (2/36)    |
|                                                                         | No Preference | 6% (2/36)  |
| (2) Please select a preferred outer covering of the HRI robot platform.| Hard        | 50% (18/36)  |
|                                                                         | Soft        | 28% (10/36)  |
|                                                                         | No Preference | 11% (4/36)  |
|                                                                         | Fabric      | 6% (2/36)    |
|                                                                         | Fuzzy       | 6% (2/36)    |
| (3) Please select height preferences for the HRI robot platform.       | Minimum     | 0.71±0.36 m  |
|                                                                         | Preferred   | 1.14±0.35 m  |
|                                                                         | Maximum     | 1.48±0.33 m  |
| (4) Please select a preferred gender of the HRI robot platform.        | Customizable | 49% (18/37) |
|                                                                         | Neutral     | 41% (15/37)  |
|                                                                         | No Preference | 5% (2/37)   |
|                                                                         | Female      | 3% (1/37)    |
|                                                                         | Male        | 3% (1/37)    |
| (5) Please rank the following actuation capabilities in order of preference for the HRI robot platform. | Head | 6.00 |
|                                                                         | Face        | 5.00 |
|                                                                         | Arms        | 4.58 |
|                                                                         | Mobile Base | 4.08 |
|                                                                         | Trunk/Spine | 3.28 |
|                                                                         | Hands       | 3.08 |
|                                                                         | Shoulders   | 2.00 |
| (6) Please rank the following face actuation capabilities in order of preference for the HRI robot platform. | Eyes | 6.43 |
|                                                                         | Eyelids     | 4.91 |
|                                                                         | Eyebrows    | 4.69 |
|                                                                         | Lips        | 4.56 |
|                                                                         | Jaw         | 3.82 |
|                                                                         | Ears        | 2.79 |
|                                                                         | Nose        | 1.41 |
| (7) Please rank the following head actuation capabilities in order of preference for the HRI robot platform. | Nodding | 3.59 |
|                                                                         | Shaking     | 2.82 |
|                                                                         | Tilting     | 2.45 |
|                                                                         | Squashing / Stretching | 1.18 |
| (8) Please rank the following trunk/spine actuation capabilities in order of preference for the HRI robot platform. | Leaning Forward / Back | 3.41 |
|                                                                         | Leaning Left / Right | 2.57 |
|                                                                         | Twisting Left / Right | 2.42 |
|                                                                         | Squashing / Stretching | 1.77 |
| (9) Please rank the following mobile base actuation capabilities in order of preference for the HRI robot platform. | Omni-drive | 2.64 |
|                                                                         | Diff-drive  | 2.06 |
|                                                                         | Legs        | 1.34 |
### TABLE V
**Survey #1: Robot Sensing and Behavioral Considerations**

| Prompt                                                                 | Responses                  | Result |
|-----------------------------------------------------------------------|----------------------------|--------|
| (1) Please rank the following autonomous behavior generation capabilities in order of preference for the HRI robot platform. | Speech & Dialog: 7.63      |        |
|                                                                       | Eye Gaze & Attention: 7.37 |        |
|                                                                       | Turn-taking & Back-channeling: 6.06 |        |
|                                                                       | Environmental Navigation: 6.00 |        |
|                                                                       | Social Navigation: 5.67   |        |
|                                                                       | Gestures: 5.53            |        |
|                                                                       | Emotion: 4.89             |        |
|                                                                       | Prosody: 4.89             |        |
|                                                                       | Object Manipulation: 4.51 |        |
|                                                                       | Touch: 3.49               |        |
| (2) Please rank the following autonomous behavior recognition capabilities in order of preference for the HRI robot platform. | Speech: 8.75              |        |
|                                                                       | Eye Gaze & Attention: 7.44 |        |
|                                                                       | Person Identification: 7.06 |        |
|                                                                       | Gesture: 6.89             |        |
|                                                                       | Object Identification: 6.15 |        |
|                                                                       | Mapping & Localization: 6.06 |        |
|                                                                       | Social Navigation: 5.89   |        |
|                                                                       | Turn-taking & Back-channeling: 5.86 |        |
|                                                                       | Emotion: 5.44             |        |
|                                                                       | Touch: 3.91               |        |
|                                                                       | Prosody: 3.42             |        |
| (3) Please rank the following sensing capabilities in order of preference for the HRI robot platform. | RGB+Depth Camera(s): 5.53 |        |
|                                                                       | Microphones: 4.81         |        |
|                                                                       | Distance: 3.49            |        |
|                                                                       | Tactile: 2.86             |        |
|                                                                       | RGB-only Camera(s): 2.32  |        |
|                                                                       | Depth-only Camera(s): 2.23 |        |

### TABLE VI
**Survey #1: Robot Cost Considerations**

| Prompt                                                                 | Responses                  | Result |
|-----------------------------------------------------------------------|----------------------------|--------|
| (1) How much would you expect to pay for the HRI robot platform.      | $25K-50K: 25% (9/36)       |        |
|                                                                       | $5K-10K: 22% (8/36)        |        |
|                                                                       | $2.5K-5K: 19% (7/36)       |        |
|                                                                       | $10K-25K: 14% (5/36)       |        |
|                                                                       | $100K-250K: 8% (3/36)      |        |
|                                                                       | $1K-2.5K: 6% (2/36)        |        |
|                                                                       | $50K-100K: 6% (2/36)       |        |
|                                                                       | <$1K: 0% (0/36)            |        |
|                                                                       | >$250K: 0% (0/36)          |        |
| (2) How much would you be willing to pay for the HRI robot platform.  | $1K-2.5K: 17% (6/36)       |        |
|                                                                       | $5K-10K: 17% (6/36)        |        |
|                                                                       | $10K-25K: 17% (6/36)       |        |
|                                                                       | $25K-50K: 17% (6/36)       |        |
|                                                                       | $50K-100K: 17% (6/36)      |        |
|                                                                       | $2.5K-5K: 11% (4/36)       |        |
|                                                                       | <$1K: 6% (2/36)            |        |
|                                                                       | $100K-250K: 0% (0/36)      |        |
|                                                                       | >$250K: 0% (0/36)          |        |
TABLE VII
SURVEY #1: DEMOGRAPHICS

| Prompt                                                                 | Responses | Result          |
|-----------------------------------------------------------------------|-----------|-----------------|
| (1) What is your age (in years)?                                      | 25-34     | 53% (18/34)     |
|                                                                        | 18-24     | 18% (6/34)      |
|                                                                        | 45+?      | 15% (5/34)      |
|                                                                        | 35-44     | 15% (5/34)      |
| (2) What is your gender?                                              | Female    | 50% (17/34)     |
|                                                                        | Male      | 47% (16/34)     |
|                                                                        | Did not specify | 3% (1/34)  |
| (3) What is your ethnicity? (Please select all that apply.)           | White / Caucasian | 76% (26/34) |
|                                                                        | Asian or Pacific Islander | 21% (7/34) |
|                                                                        | Black or African American | 3% (1/34)  |
|                                                                        | Prefer not to specify | 3% (1/34)  |
| (4) How much do you know about the following topics?                  | Human-Robot Interaction | 6.38        |
|                                                                        | Robotics  | 6.24           |
|                                                                        | Artificial Intelligence | 5.85       |
|                                                                        | Psychology | 4.65        |
|                                                                        | Signal Processing | 4.59        |
|                                                                        | Animation (2D/3D) | 3.97        |
|                                                                        | Anatomy   | 3.18           |
|                                                                        | Anthropology | 2.91        |
| (5) Would you be willing to provide a letter of support for the       | No        | 50% (15/30)    |
| development of the HRI robot platform?                                | Yes       | 50% (15/30)    |
The second web survey was sent to 50 HRI researchers in Fall 2015. Result ratios below 50 represent the prompt includes skipped response(s). Table VIII-X present the prompts, responses, and results from the second survey that are related to Quori’s design. Table VIII presents the survey data for robot appearance and actuation considerations in support of Section III-A1. Table IX presents the survey data for robot sensing and behavior considerations in support of Section III-A2. Table X presents demographic data of Survey #2.

### TABLE VIII
**Survey #2: Appearance and Actuation Considerations**

| Prompt | Responses | Result |
|--------|-----------|--------|
| (1) By default, Quori will have a mechanical appearance (clearly a robot, not a human); however, users may prefer a cartoonish appearance (similar to a 3D-animated character). What features would best express a cartoonish appearance? | Projected Face (e.g., 3D-animated vs. “robotic”) | 7.67 |
| | Vocal Behavior (e.g., how things are said) | 7.36 |
| | Visual Behavior (e.g., movement is different) | 7.33 |
| | Arms (e.g., exaggerated plastic arm covers) | 5.29 |
| | Chest (e.g., exaggerated plastic chest cover) | 5.00 |
| | Hands (e.g., exaggerated plastic hand covers) | 4.82 |
| | Mobile Base (e.g., exaggerated plastic base cover) | 3.60 |
| | Other (please explain) | 2.31 |
| | Do Not Support (please explain) | 2.00 |

| Prompt | Responses | Result |
|--------|-----------|--------|
| (2) By default, Quori’s upper body will be approximately 0.6-0.9 meters (2-3 feet) tall and the mobile base will be approximately 0.3 meters (1 foot) tall. Combined, Quori will be 0.9-1.2 meters (3-4 feet) tall. How would you prefer the HRI robot platform’s height to change? | Telescoping Spine (e.g., the robot can translate up and down a pole.) | 45% (17/38) |
| | Flexible Spine (e.g., the robot can lean forward and backward) | 26% (10/38) |
| | No Preference | 18% (7/38) |
| | Fixed Sizes (choose from: small, medium, and large) | 11% (4/38) |

| Prompt | Responses | Result |
|--------|-----------|--------|
| (3) How many hours does the platform need to operate on a single full charge? | 4 - 6 hours | 35% (13/37) |
| | 2 - 4 hours | 30% (11/37) |
| | 6+ hours | 27% (10/37) |
| | 0 - 2 hours | 8% (3/37) |

| Prompt | Responses | Result |
|--------|-----------|--------|
| (4) Quori will feature a low-cost holonomic base by default. What modular mobility options would you prefer? | Mobile | 1.74 |
| | Tabletop | 1.33 |

| Prompt | Responses | Result |
|--------|-----------|--------|
| (5) What data storage options would you prefer? | Removable | 2.59 |
| | Internal | 1.82 |
| | External | 1.71 |

| Prompt | Responses | Result |
|--------|-----------|--------|
| (6) What computing options would you prefer? | On-board | 72% (26/36) |
| | Off-board | 28% (10/36) |
### TABLE IX
**Survey #2: Robot Sensing and Behavioral Considerations**

| Prompt                                                                 | Responses          | Result          |
|------------------------------------------------------------------------|--------------------|-----------------|
| (1) By default, Quori will include several low-level software “drivers” (speech, eye gaze, gestures, proxemics, emotion, etc.) for behavior generation and recognition that are compatible with Robot Operating System (ROS). How much do you intend to modify Quori’s behavior generation and recognition systems? | Might modify: 56% (19/34) | Will modify: 26% (9/34) |
|                                                                        | Will not modify: 18% (6/34) |
| (2) Some (but not all) commercial (i.e., closed source, often for purchase, often with user/customer supported) software outperforms existing non-commercial (i.e., open source, often for free, sometimes with user support) software solutions for autonomous behavior generation and recognition. How do you rank commercial vs. non-commercial software? Would you use commercial software if it performs better and/or has user/customer support? | No: 55% (17/31) | Yes: 45% (14/31) |
|                                                                        | Spanish: 11.21    | Mandarin: 10.67 |
|                                                                        | Japanese: 8.92    | German: 8.46    |
|                                                                        | Hindi: 6.64      |
| (3) By default, Quori will speak English; however, some users may need to support other spoken languages. What spoken languages (other than English) should be supported by the platform? | Spanish: 11.21    | Mandarin: 10.67 |
|                                                                        | Japanese: 8.92    | German: 8.46    |
|                                                                        | Hindi: 6.64      |

### TABLE X
**Survey #2: Demographics**

| Prompt                                                                 | Responses          | Result          |
|------------------------------------------------------------------------|--------------------|-----------------|
| (1) How often do you personally (i.e., as an individual) ...           | ... modify, but do not publicly release, existing open-source software (i.e., software that originated from another individual or organization outside of your workplace)? | {55, 60, 35, 22}% |
| {Weekly, Monthly, Yearly, Never}                                       | ... modify and publicly release existing open-source software (i.e., software that originated from another individual or organization outside of your workplace)? | {10, 20, 25, 47}% |
|                                                                        | ... create and publicly release new open-source software (i.e., software that originated from your own ideas or workplace)? | {35, 20, 40, 31}% |
| (2) What is your age (in years)?                                       | 25-34: 13% (13/36) | 20-34: 10% (10/36) |
|                                                                        | 45-54: 6% (6/36)   | 40-44: 5% (5/36)  |
|                                                                        | 18-24: 5% (5/36)   | 35-44: 1% (1/36)  |
|                                                                        | 55-64: 1% (1/36)   | 50-64: 1% (1/36)  |
| (3) What is your gender?                                               | Male: 23% (23/36)  | Female: 11% (11/36) |
|                                                                        | Did not specify: 2% (2/36) |
| (4) What is your ethnicity? (Please select all that apply.)            | White/Caucasian: 77% (27/35) | Asian or Pacific Islander: 14% (5/35) |
|                                                                        | Did not specify: 9% (3/35) | Hispanic or Latino: 6% (2/35) |
| (5) How much do you know about the following topics?                   | Human-Robot Interaction: 5.31 | Artificial Intelligence: 5.00 |
|                                                                        | Robotics: 4.94     | Psychology: 4.00 |
|                                                                        | Signal Processing: 3.46 | Animation (2D/3D): 3.06 |
|                                                                        | Anthropology: 2.80 | Anatomy: 2.63 |