

# Robonaut 2 and Watson: Cognitive Dexterity for Future Exploration

**Julia M. Badger, Philip Strawser,  
Logan Farrell, S. Michael Goza**  
NASA Johnson Space Center  
2101 NASA Parkway  
Houston, TX 77058  
281-483-2277  
[julia.m.badger@nasa.gov](mailto:julia.m.badger@nasa.gov)

**Raphael Chancey**  
Apricity Health  
3262 Westheimer Rd. #762  
Houston, TX 77098  
512-773-9731  
[raphael.chancey@gmail.com](mailto:raphael.chancey@gmail.com)

**Charles Claunch**  
GeoControl Systems  
2900 Woodridge Dr. Suite 100  
Houston, TX 77087  
281-483-1580  
[charles.a.claunch@nasa.gov](mailto:charles.a.claunch@nasa.gov)

**Russell Potapinski**  
Woodside Energy Ltd.  
Woodside Plaza  
240 St Georges Terrace  
Perth WA 6000  
Australia  
+61 (0) 8 9348 5239  
[russell.potapinski@woodside.com.au](mailto:russell.potapinski@woodside.com.au)

**Abstract**—Future exploration missions will dictate a level of autonomy never before experienced in human spaceflight. Mission plans involving the uncrewed phases of complex human spacecraft in deep space will require a coordinated autonomous capability to be able to maintain the spacecraft when ground control is not available. One promising direction involves embedding intelligence into the system design both through the employment of state-of-the-art system engineering principles as well as through the creation of a cognitive network between a smart spacecraft or habitat and embodiments of cognitive agents. The work described here details efforts to integrate IBM’s Watson and other cognitive computing services into NASA Johnson Space Center (JSC)’s Robonaut 2 (R2) anthropomorphic robot. This paper also discusses future directions this work will take. A cognitive spacecraft management system that is able to seamlessly collect data from subsystems, determine corrective actions, and provide commands to enable those actions is the end goal. These commands could be to embedded spacecraft systems or to a set of robotic assets that are tied into the cognitive system. An exciting collaboration with Woodside provides a promising Earth-bound testing analog, as controlling and maintaining not normally manned off-shore platforms have similar constraints to the space missions described.

common concept of operations for these missions is to emplace habitats, spacecraft, and logistics in advance of the arrival of the crew. This important equipment will remain in place between crewed missions, but during this time, it is essential that the health of these assets is maintained. Ground support will clearly play a role in this, but with reduced communication bandwidth and increased latency, operations must advance beyond the paradigm of the International Space Station (ISS). As such, research into what technologies are needed to enable the autonomous operation of not always crewed human spacecraft is underway.

One promising direction involves embedding intelligence into the system design both through the employment of state-of-the-art system engineering principles as well as through the creation of a cognitive network between a smart spacecraft or habitat and embodiments of cognitive agents. The work described here details efforts to integrate IBM’s Watson and other cognitive computing services into NASA Johnson Space Center (JSC)’s Robonaut 2 (R2) anthropomorphic robot. R2 is a testbed for developing robotic technologies for astronaut assistance or human spacecraft caretaking. It was designed to be human-safe and dexterous enough to be able to manipulate the same tools and interfaces as humans do. Watson and other cognitive computing products are sets of technologies that ingest, analyze, and make connections between various types of data. An open source technology called Intu provides utilities for the embodiment of the intelligence, the connection of data sources, and the means of actuation with the cognitive services.

## TABLE OF CONTENTS

<b>1. INTRODUCTION</b> .....	<b>1</b>
<b>2. BACKGROUND INFORMATION</b> .....	<b>2</b>
<b>3. DESCRIPTION OF TASK AND PIPELINE</b> .....	<b>3</b>
<b>4. LESSONS LEARNED</b> .....	<b>5</b>
<b>5. CONCLUSIONS AND FUTURE WORK</b> .....	<b>5</b>
<b>ACKNOWLEDGEMENTS</b> .....	<b>6</b>
<b>REFERENCES</b> .....	<b>6</b>
<b>BIOGRAPHY</b> .....	<b>6</b>

## 1. INTRODUCTION

Future exploration missions that will send humans beyond near Earth orbit are in the planning stages at NASA. A

This paper will detail the efforts to combine R2 with cognitive computing using concepts like Watson's Intu<sup>1</sup> and discuss future directions this work will take. A cognitive spacecraft management system that is able to seamlessly collect data from subsystems, determine corrective actions, and provide commands to enable those actions, whether to embedded spacecraft systems or to a set of robotic assets that are tied into the cognitive system, is the end goal. An exciting collaboration with Woodside provides a promising Earth-bound testing analog, as controlling and maintaining not normally manned off-shore platforms have similar constraints to the space missions described. Robonaut units at JSC in Houston, at Woodside headquarters in Perth, Australia, and on the ISS are all important testbeds in the development of this network of cognitive services and embodiments.

This paper is organized as follows. Section 2 will give some background on previous work on robotic caretaking, cognitive computing, and using learning to solve cyber-physical problems. Section 3 will describe the integration of cognitive technologies with Robonaut 2 to demonstrate human-robot interaction in gathering tools. Section 4 will describe lessons learned from this exercise, the potential for applying this technology in industry, and introduce an autonomy framework under development at NASA JSC. Section 5 will conclude the paper with a focus on the vision of future work along this promising path.

## 2. BACKGROUND INFORMATION

The Robonaut 2 project started as a collaboration with General Motors. The intent was to develop a robot that could interact safely with humans, tools, and interfaces while still maintaining its ability to handle significant payloads (approximately 9 kg). For NASA, Robonaut 2 (shown in Figure 1) was envisioned to be a robotic astronaut assistant or a human spacecraft caretaker. An R2 unit was launched to the International Space Station in February 2011, and has spent many years as a testbed for robotic technology development.



**Figure 1: Robonaut 2**

Several advances were made learning from the task board and tool use activities conducted on the International Space Station [1,2]. In 2014, robotic manipulators, or legs, were flown to the ISS to allow R2 to develop climbing technologies in the next step of discovery along the path of robotic caretaking. This pushed an advancement of the control and safety system on Robonaut [3,4], as well as multiple advancements in the methods of control of the robot. While the first set of robotic experiments focused on controlling Robonaut through tele-operation and direct control, new methods of commanding dexterous manipulation and climbing using more supervisory control and autonomy have been the recent research focus [5].

Recent Robonaut demonstrations showcasing robotic caretaking technology development have centered around spacecraft maintenance and logistics tasks [6]. Using the concept of Affordance Templates [7], operators can create plans to interact with objects and tools in the robot's environment on an activity level. The combination of the Affordance Templates framework with a powerful yet generic task-based interface called TaskForce makes a powerful scripting and autonomous control tool chain that enables many types of robot-environment interactions.

Cognitive computing is generally defined as a set of machine learning and artificial intelligence techniques that are roughly modeled after human cognition. They are meant to scale and interact naturally with humans. These techniques are generally good at quickly analyzing large unstructured data sets and drawing conclusions about them with respect to some query. IBM's Watson technology is one cognitive computing platform, and examples of some of its services include Speech-to-Text, Conversation, Personality Assessment, and Tone Analyzer, among others. While these services have been vetted using a computer interface, the application of these technologies to interactions with the world is largely unexplored. This paper will describe the

<sup>1</sup> <https://www.ibm.com/watson/developercloud/project-intu.html>

successful integration of Robonaut 2 with cognitive services, and will discuss promising future directions of research.

### 3. DESCRIPTION OF TASK AND PIPELINE

Cognitive services provide nearly endless possibilities for robotic integration. A plan was developed to integrate services, starting with those that had the least barrier to integration, based on the amount of development that would be needed. The initial integration of cognitive abilities with the Robonaut platform started with natural language processing, followed by object classification and recognition. A series of demonstrations were selected as integration goals. The first demonstrations focused on teaching the robot simple free-space motions, such as waving to the crowd, based on natural language commands. Next, a demonstration where Robonaut localized, grasped, and used a drill using natural language processing but no vision-related cognitive services. The final demonstration also centered around Robonaut retrieving tools at the request of a human team member, but this time while using cognitive services in the vision pipeline.

The demonstration features multiple command modalities for the human-robot system. First, a human counterpart will request that Robonaut hands him or her a tool. Several tools will be available to Robonaut, near or in its workspace. The tool cabinet is shown in Figure 2. Robonaut will use its vision sensors and processing pipeline to identify and localize the tool requested. If the tool is not found or is outside of Robonaut’s workspace, it will communicate those outcomes to the human counterpart. If the tool is available and accessible, Robonaut will plan a path to the tool, grasp the tool, and then attempt to hand the tool to the human counterpart. When the human is ready, he or she will simply have to grab onto the tool to trigger Robonaut’s release response to complete the transfer.



**Figure 2: Tool cabinet for the Interactive Tool Gather Demonstration**

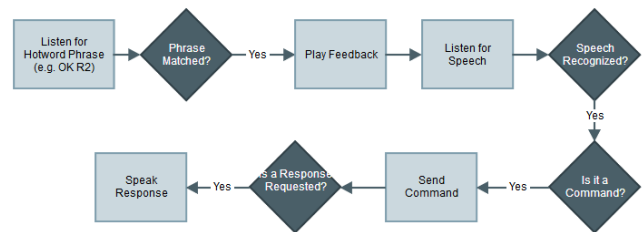
This section will cover the tool pipelines for the natural language processing, the vision processing, and the manipulation tasks that were essential to this integrated demonstration.

#### Natural Language Interface

Robonaut’s voice interaction is a mixture of hotword listening, phrase recognition, and intent reasoning. The hotword process uses the Python PyAudio library in combination with Google Cloud Speech in a continuous loop, looking for the words “Okay Robonaut.” When the hotwords are detected, a message is published using Robot Operating System (ROS<sup>2</sup>) to another process, which then listens again using the same libraries, except for a given period of time (typically about three to five seconds). This chunk of audio is sent to and converted to text by the same Google Cloud Speech API. Any detected speech is then published in a ROS message, which is consumed by a third process, the Voice Commander. The Voice Commander process listens for recognized speech to match to a user specified command list in order to perform TaskForce commands. In cases where an exact command match is not found to fit the phrase, the cognitive services can reason about the intent of the user, which can additionally be used to determine an appropriate action. These actions can be validated by the user. The Voice Commander could then learn from the user validations.

Text to Speech is a fourth process, which provides a ROS Service (a request-reply remote procedure call that uses ROS messages) that can be triggered with a text string to be spoken. The process uses IBM Watson’s Text to Speech API to generate an audio file, which is then played using the Advanced Linux Sound Architecture (ALSA). The Voice Commander process also has a direct instantiation of the Text to Speech object and provides an immediate response capability when a particular command is matched.

A flow chart of the natural language interface is shown in Figure 3.



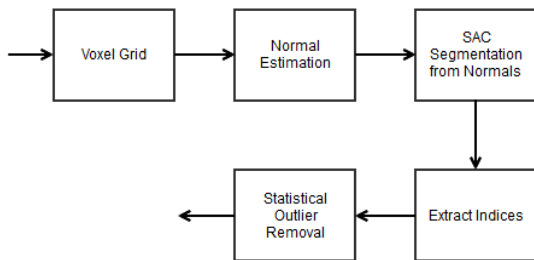
**Figure 3: Flow Chart of Speech Recognition**

<sup>2</sup> <http://www.ros.org/>

There do exist some speech to text libraries, such as two libraries developed by Carnegie Mellon University called Sphinxbase<sup>3</sup> and PocketSphinx<sup>4</sup>, that do not require the cloud to run. These have been explored since deployed robots on space missions will not have access to cloud services. However, it was found that the quality of sound recorded by Robonaut’s microphone was not sufficient for robust word recognition. Since the microphone is located inside the head along with fans, sensors, and motors, the electrical noise would need to be overcome by careful filtering in order to use these promising libraries. Time constraints precluded investigating ways to eliminate or reduce this problem.

### Vision Processing

The ASUS Xtion Pro-Live sensor broadcasts the raw point cloud and RGB image data. A pipeline of Point Cloud Library (PCL<sup>5</sup>)-ROS nodes, shown in Figure 4, are used down-sample the point cloud, to remove the plane of the tool rack, and to remove outlier points.



**Figure 4: Point Cloud Processing Pipeline**

A clustering node takes in the raw point cloud from the ASUS, the corresponding RGB image associated with the point cloud, and the output of the pre-processing pipeline. First, the unstructured point cloud is broken up into clusters using the `pcl::EuclideanClusterExtraction` node. Then, for each cluster, a mask is applied to the RGB image such that only one cluster is still visible. This involves using a kinematic tree library to match the filtered points with indices of the full point cloud. Those correspondences are then mapped to pixel positions. Then, morphological operators are used to improve the mask and extract the individual tools. That processed image is sent via ROS service call to a classifier node to label the object in the masked image. The cluster corresponding to the selected label is then sent to the Iterative Closest Point (ICP) localizer node. This node is written in C++ and uses PCL, OpenCV<sup>6</sup>, and ROS.

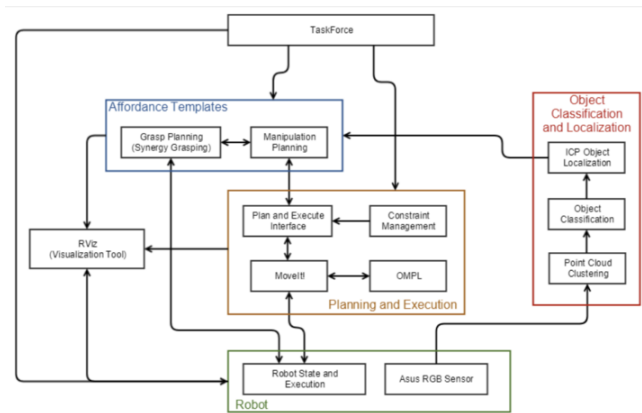
The classifier node is a 2D image classifier written in Python. It is built on top of the TensorFlow<sup>7</sup>

library. TensorFlow is an open source software library for numerical computation using data flow graphs, and is meant for neural network and deep learning research. The classifier node provides a ROS service call interface that takes in an RGB image, and returns a classification label (a string) and a confidence value. The training images were generated by placing each tool in discrete locations all along the tool cabinet and collecting images of the front and back of the tool. The classifier was then trained for optimal performance.

Once a cluster is labeled, it goes to the ICP node. The ICP node is a ROS node written in C++ and uses PCL (`pcl::IterativeClosestPoint`). It attempts to register a model (which is selected via an input topic) with the input point cloud. Once matched, the transform is broadcast out on TF, the transform ROS topic. Models from the internet were found and converted to point-cloud files to support this step.

### Manipulation Pipeline

The manipulation pipeline combines the vision processing pipeline described in the previous subsection with the Affordance Templates framework and some planning and execution nodes to create the overall behavior. The tools are collected in the TaskForce task execution engine and concurrently run with the natural language processing pipeline previously described. The overall manipulation pipeline and its connections to the vision processing pipeline are shown in Figure 5.



**Figure 5: Manipulation Pipeline**

In the Planning and Execution nodes, OMPL<sup>8</sup> is the Open Motion Planning Library and MoveIt!<sup>9</sup> is a planning interface associated with ROS. The other two nodes have been developed in house to work with the highly constrained, high degree of freedom motions that Robonaut typically makes [8].

<sup>3</sup> <https://github.com/cmusphinx/sphinxbase>

<sup>4</sup> <https://github.com/cmusphinx/pocketsphinx>

<sup>5</sup> <http://pointclouds.org/>

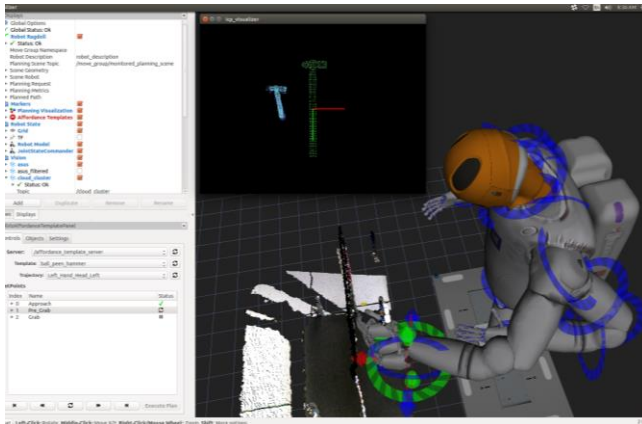
<sup>6</sup> <https://opencv.org/>

<sup>7</sup> <https://www.tensorflow.org/>

<sup>8</sup> <http://ompl.kavrakilab.org/>

<sup>9</sup> <http://moveit.ros.org/>

The Affordance Templates framework, shown in Figure 6, associates a manipulation plan (paths between a series of set points on the approach to grasp an object) and a grasp plan with an object. The manipulation plan makes use of OMPL and MoveIt!. The grasp plan uses a concept that allows several types of objects to be associated with a grasp type by focusing on the essential parameters of the object's shape with respect to the particular end effector to be used. The approach is called synergy grasping [9], and allows the complex Robonaut hand to be simply controlled for many grasp types.



**Figure 6: Affordance Templates Framework**

#### 4. LESSONS LEARNED

The process of creating the demonstration discussed in the previous section provided experience in combining robotic manipulation with cognitive computing. While the cognitive computing was limited to the robot's direct interactions with the environment (vision processing and natural language processing only), the potential for the expansion of this technology to more parts of the robotic task was realized. In particular, a next step would be to include cognitive computing into the Affordance Templates framework by allowing general models to be associated with various types of tools (i.e., all types of screwdriver would share a model). This could be enabled by allowing learning to affect object models. If that is possible, the learning could influence how the object is visualized, approached, and grasped, creating a multiplication of efforts and reducing the time needed to teach the robot a new task.

The concept of using cognitive computing to find connections in unstructured data is becoming popular for autonomous procedure management, whether that procedure is for controlling maintaining a spacecraft or for analysis of processes in a Liquefied Natural Gas (LNG) plant. The combination of robotics with smart spacecraft, LNG plants, or off-shore platforms is compelling, and has brought together a collaboration between NASA and Woodside that has been very fruitful. The robotic caretaking of not-normally-manned (NNM) offshore platforms is a topic of research for Woodside, as it promises to increase situational awareness and safety when maintaining these locations.

Similarly, the not-always-crewed concepts for future exploration spacecraft will benefit from having intelligent agents (both subsystems and robots) to respond to contingencies and provide maintenance and logistics support. The tie that binds the two organizations together is the shared interest in using the intersection between robotics and cognitive computing to accomplish these goals.

#### 5. CONCLUSIONS AND FUTURE WORK

The embodiment of cognitive computing technologies into robotic agents to allow these cognitive solutions to interact with the world has been successfully demonstrated on the Robonaut 2 testbed, both at the Johnson Space Center and at Woodside headquarters in Perth, Australia. Future work focuses on the continued integration of these technologies specifically to enable maintenance and logistics functions on board future exploration spacecraft or NNM offshore platforms. In both cases, the facility will be heavily instrumented and will provide rich data to the cognitive system. The cognitive system will also need to control the movements and actions of the robotic agents providing the maintenance or logistics functions. As such, several directions of future research are possible.

First, cognitive computing has the potential to be extremely valuable for system health management. The collection and analysis of data to determine the most likely health state of the overall system could be generated by an intelligent agent, and then that state could be connected to maintenance or inspection procedures. For example, if an acoustic signature is detected by a mobile robot on a fan in the spacecraft (or in an LNG plant), that fan could be turned off and scheduled for maintenance, and a redundant fan could be used in the interim.

Secondly, extensions such as storing the relative location of classified objects into a 'world' knowledge base can enable additional use cases where the human counterpart can then ask Robonaut the location of a specific tool. The path planning can then be verbalized as well. Continual updating as objects are reclassified should improve accuracy and utility.

The ability of the intelligent system to generate procedures (say, from a collection of relevant procedures) and to generate a task plan for the multi-agent team is feasible using cognitive solutions. Task planning for robotic agents is often much more difficult than planning for smart facilities, as an ontology of mappings between words and common sense actions are required to make robots go. For example, telling a robot to vacuum a filter requires informing the robot to do a coverage motion plan over a specific area while maintaining a set distance from or force on the filter. However, most humans would understand how to interpret the initial instruction without any extra information. Work has been done to show that cognitive services connected by Watson's Intu can allow robotic agents to build up new skills from a

learned combination of task primitives. This work has set the stage for future work in cognitive task planning for robotics.

## ACKNOWLEDGEMENTS

The authors thank the IBM Watson Labs, Alison Barnes, Anthony Biviano, Mary Hewitt, and Hide Muta for their contributions to this work.

## REFERENCES

- [1] M.A. Diftler et al. "Robonaut 2 – Initial Activities On-Board the ISS," in the Proc. of the IEEE Aerospace Conference, 2012.
- [2] T.D. Ahlstrom, M.A. Diftler, R.B. Berka, J.M. Badger, S. Yayathi, A.W. Curtis, and C.A. Joyce. Robonaut 2 on the international space station: Status update and preparations for IVA mobility. In AIAA SPACE 2013 Conference and Exposition, San Diego, CA, pages 1–14, 2013.
- [3] J. M. Badger, A. M. Hulse, R. C. Taylor, A. W. Curtis, D. R. Gooding, and A. Thackston. Model-based robotic dynamic motion control for the Robonaut 2 humanoid robot. In Proc. of IEEE-RAS International Conference on Humanoid Robots. IEEE, 2013.
- [4] J. M. Badger, A. M. Hulse, and A. Thackston. Advancing safe human-robot interactions with Robonaut 2. In Proc. of the 12th International Symposium on Artificial Intelligence, Robotics and Automation in Space, 2014.
- [5] J. M. Badger, S. W. Hart, and J. Yamokoski. Towards autonomous operation of Robonaut 2. In Infotech@Aerospace 2012. 2011.
- [6] L.C. Farrell, P. Strawser, K. Hambuchen, W. Baker, and J. Badger. Supervisory control of a humanoid robot in microgravity for manipulation tasks. In Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2017.
- [7] S. Hart, P. Dinh, and K. Hambuchen. The affordance template ROS package for robot task programming. In Proc. of the 2015 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2015.
- [8] W. Baker, Z. Kingston, M. Moll, J. Badger, and L. Kavraki. Robonaut 2 and you: Specifying and executing complex operations. In Proc. of the IEEE Workshop on Advanced Robotics and its Social Impacts. IEEE, 2017.
- [9] L.C. Farrell, T.A. Dennis, J. Badger, and M.K. O'Malley. Simply grasping simple shapes: Shape based synergy commanding of a humanoid hand. In Proc. of the International Symposium on Robotics Research. Microsoft, 2017.

## BIOGRAPHY



**Dr. Julia Badger** is the Project Manager for the Robonaut Project at NASA-Johnson Space Center in Houston, TX. She is responsible for the research and development of humanoid robotic capabilities, both on the Earth and on the International Space Station, that include dexterous manipulation, robotic autonomy, and human-robot interfaces. She is also the Autonomous Spacecraft Manager Project Manager, which is developing provably correct autonomy frameworks for spacecraft. Julia has a BS from Purdue University, and an MS and a PhD from the California Institute of Technology, all in Mechanical Engineering.



**Philip Strawser** received a B.S. in Computer Engineering at Georgia Tech in 2002. In 2002, he joined NASA's Johnson Space Center (JSC) in Houston, Texas. At NASA, Mr. Strawser works in the Robotic Systems Technology branch of the Software, Robotics, and Simulation division. Initially focused on avionics and embedded systems, he helped to design and implement several robotic platforms including Robonaut, Spidernaut, and Centaur. He later focused on software development, and in 2007, he led the software team to develop Robonaut 2. He led the software certification effort which allowed R2 to go to the International Space Station in 2011. In 2012, Mr. Strawser worked with a JSC team to design and develop the Valkyrie robot for the DARPA Robotics Challenge. Mr. Strawser is currently the Perception and Cognition lead for the Robonaut 2 system. His interests are in computer vision, machine learning, and robotic task design and execution.



**Logan Farrell** is the Mobility and Manipulation lead on the Robonaut 2 project and works at NASA, Johnson Space Center in the Robotics Division in development of rovers and humanoid robotics as both an actuator designer and an application programmer. Logan received his B.S. in Mechanical Engineering from the University of Illinois and is currently pursuing his M.S. in Mechanical Engineering at Rice University.



**S. Michael Goza** received his B.S. in Aerospace and Ocean Engineering from Virginia Polytechnic Institute and State University in 1984. From 1980 to the present he has worked at NASA Johnson Space Center on a number of projects ranging from aerodynamic analysis, to computer graphics, to robotics.

Currently, Mr. Goza is part of the Robonaut team in charge of telepresence control software, human – machine interfaces, mobile platform integration, spider gait generation, and computer simulation. He is currently branching out into the field of autonomous flight vehicles. Mr. Goza has developed control software for a number of robots such as Robonaut, Robonaut 2, Centaur, Centaur 2, Spidernaut, and the Space Exploration Vehicle.

including engineering, mergers and acquisitions, corporate strategy, culture change and business leadership.



**Charles "Chuck" Claunch** is a software engineer with expertise in systems integration and software design. He has a B.S. in Computer Science from Texas Tech University and worked in the telecom industry for a few years before his nearly ten years at NASA. At NASA, Chuck has worked on several projects which

have flown to the International Space Station.



**Raphael P Chancey (Ray)** is currently the CTO for Apricity Health. He has always been inspired by science fiction and has always embraced the challenge of taking ordinary use cases and transforming them by imagining the world of tomorrow. He endeavors to leave my mark on society by working on projects

which transform everyday lives. He spent the last 19 years in IT and have deep technical expertise building client solutions. He most recently was the Director of the Watson Labs in IBM pioneering transformational use cases merging traditional technologies and AI.



**Russell Potapinski** is the Head of Cognitive Computing and Robotics for Woodside Energy, responsible for leading the vision and implementation of these new technologies throughout the business. Russ has more than 20 years' industry experience. Russ joined Woodside in 1998 and has held a variety of roles in areas

