Abstract

SBIR A18-030 : "Develop a service-oriented-architecture that permits the use of measures from unobtrusive COTS sensors, training data, and other measures of health and wellbeing to understand, manage, and optimize the well being and performance of Army enlistees in an initial entry training environment."

In addition the above objective, through our questions with the Program Manager, we have added the element of identifying how not only individual trainees perform but also how they perform together in concert by varying the composition of trainee "teams" during instruction. Data collection around these points will be critical for future (outside of the scope of this SBIR) extension the unit assignment algorithm to consider unit cohesiveness.

The way we look at this problem is fourfold:

- Identifying a suite of wearable devices that capture and disseminate various biometrics in real-time. These wearables must be comfortable and not interfere with training activity while simulatniously be able to withstand the harsh realities of a Basic Military Training environment.
- A Service-Oriented-Architecture Azure "Edge" Service Bus capable of capturing the massive amount of real-time data and feeding it to Mil Azure in near-real-time for further processing.
- Trainable Artificially Intelligent Models in Mil Azure that will not only slice the data into relevant structures but also discover predictive information on how the captured biometrics relate to performance in other areas besides the current task. This is includes platoon and squad cohesiveness metrics.
- A near-real time reporting dashboard at all levels of Training Command which reports and alerts to instructors/commanders projected deficiencies in training to make alterations to individual and squad plans in near-real-time.

Proposed Phase I Solution

We have identified three wearables to capture the data required to perform a proper feasibility study on the machine learning end, being the primary deliverable of Phase I. From our point of view it is impossible to complete Phase I without developing a near-production ready system for Phase II due to the fuzzy nature of machine learning: We need to at least a rudimentary understanding of how these biometrics correlate to specified performance metrics to indeed see if there is any correlation at all to justify a Phase II.

Phase I Wearable Strategy

After evaluation of the marketplace for near-industrial grade sensors we have settled on empatica, a Massachusetts Institute of Technology spin-off, to provide two wrist bands that are to be alternated based on training environment. The more advanced E4 is informationally not water resistant therefore would be inappropriate for a small subset of training exercises. Their less capable embrace product is rated for IP67 meaning anything up to submersion is covered.

Neither of these devices support Wifi as they are intended to be paired with a phone, instead due to the training environment we have gone with Misfit Vapor, a Fossil product, to relay the signals to the secure cloud, provide a user interface and GPS capabilities.

empatica E4
American/Italian Design
South Korean Made
No IP Rating

empatica embrace
American/Italian Design
South Korean Made
IP67

Mistfit Vapor
American/Swiss Design
?? Made
IP68

	E4	embrace	Vapor
EDA (Measures nervous system activity)	○	○	⊘
Gyroscope	??	○	○
3-Axis Accelerometer	○	○	○
Peripheral Temperature Sensor	○	○	⊘
PPG Sensor (Measures Blood Volume Pulse/Heart rate)	○	⊘	⊘
Optical Heart Rate Sensor	⊘	⊘	○
GPS	⊘	⊘	○
Bluetooth	○	○	○
WiFi	⊘	⊘	○
Water Resistant	⊘	○	○
Water Proof	⊘	⊘	○

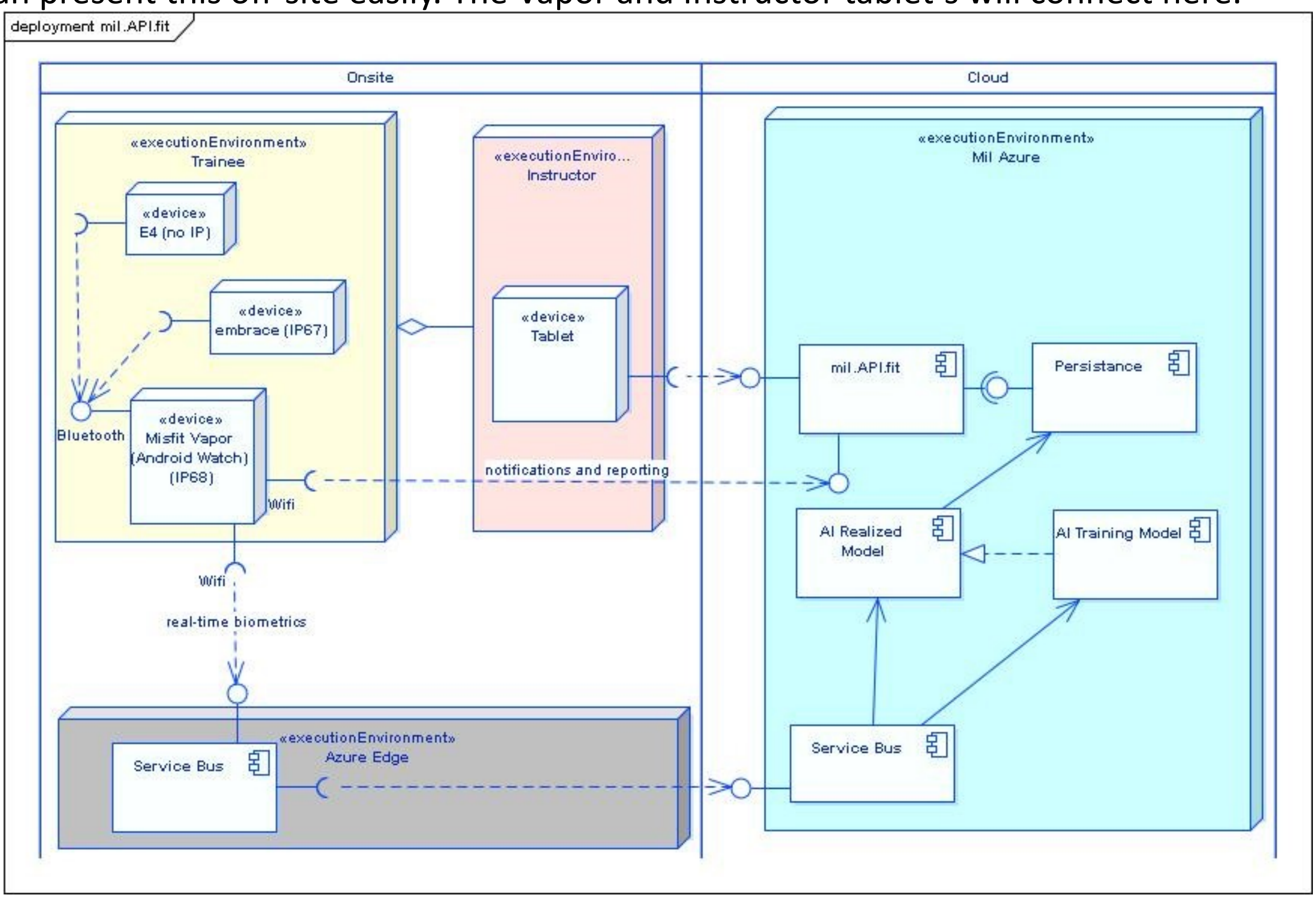
Simultaneous Phase II Prep

We believe that a two-primary sensor array with no true water proofing is not sufficient for a Basic Military Training deployment. As such we have initiated discussions with empatica, backed by Leidos investment, to develop a IP68 version of the E4 with prototypes due in Q1 2019. While we cannot rely on this occurring given the timeframes of SBIR bid development; we believe it to be a very real possibility.

Phase I Systems Strategy

The overarching goal of the mil.API.fit system is to extremely reliably capture biometric data in real-time and run that data through a near-real-time artificially intelligent model to capture previously identified metric as well as feed the data into an AI model to identify new correlations, realizing those back into the model for future data sets. The amount of data throughput we will be dealing with on this project, even with 25-50 test subjects, is signficant and thus requires robust hybrid architecture.

Our plan is deploy an Azure "Edge" cloud on-premise that is a mirror in functional capability of Mil Azure. This will have a Service Bus that collects the data in real-time; from the Vapor which is streaming its own data and the E4/embrace's as well (if connected); and then performs a lightweight transform of the data before it ships it off to Mil Azure. Mil Azure is required as it has significantly more compute power and is globally available so, after run through the modeling and persisted, mil.API.fit can present this off-site easily. The Vapor and Instructor tablet's will connect here.



empatica Sensor Capabilities

There are two sensor devices that recruits would wear mounted on a wristband. The E4 and the Embrace which are both produced by Empatica Inc., Cambridge, MA, USA and Milano, Italy. The Embrace is a Class IIa Medical Device, according to Directive 93/42/EEC, for patients with epilepsy or at risk of having epilepsy. The Embrace has been approved this month (02/2018) by the FDA as a medical device for early detection of epileptic seizures; however, at the time of this writing an official notification of approval is not yet available since they are published at the end of each month. There are four sensors on the E4: accelerometer, infrared thermopile, PPG and EDA. There are four sensors on the Embrace: accelerometer, infrared thermopile, gyroscope and EDA. The accelerometer captures motion in all three axes.

The infrared thermophile reads peripheral skin temperature. The gyroscope measures rotational motion. Photoplethysmographic (PPG) signals measured via blood volume pulse oximetry are typically used for measuring heart rates. Such wearable sensors may be used for early detection of abnormal conditions for preventive actions in monitoring individual health. However, it is challenging to estimate heart rate with high accuracy using PPG signals unless the subject is at rest due to motion artifacts. Electrodermal activity (EDA) is measured via galvanic skin response. Changes in vital signs such as heart rate occur as a result of both the sympathetic and parasympathetic nervous system activation.

The skin is the only organ that is purely innervated by the sympathetic nervous system and not affected by parasympathetic activation. Sympathetic activation increases when you experience surprise or excitement; it can also increase with stress whether physical, emotional or cognitive. In some medical conditions (e.g., epilepsy), sympathetic activation shows significant increases that are related to specific brain structures activation (Dlouhy et al. 2015). Conversely, parasympathetic activation has a calming influence on vital signs such as slowing the heart rate.

These sensors will give us the capability to monitor recruits in various ways. In real-time, we will have access to their vital signs safety profiles for temperature and heart rate. For example, those recruits who are acutely overheating could be quickly identified and given first aid. Alternatively, once profiles associated with good boot camp outcomes are identified via machine learning, then near real-time feedback could be supplied to drill instructors.

Dlouhy, B. J., Gehlbach, B. K., Kreple, C. J., Kawasaki, H., Oya, H., Buzza, C., Granner, M. A., Welsh, M. J., Howard, M. A., Wemmie, J. A., Richerson, G. B. (2015). Breathing Inhibited When Seizures Spread to the Amygdala and upon Amygdala Stimulation. Journal of Neuroscience, 35(28), 10281-10289.

Machine Learning Strategy

We propose to correlate the biometric signals captured by wearable sensors to tangible outcomes like the number of sit-ups performed in two minutes at the end of boot camp. We will expand the data available beyond the biometric sensors alone to include relevant baseline data available such as age, gender, height, weight, body fat proportion, ASVAB score, etc. Artificial intelligence machine learning is a broad computational statistical concept which allows computer algorithms to correlate inputs like heart rate and body temperature to the desired outputs like sit-ups or body fat proportion.

Machine learning methods derive complex black-box model relationships of inputs to outputs. We call them black-box models because the models are so complex that you can't readily evaluate the models based on the details of the models themselves; rather, we will base the model's performance on the accuracy of their predictions. A common critique of machine learning is that a black-box model might be very predictive, but those responsible for making decisions based on these models might reasonably require an explanation of how and why the model works. Several techniques exist which allow us to peer inside the black-box such as the partial dependence function (Friedman 2001); Decoupling, Shrinkage and Selection (DSS) (Hahn and Carvalho 2015); and sparse prior variable selection (Linero 2016).

Unfortunately, the data necessary to train these models does not currently exist since it has not been captured to date. Therefore, during phase 2 of this contract, we would acquire the necessary data from a large number of basic training recruits. And we would train our machine learning models on the baseline and recorded sensor data in order to predict outcomes of interest. However, how predictive are these models? Generally, machine learning falls in the class of ensemble models which are the best known predictive models at this time (Krogh and Solich 1997, Baldi and Brunak 2001, Kuhn and Johnson 2013). We will assess predictive ability by training on say, 400 recruits, and then use those models to predict the outcomes of another 100 recruits to assess the accuracy.

Note that missing data is an important issue in data collection projects like this. We adopt a two-pronged strategy. First, for various reasons such as illness or sensor malfunction, a given recruit's sensor data might not be captured for a particular training session. Machine learning is flexible enough to allow us to include each recruit for as much data as has been observed which will often be adequate since there will be many training opportunities. Second, for missing baseline data, we will impute the missing variables with a machine learning missing data technique called Sequential BART (Xu et al. 2016).

We will explore various machine learning methods to see which performs the best for this problem including neural networks (Baldi and Brunak 2001), gradient boosting (Friedman 2001), random forests (Brieman 2001), support vector machines (Dimitriadou et al. 2008) and Bayesian Additive Regression Trees (Chipman et al. 2010). Well tested and validated software for all of these techniques is available as part of the Microsoft R Server product available in the Mil Azure cloud.

Baldi P, Brunak S. Bioinformatics: The Machine Learning Approach. MIT Press: Cambridge, MA, 2001.
BREIMAN, L. (2001). Random forests. Machine Learning 45 5–32.
Chipman HA, George EI, McCulloch RE. BART: Bayesian additive regression trees. Annals of Applied Statistics 2010; 4:266–98.
DIMITRIADOU, E., HORNIK, K., LEISCH, F., MEYER, D. and WEINGESSEL, A. (2008). e1071: Misc functions of the Department of Statistics (e1071), TU Wien. R package version 1.5-18.
FRIEDMAN, J. H. (2001). Greedy function approximation: A gradient boosting machine. Ann. Statist. 29 1189–1232.
Hahn, PR and Carvalho, CM. (2015). Decoupling shrinkage and selection in Bayesian linear models: a posterior summary perspective. JASA 110, 435–48.
Krogh A, Solich P. Statistical mechanics of ensemble learning. Physical Review E 1997; 55:811–25.
Kuhn M, Johnson K. Applied Predictive Modeling. Springer: New York, NY, 2013.
Linero, AR. (2016). Bayesian regression trees for high dimensional prediction and variable selection. JASA. <doi:10.1080/01621459.2016.1264957>.
Xu, D, Daniels, MJ and Winterstein, AG. (2016). Sequential BART for imputation of missing covariates. Biostatistics 17, 589–602.

Some Possible Available Physical Metrics

Metric	Likely Sensors	Method	Certainty	Comment
Hypo/Hyperthermia	Infrared thermopile		↑	
Seizure detection	EDA and accelerometer		↑	approved by the FDA 01/26/2018
Cardiac irregularity	PPG		↗	No data to back up direct measurement
Situps in 2 minutes	PPG and accelerometer		↗	
Pushups in 2 minutes	PPG and accelerometer		↗	
Two-mile run time	PPG and accelerometer		↗	
Marksmanship	PPG, EDA and accelerometer		→	
Hydration	PPG, EDA and infrared		→	

Direct Measurement **Machine Learning**

Psychometric Instruments

The machine learning approaches will be augmented through use of traditional psychometric self-report instruments. We will focus primarily on a broad conceptualization of wellbeing. This approach will incorporate three areas, including individual psychological strengths important to military leadership (Matthews et al, 2006), individual psychological weaknesses, and team cohesion. Validated measures including the Values in Action (VIA) inventory (Peterson & Seligman, 2006), the Brief Resilience Scale or Connor-Davidson Resilience Inventory, the Duckworth Grit Scale and the Wellness Inventory will be used to assess individual strengths, coping approaches and social connectedness.

The Beck Depression Inventory II, and the PTSD Check List (PCL-5) will be used to assess negative aspects of psychological wellbeing. Team cohesion will be assessed using the Group Environment Questionnaire (Callow, 2009). Psychometric assessments will be delivered using a tablet, with data capture directly into MS Azure framework immediately following the consent process. These data will be linked to sensor data using a participant ID number.

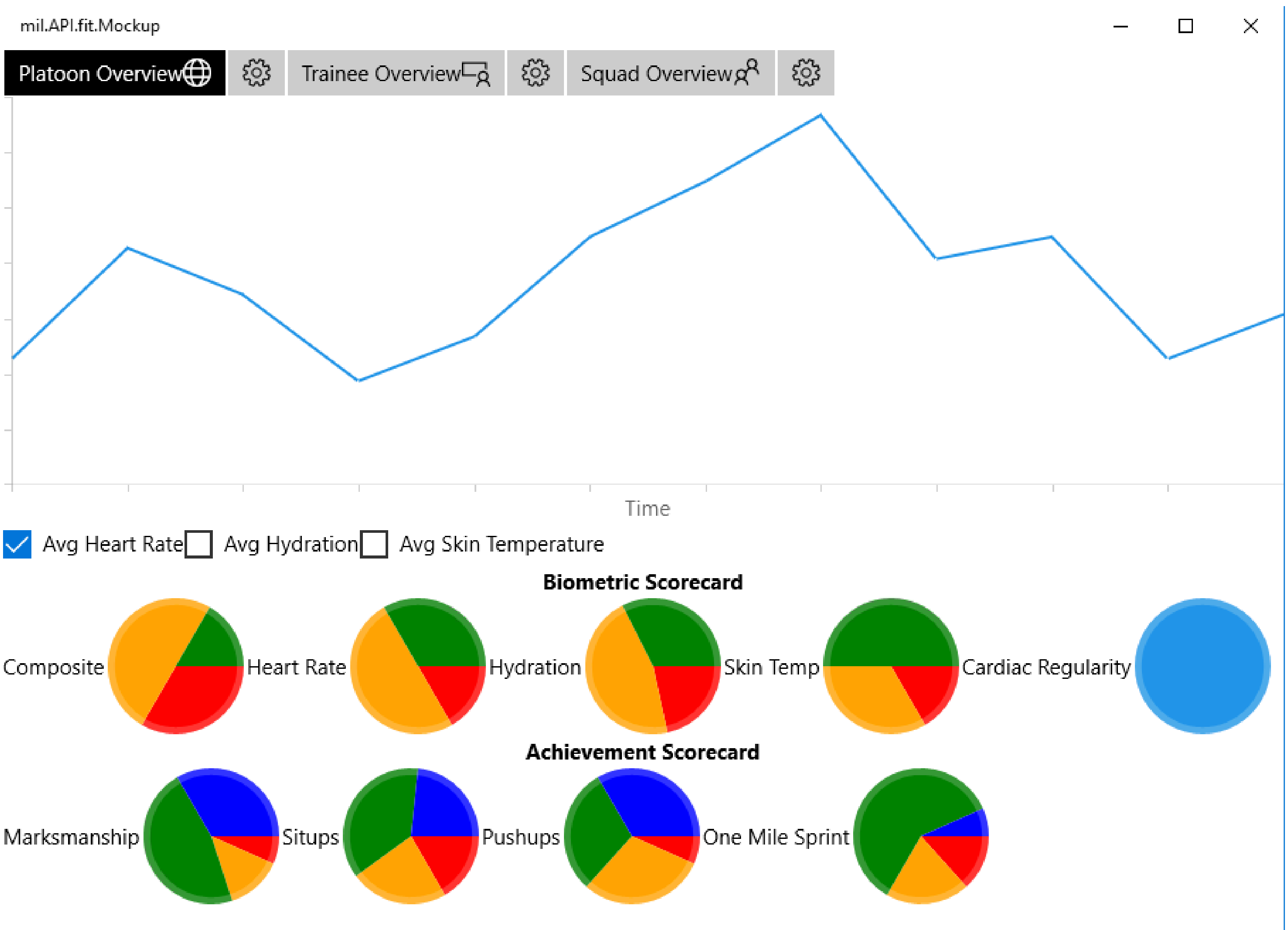
Some Possible Psychometric Metrics

Psychometric Instrument	Sensor Detected Performance Metric	Comment
Sleep quality	Hypo/Hyperthermia	some evidence in cadet training with extended sleep deprivation of daytime impact on thermoregulation
	Seizure detection	
Sleep quality, stress measure	Cardiac Irregularity	
	Marksmanship	Sleep deprivation may cause hand tremor, nystagmus
Grit, BDI depression, team cohesion	Situps in 2 Min	
	Pushups in 2 min	
	2 mile run	

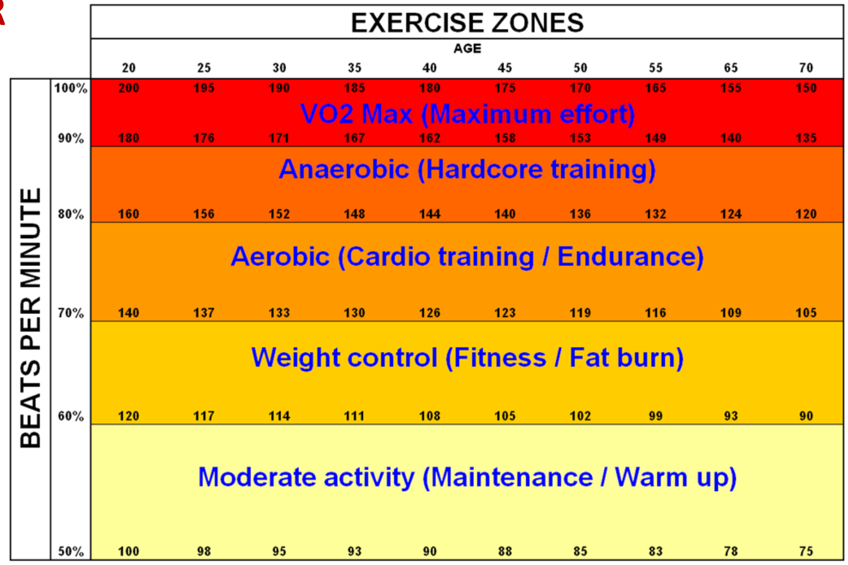
Real-Time Training Modification Strategy

From our perspective the prime objective of US Army Basic Combat Training (BCT) is to transform civilians into combat effective military personnel. While the physical, tactical and survival aspects of BCT are equally important, in our opinion, so are the psychological and safety aspects. We propose that with this new sensor data that Instructors be empowered to apply, in both positive and negative multipliers of magnitude, the level of all aspects of drill to both individual trainees and concerted squadrons within a platoon.

In order to achieve this we must build a configurable dashboard for the Training Instructor with usable metrics to understand where their platoon, and then drilling down into the individual squadrons and indeed the trainees themselves, are at metric wise delivering real insight into the level of effort being put forth while simulatniously providing warning signs such as irregular heart beat, dangerous hydration level and even psychological factors leading to a mental breakdown and washout. The mocked up dashboard below is designed to fit on a tablet for the Training Instructor to regularly, if not constantly, refer to.



Interpretation of Heart Rate for Real-Time Training Modification



The Training Instructor can also use the dashboard to push trainees to their maximum potential. To improve fitness, it is important to reach your bodies Target Heart Rate, or training heart rate. The THR is a desired range of heart rate reached during aerobic exercise, which enables your heart and lungs to receive the most benefit from a workout. In turn, this will improve your endurance. Experts recommend that you monitor your heart rate during your exercise routine. Backed by the science of Excess Post-Exercise Oxygen Consumption (or EPOC), heart rate monitored training is designed to maintain a target zone that stimulates metabolism and increases energy. When a trainee exercises, are they working hard or hardly working? Exercising at the correct intensity can help a soldier get the most out of their physical activity. The Training Instructor will know which trainees to push harder and those who have already maximized their heartrates.

The U.S Army Research Lab report ARL-RP-0595 published in April 2017 “The Effect of Soldier Marching, Rucksack Load, and Heart Rate on Marksmanship” concluded increases in shooting HR after marching result in lower probability of hitting the target. Through an instructor’s dashboard, trainees can use different techniques to lower the heartrate while shooting. This will be an individualized approach to the technique to use. The result will be more combat effective soldiers.

Trainee Heart Rate	Activity	Instructor Response
Moderate Activity	Aerobic	Push trainee
	Marksmanship	Maintain current level
Weight Control	Aerobic	Monitor performance
		Monitor performance
Aerobic	Aerobic	Monitor performance
		Work with trainee
Anaerobic	Aerobic	Monitor performance
		Work on techniques
VO2 Max	Aerobic	Carefully monitor
		Work on techniques

Integrated Warfighting Organization Strategies

Recent research on group dynamics in society suggests that individuals fall into one of three types:

i) cooperators who contribute to generating group benefits at some cost to themselves,

ii) free-riders who are unwilling to incur said costs and (iii) reciprocators who respond to others' behavior rather than being pure cooperators or free-riders. Furthermore, these traits appear to be rather evenly distributed suggesting that each have their own evolutionary advantages. Therefore, classes of boot camp recruits and units of the military are likely to be composed of these diverse groups of individuals.

Similar patterns have been observed in the athletic settings of sports teams and exercise groups.

For example, much of human behavior observed in athletics is embedded within groups, where the composite of individuals' cognitions, emotions and behaviors influence, and are influenced by, other group members. Key issues involved in successful groups are social identity, families, coordination/shared knowledge, the group as a vehicle for facilitating individual behavior change, social support regulating emotions, peer leadership and cultural perspectives in relation to group dynamics.

Therefore, a better understanding of group dynamics is key to raising individual and group performance. The capture of unit cohesive metrics allows new insights into the basic combat training pipeline from recruiting to performance in their first operational unit. Are we recruiting the right mix of talents? How can we adjust training to ensure a lower wash out rate while maintaining standards? Training Instructors will have additional quantitative measures for peer-to-peer evaluation and best practice analysis. We can analyze biometric information to identify our better team players. There will be a higher success rate allowing better use of resources. We look to advance the successes of the sports and various other communities in machine learning and data analytics to increase the combat effectiveness of the U.S. Army at an appropriate price point.

The capture of unit cohesive metrics allows new insights into the basic combat training pipeline from recruiting to performance in their first operational unit. Are we recruiting the right mix of talents? How can we adjust training to ensure a lower wash out rate while maintaining standards? Training Instructors will have additional quantitative measures for peer-to-peer evaluation and best practice analysis. We can analyze biometric information to identify our better team players. There will be a higher success rate allowing better use of resources. We look to advance the successes of the sports and various other communities in machine learning and data analytics to increase the combat effectiveness of the U.S. Army at an appropriate price point.

Differing commanders and their chain have different cultures and objectives and therefore should be able to demand, weighted against solider preferences, troops that best fit their specific needs. In deed most likely different squad formations of a BCT platoon could, should?, be deployed together if they have demonstrated cohesive traits that fit with open slots as their experience in BCT has already created a comradery that will further enhance their integration in the unit.

Our hope is in Phase I that we can interview commanders at various levels to get an idea on the qualitative metrics they are looking for that we can then look to map quantitatively to our senor data with machine learning. We have proposed several self report measures that get at individual strengths, psychological issues, and team cohesion. Further work in Phase II and III will focus on providing command customizable assessment modules based on this input. Further, we will develop approaches to incorporate unit level assessment data as a key input for machine learning, thus expanding beyond an individual view to key groupings of soldiers.

Phase III Commercialization Opportunities

Military Training Beyond Enlisted in the US Army into the Other Branches and Collation Partners

Integration into Operational Military with the Army and Beyond

Organized Team Sporting Applications

Real-Time Customer-Servers Metrics to Alert Managers to Disputes that need Escalated Resolution

In closing, the mil.API.fit team has compiled this presentation largley over the past 4 days and it has been the most exciting and productive environment I have ever been apart of. We hope that you find this as compelling as we do and look forward to working with US Army on this project.

Thank you for the thought provoking requirement.

-Jason Lawrence Lind

7 February 2018 18:23 CST

Principal Biographies

Melissa Flinn

Project Manager

Melissa Flinn earned a B.S. in Biomechanical Engineering from Marquette University with minors in mechanical engineering and biology. She has been working in market research as a project manager conducting studies with advanced analytical components for nearly 10 years. She currently leads a team of research staff to ensure deliverables and all milestones are hit as well as oversees projects to make sure that objectives are met and results are communicated in an actionable way.

Rodney Sparapani

Principal Investigator

Dr. Rodney Sparapani, PhD, is the Executive Vice President for Research at Transformation.run and an Assistant Professor of Biostatistics at the Medical College of Wisconsin (MCW) Milwaukee campus. Rodney has developed a research program in Bayesian machine learning at MCW with a particular emphasis in biomedical research. Mainly, his research interests involve artificial intelligence/ machine learning, Bayesian nonparametrics, big data and the surrounding computational challenges. He is co-author of the Bayesian Additive Regression Trees (BART) package which is free, open-source software for Bayesian machine learning: https://cran.r-project.org/package=BART. He has been invited to present at major international statistical meetings in New York (2002), Hawaii (2014) and Sardinia, Italy (2016) as well as presenting at major national/international meetings: San Diego (2012), Montreal (2013), Madison WI (2016), Chicago (2016), Storrs CT (2017), Houghton MI (2017) and Milwaukee (2017). He has recently published a first author article in the journal Statistics in medicine on Bayesian machine learning with survival analysis: http://dx.doi.org/10.1002/sim.6893. He has 36 other publications in high impact journals such as NEJM, JAMA, Lancet, Journal of Clinical Oncology, Medical Care, Cancer, Circulation, Statistical Methods in Medical Research and the Journal of Computational & Graphical Statistics.