U.S. ARMY COMBAT CAPABILITIES DEVELOPMENT COMMAND SOLDIER CENTER

Enhancing Data Drive Decision Making through Optimized Performance

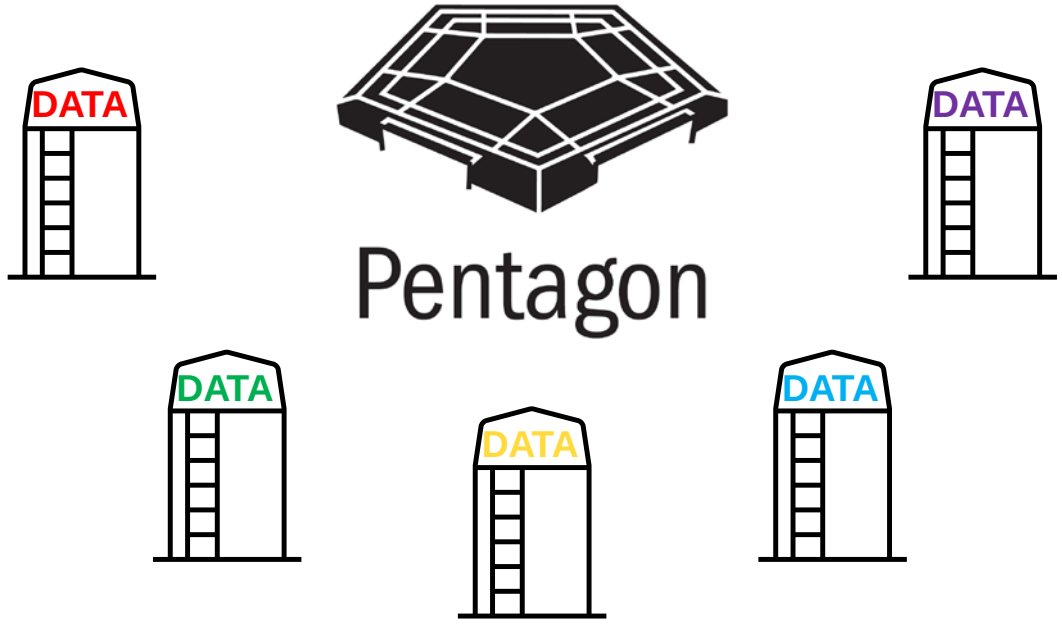
Seth Elkin-Frankston, Ph.D.

Cognitive Scientist

U.S. Army DEVCOM Soldier Center, Soldier Effectiveness Directorate, Performance Division, Cognitive Science and Applications Team

INTRODUCTION

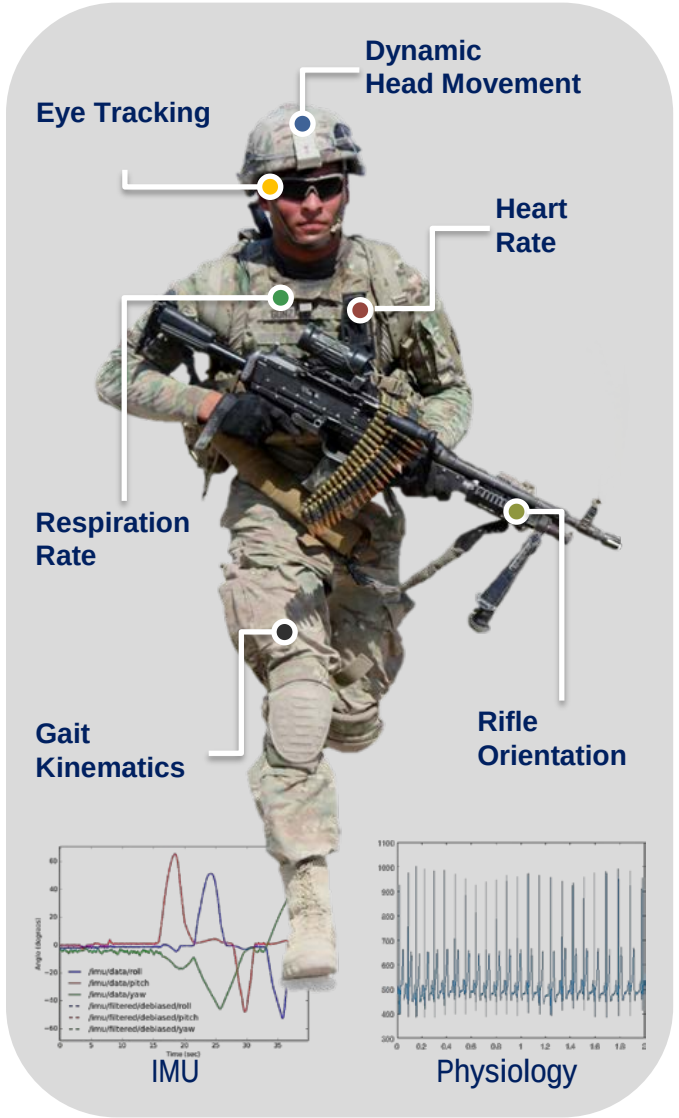
- The Department of Defense (DoD) has historically taken a conservative approach to data.
- Current practices silo data – thereby limiting effective data access reducing the overall value defense research and development programs.
- As the world becomes increasingly reliant on rapid access to trusted data, this position impedes development (e.g., data-driven decision aids, modeling and simulation, and artificial intelligence and machine learning tools).



HUMAN PERFORMANCE DATA

- **Human performance (HP) data:**
  - Physiological
  - Health
  - Physical
  - Cognitive
  - Lifestyle
  - Demographic
  - Social-Emotional
  - Performance

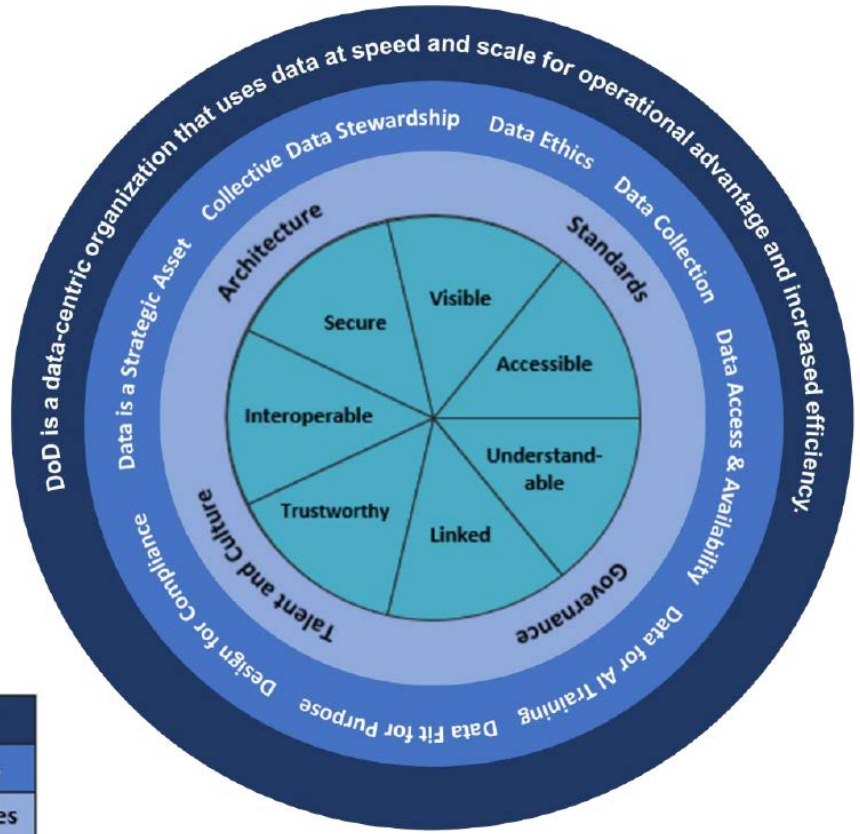
- **Contributing Factors:**
  - **Complexity:** HP data is complex without well-specified structure or markings – difficult to derive actionable in-sights without deep expertise
  - **Security:** Data may include information about Soldier performance and inadvertently categorized as personal identifiable information (PII)
  - **Prioritization:** A lack of data-focused prioritizations and infrastructure

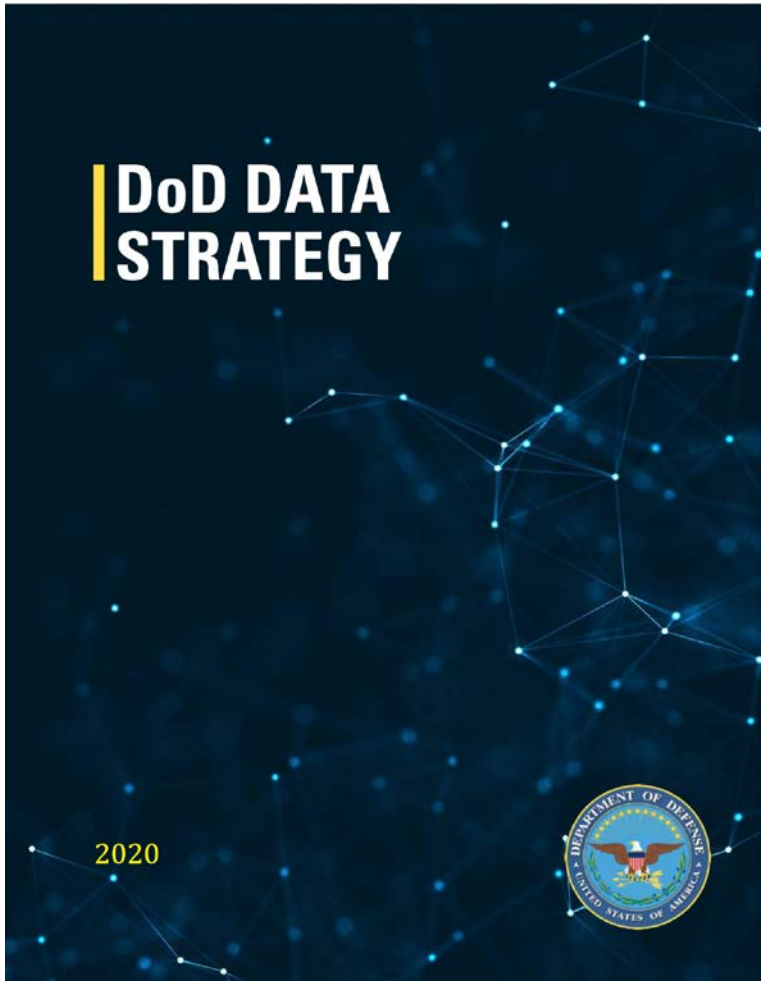


DOD DATA STRATEGY

- **In a cultural shift the DoD has begun to prioritize data as a strategic asset through improved data workflows and infrastructure**

DoD Data Strategy Framework:





https://media.defense.gov/2020/Oct/08/2002514180/-1/-1/0/DOD-DATA-STRATEGY.PDF

MASTR-E

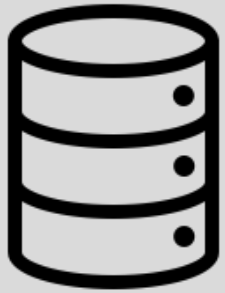
Measuring and Advancing Soldier Tactical Readiness and Effectiveness

MASTR-E OVERVIEW

BASELINE PREDICTORS
(X-Factors)

Holistic suite of cognitive, physical, social-emotional, and health measures (Baseline Battery)

DATABASE



USE CASES
(Outcome Measures)

Lab Studies

PEP-Lab

PEP-VR

Field Studies

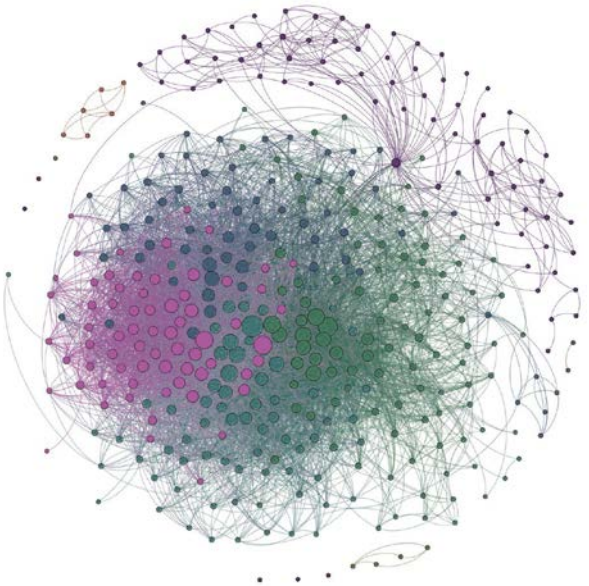
TSMA

SUPRA

Field Studies

72HFS

PREDICTIVE MODELS



MASTR-E OVERVIEW

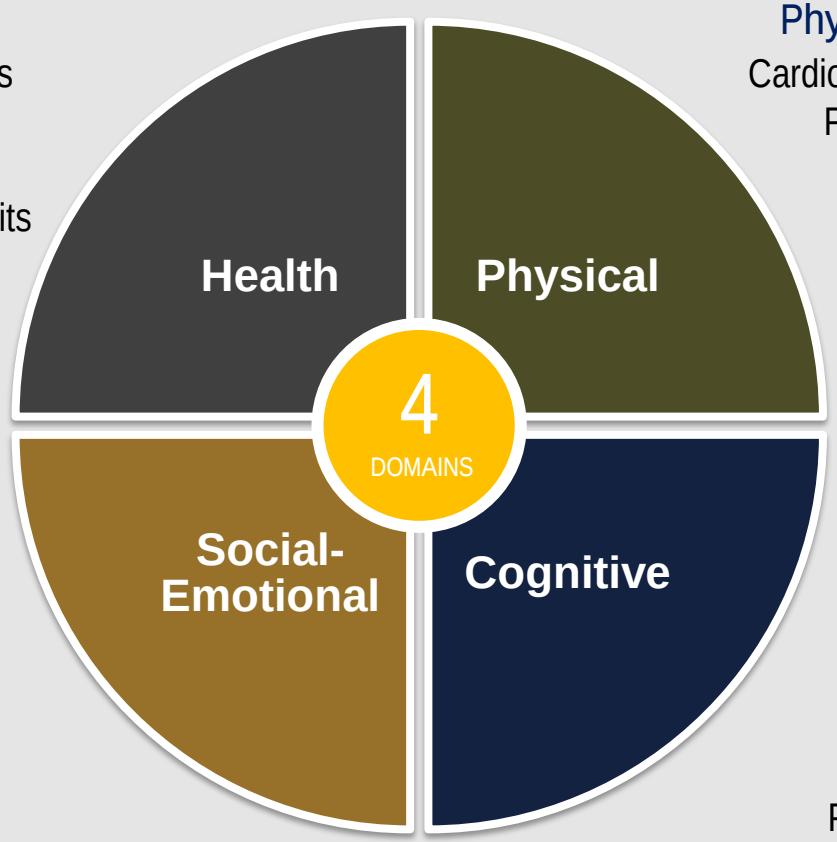
Work Packages	Protocols	Assessments	Metrics
Field Study	18-0XX	68	568
Decision Making	19-0XX	12	43
	OD-XX	19	46
SUPRA	20-0XX	25	766
TSMA	20-0XX	21	44
Prediction	17-0XX	27	143
	18-0XX	36	320
	19-0XX	54	959
	20-0XX	9	99
4	**9**	**271**	**2988**

HUMAN PERFORMANCE-CENTRIC DATABASE

MEASURES

Health Constructs
Immune System Status
Gut-Microbiome
Nutrition/Metabolism
Nutrition & Eating Habits
Hearing/Vision
Sleep

Physical Constructs
Cardiorespiratory Fitness
Power & Endurance
Movement Quality
Body Dimensions
Flexibility
Balance
Strength
Agility



Social-Emotional Constructs
Lifestyle
Motivation
Resilience
Personality
Trait Affect
Mindfulness
Impulsivity
Stress Reactivity
Emotion Regulation

Cognitive Constructs
Attention
Mental Flexibility
Decision Making
Spatial Cognition
Working Memory
Executive Control
Physical Constructs

- **Duration: 4 years**
- **Total Participants: ~800**
- **Independent Studies: 9**
- **Assessments: 271**
- **Metrics: ~3,000+**
- **Data Points: ~10M/Soldier**

MASTR-E

DATA DICTIONARY

Example of a completed data dictionary

Assessment_Task_Tool	Metric_Column_Variable	Description	Datatype	Units	Format	Min_Value	Max_Value	Response_Options	FileType	is_Required	Primary_Key	nullable	is_PII
Demographics	Age (Demographics Questionna	Age	integer	years					.csv	YES	NO	NO	YES
Demographics	Age (Demographics Questionna	Age	integer	years					.csv	YES	NO	NO	YES
GPS	BLFOR_longitudeUTM	The geographic	float	meters	########.###				.csv				NO
GPS	BLFOR_latitudeUTM	The geographic	float	meters	########.###				.csv				NO
GPS	OPFOR_longitudeUTM	The geographic	float	meters	########.###				.csv				NO
GPS	OPFOR_latitudeUTM	The geographic	float	meters	########.###				.csv				NO
Biomarker_BloodSampling_iSt	BloodBiomarker_sodium	blood electrolyt	float	mmol/L	###.##	100	180						NO
Biomarker_BloodSampling_iSt	BloodBiomarker_HCO2	blood gasses tak	float	mmol/L	##.##	1	85						NO
Attention Blink Task	Attentional Blink Response Late	Ability to select	float	ms	time	100	800	key	.csv	YES	NO	NO	NO
Speed of Processing Task	Speed of Processing Response L	Ability to detect	float	ms	time	100	500	key	.csv	YES	NO	NO	NO
Visual Search Task	Conjunctive Search High Demar	Ability to select	float	ms	time			key	.csv	YES	NO	NO	NO
Visual Search Task	Conjunctive Search Low Deman	Ability to select	float	ms	time			key	.csv	YES	NO	NO	NO
Visual Search Task	Feature Search High Demand Re	Ability to select	float	ms	time			key	.csv	YES	NO	NO	NO

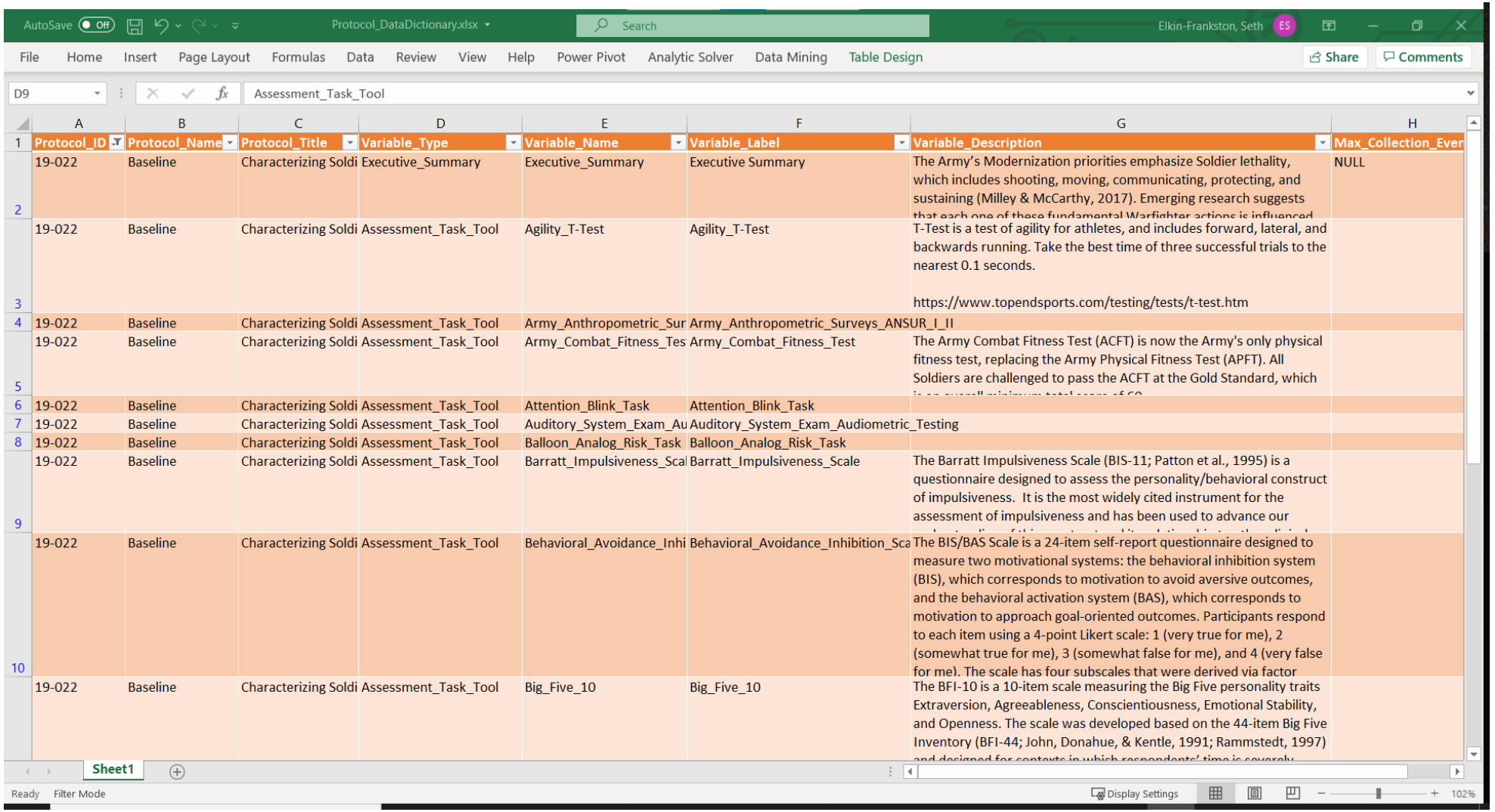
YES
NO
Other

Definitions

Column	Description	Use case	Example	Options	Notes
Assessment_Task_Tool	The name of the assesment, task, or data ta				
Metric_Column_Variable	The name of hte attribute, data field or varia				Assessment_Task_Tool will have one or
Description	Write a brief description of the data field	Education	Number of years		
Datatype	Data type of data field, for example, indicate	Years of Education	integer	text|string|ir	
Units	Units of data field, if known	Weight	kilogram		
Format	General format for data field, if known	Date of Birth	YYYYMMDD		For numeric values use "#" to indicate
Min_Value	Minimum value of data field, if known	Marksmanship Score	0		Not all variables will have a minimum
Max_Value	Maximum value of data field, if known	Marksmanship Score	100		Not all variables will have a maximum
Quantitative_Classification	Classification of quantitative data fields, if a	Acceleration - x	time-series	discrete|time	Not all variables will be quantiatitive or
Response_Options	Options for response, if any	Gender	M|F|O		
Code	Provide the key to any coded variables, if kn	Gender	M	M|F|O	M=1; F=0; O=-1
FileType	The file type or data extention of the data, if	Demographics	.csv		This is more of table level question but it is
is_Required	Enter y/n to indicate whether this field is re	date of birth	y	y|n	
Primary_Key	Is this column used to uniquely identify row	User_ID	y	y|n	A Primary Key is used to uniquely identify
nullable	Indicate y/n if null value is allowable	Gender	n	y|n	
is_PII	Does the variable contain Personal Identifial	date of birth	y	y|n	y=1; n=1

DATA DICTIONARY

Protocol Dictionary



AutoSave Off | Protocol_DataDictionary.xlsx | Search | Elkin-Frankston, Seth

File Home Insert Page Layout Formulas Data Review View Help Power Pivot Analytic Solver Data Mining Table Design | Share | Comments

D9 | Assessment_Task_Tool

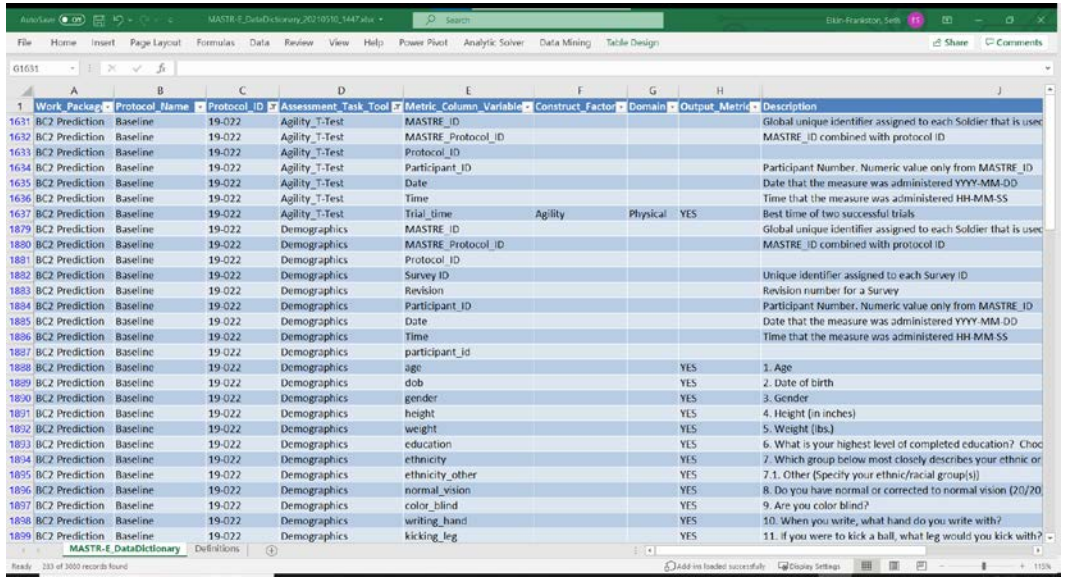
	A	B	C	D	E	F	G	H
1	Protocol_ID	Protocol_Name	Protocol_Title	Variable_Type	Variable_Name	Variable_Label	Variable_Description	Max_Collection_Ever
2	19-022	Baseline	Characterizing Soldi	Executive_Summary	Executive_Summary	Executive Summary	The Army's Modernization priorities emphasize Soldier lethality, which includes shooting, moving, communicating, protecting, and sustaining (Milley & McCarthy, 2017). Emerging research suggests	NULL
3	19-022	Baseline	Characterizing Soldi	Assessment_Task_Tool	Agility_T-Test	Agility_T-Test	T-Test is a test of agility for athletes, and includes forward, lateral, and backwards running. Take the best time of three successful trials to the nearest 0.1 seconds. https://www.topendsports.com/testing/tests/t-test.htm	
4	19-022	Baseline	Characterizing Soldi	Assessment_Task_Tool	Army_Anthropometric_Sur	Army_Anthropometric_Surveys_ANSUR_I_II		
5	19-022	Baseline	Characterizing Soldi	Assessment_Task_Tool	Army_Combat_Fitness_Tes	Army_Combat_Fitness_Test	The Army Combat Fitness Test (ACFT) is now the Army's only physical fitness test, replacing the Army Physical Fitness Test (APFT). All Soldiers are challenged to pass the ACFT at the Gold Standard, which	
6	19-022	Baseline	Characterizing Soldi	Assessment_Task_Tool	Attention_Blink_Task	Attention_Blink_Task		
7	19-022	Baseline	Characterizing Soldi	Assessment_Task_Tool	Auditory_System_Exam_Au	Auditory_System_Exam_Audiometric_Testing		
8	19-022	Baseline	Characterizing Soldi	Assessment_Task_Tool	Balloon_Analog_Risk_Task	Balloon_Analog_Risk_Task		
9	19-022	Baseline	Characterizing Soldi	Assessment_Task_Tool	Barratt_Impulsiveness_Scal	Barratt_Impulsiveness_Scale	The Barratt Impulsiveness Scale (BIS-11; Patton et al., 1995) is a questionnaire designed to assess the personality/behavioral construct of impulsiveness. It is the most widely cited instrument for the assessment of impulsiveness and has been used to advance our	
10	19-022	Baseline	Characterizing Soldi	Assessment_Task_Tool	Behavioral_Avoidance_Inhi	Behavioral_Avoidance_Inhibition_Sca	The BIS/BAS Scale is a 24-item self-report questionnaire designed to measure two motivational systems: the behavioral inhibition system (BIS), which corresponds to motivation to avoid aversive outcomes, and the behavioral activation system (BAS), which corresponds to motivation to approach goal-oriented outcomes. Participants respond to each item using a 4-point Likert scale: 1 (very true for me), 2 (somewhat true for me), 3 (somewhat false for me), and 4 (very false for me). The scale has four subscales that were derived via factor	
	19-022	Baseline	Characterizing Soldi	Assessment_Task_Tool	Big_Five_10	Big_Five_10	The BFI-10 is a 10-item scale measuring the Big Five personality traits Extraversion, Agreeableness, Conscientiousness, Emotional Stability, and Openness. The scale was developed based on the 44-item Big Five Inventory (BFI-44; John, Donahue, & Kentle, 1991; Rammstedt, 1997)	

Sheet1

Ready Filter Mode | Display Settings | 102%

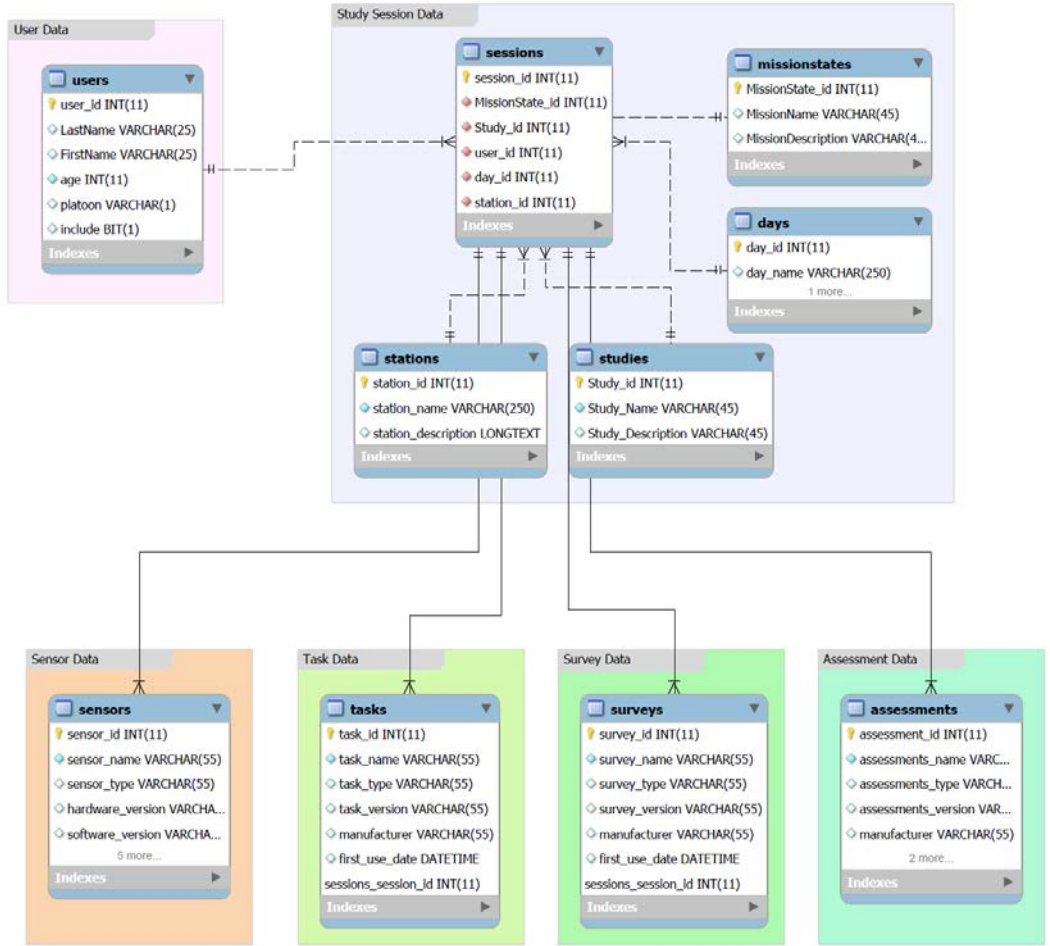
DATA MODEL

Data Dictionary



	A	B	C	D	E	F	G	H	J
1	Work_Packag	Protocol_Name	Protocol_ID	Assessment_Task_Tool	Metric_Column_Variable	Construct_Factor	Domain	Output_Metric	Description
1631	BC2 Prediction	Baseline	19-022	Agility_T-Test	MASTRE_ID				Global unique identifier assigned to each Soldier that is usec
1632	BC2 Prediction	Baseline	19-022	Agility_T-Test	MASTRE_Protocol_ID				MASTRE_ID combined with protocol ID
1633	BC2 Prediction	Baseline	19-022	Agility_T-Test	Protocol_ID				
1634	BC2 Prediction	Baseline	19-022	Agility_T-Test	Participant_ID				Participant Number. Numeric value only from MASTRE_ID
1635	BC2 Prediction	Baseline	19-022	Agility_T-Test	Date				Date that the measure was administered YYYY-MM-DD
1636	BC2 Prediction	Baseline	19-022	Agility_T-Test	Time				Time that the measure was administered HH-MM-SS
1637	BC2 Prediction	Baseline	19-022	Agility_T-Test	Trial_time	Agility	Physical	YES	Best time of two successful trials
1879	BC2 Prediction	Baseline	19-022	Demographics	MASTRE_ID				Global unique identifier assigned to each Soldier that is usec
1880	BC2 Prediction	Baseline	19-022	Demographics	MASTRE_Protocol_ID				MASTRE_ID combined with protocol ID
1881	BC2 Prediction	Baseline	19-022	Demographics	Protocol_ID				
1882	BC2 Prediction	Baseline	19-022	Demographics	Survey ID				Unique identifier assigned to each Survey ID
1883	BC2 Prediction	Baseline	19-022	Demographics	Revision				Revision number for a Survey
1884	BC2 Prediction	Baseline	19-022	Demographics	Participant_ID				Participant Number. Numeric value only from MASTRE_ID
1885	BC2 Prediction	Baseline	19-022	Demographics	Date				Date that the measure was administered YYYY-MM-DD
1886	BC2 Prediction	Baseline	19-022	Demographics	Time				Time that the measure was administered HH-MM-SS
1887	BC2 Prediction	Baseline	19-022	Demographics	participant_id				
1888	BC2 Prediction	Baseline	19-022	Demographics	age			YES	1. Age
1889	BC2 Prediction	Baseline	19-022	Demographics	dob			YES	2. Date of birth
1890	BC2 Prediction	Baseline	19-022	Demographics	gender			YES	3. Gender
1891	BC2 Prediction	Baseline	19-022	Demographics	height			YES	4. Height (in inches)
1892	BC2 Prediction	Baseline	19-022	Demographics	weight			YES	5. Weight (lbs.)
1893	BC2 Prediction	Baseline	19-022	Demographics	education			YES	6. What is your highest level of completed education? Choc
1894	BC2 Prediction	Baseline	19-022	Demographics	ethnicity			YES	7. Which group below most closely describes your ethnic or
1895	BC2 Prediction	Baseline	19-022	Demographics	ethnicity_other			YES	7.1. Other (Specify your ethnic/racial group(s))
1896	BC2 Prediction	Baseline	19-022	Demographics	normal_vision			YES	8. Do you have normal or corrected to normal vision (20/20
1897	BC2 Prediction	Baseline	19-022	Demographics	color_blind			YES	9. Are you color blind?
1898	BC2 Prediction	Baseline	19-022	Demographics	writing_hand			YES	10. When you write, what hand do you write with?
1899	BC2 Prediction	Baseline	19-022	Demographics	kicking_leg			YES	11. If you were to kick a ball, what leg would you kick with?

Data Model



HP DATABASE ARCHITECTURE

Data Collection

- Baseline Characteristics
- Physiological Measures
- Behavioral Performance

Baseline characteristic data, real-time physiological measures, and performance during operational task scenarios provide a robust dataset of Soldier performance

Ingestion

- Standardization
- Data Aggregation
- Load

Data are standardized, aggregated, and loaded into an interoperable, extensible, and accessible database that facilitates enhanced modeling and analytics

Soldier-Centric Cloud Database

AFC MADE ENVIRONMENT

Azure

HP Database(s)

MASTR-E HP Database

Feature Engineering

Hyper-parameter Tuning

Model Selection

Build Models

Model Training

Trained Models

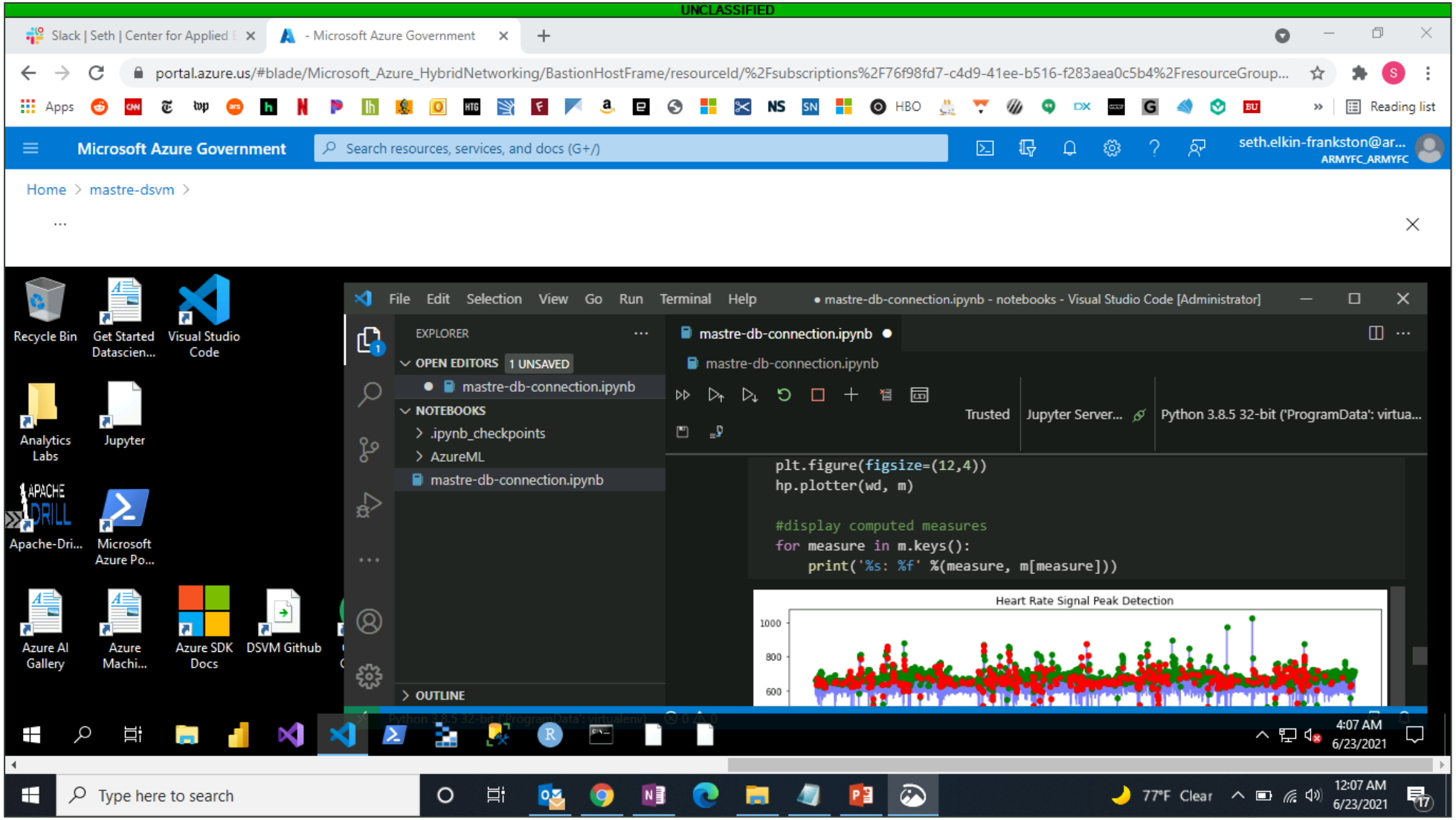
Power BI

Secured Access

Non-Gov Collaborators

Operational Community

HPDB-VM IN THE WILD

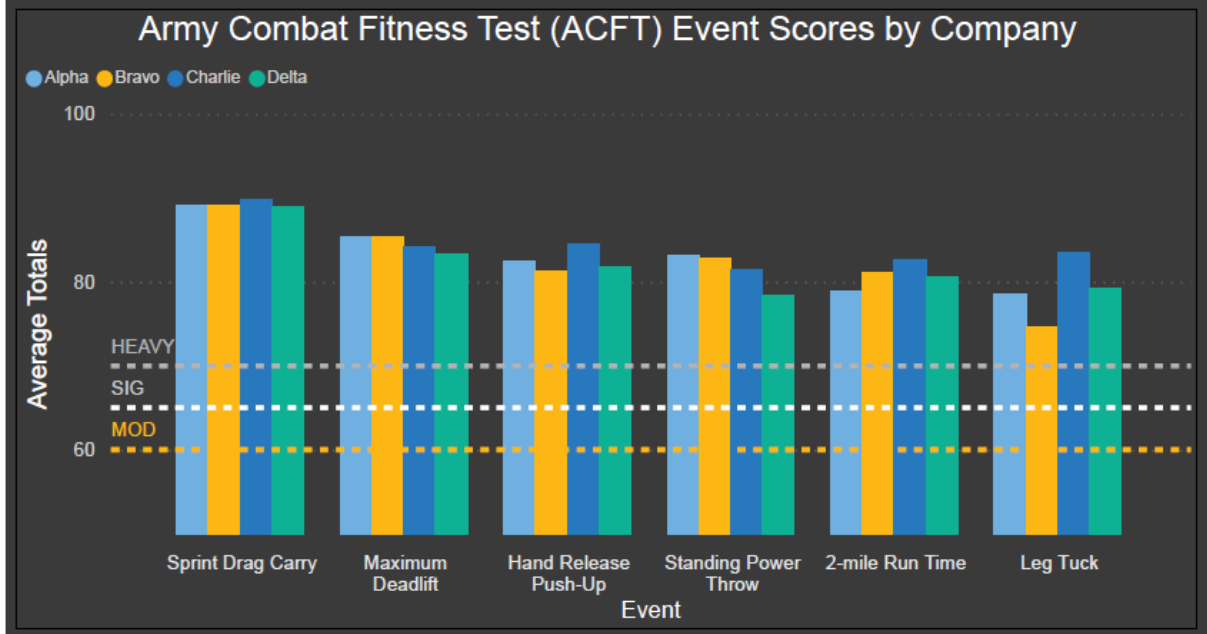


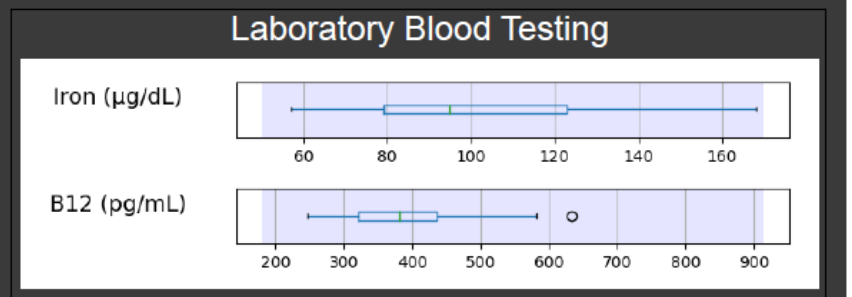
POWER BI DASHBOARD

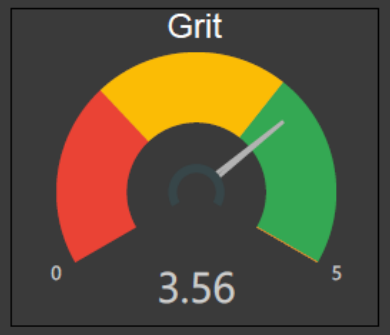
MASTR-E
MEASURING AND ADVANCING SOLDIER TACTICAL READINESS AND EFFECTIVENESS

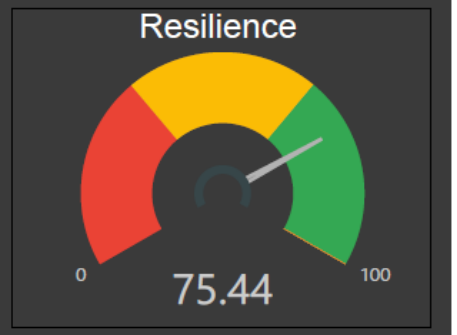
Baseline Soldier Data
Performance

150	23.1	2.6	0.31	69.5	181.8
Total Number of Soldiers	Average Age (yrs)	Avg. Years in Military	Avg. # of Deployments	Average Height (.in)	Average Weight (lbs.)









CONCLUSIONS AND NEXT STEPS

- **Successfully demonstrated ability to collect and store diverse human performance data.**
  - HP data can be readily ingested, stored, and rapidly accessed by authorized users in a DoD environment.
- **Verified functionality for roll-based authentication, and ability to analyze data across Army Science and Technology efforts.**
- **To maximize the utility of complex datasets future HP efforts must standardize and consolidate data – look to existing weapon system databases and Electronic Health Record systems for inspiration.**
- **To facilitate enhanced modeling and analytics privacy and security issues must be maintained while maintaining data that is interoperable, extensible, and accessible.**