Simulation Interoperability Standards Organization
"Simulation Interoperability & Reuse through Standards"

SISO-PN-015-2016

Product Nomination for

Human Performance Markup Language

Version 1.0

9 May 2016

SAC Approved: 02/02/2016 **EXCOM Approved: 05/09/2016**

Revision History

Version	Section	Date (MM/DD/YYYY)	Description
V1.0	All	05/09/2016	Published version

Table of Contents

SISO Product Nomination for
Human Performance Markup Language

SISO-PN-015-2016

1. Product title *

Human Performance Markup Language

2. Proponent name(s) and contact information

First name: Courtney **Last name:** Dean
Affiliation: Aptima, Inc.
Street Address: 12 Gill Street, Suite 1400
City: Woburn **State: MA** **Zip:** 01801
Phone #: 781.496.2455 **Fax #:**
Email: cdean@aptima.com

First name: Michael **Last name:** Tolland
Affiliation: Aptima, Inc.
Street Address: 12 Gill Street, Suite 1400
City: Woburn **State: MA** **Zip:** 01801
Phone #: 781.496.2458 **Fax #:**
Email: mtolland@aptima.com

Additional proponents:

Name	Organization	Email Address
Michael Tolland	Aptima, Inc.	mtolland@aptima.com
Mike Paley	Aptima, Inc.	paley@aptima.com
Charles Amburn	Army - STTC	charles.amburn@us.army.mil
Courtney Dean	Aptima, Inc.	cdean@aptima.com
Mike Hruska	Problem Solutions	mike@problemsolutions.net
Rodney Long	Army - STTC	rodney.a.long3.civ@mail.mil
Scott Galster	AFRL	scott.galster@us.af.mil
Steven Burnett	ARI	steven.f.burnett.civ@mail.mil
Timothy Burns	InnovaSystems	tburns@innovasi.com
Webb Stacy	Aptima, Inc.	wstacy@aptima.com
Beth Atkinson	NAWC-TSD	Beth.atkinson@navy.mil
Eric Watz	Lumir	eric.watz.ctr@us.af.mil
John Costa	RePubIT	jbcosta@repubit.com
Alex Walker	MedStar SiTEL	alex.walker@email.sitel.org

3. Type of product(s) *(*Balloted Products)*

X	Standard*
	Guidance*
	Reference
	Administrative

The SISO Standards Product will be produced IAW the approved SISO-STD Standards Product Template found in the SISO's Digital Library Templates folder here:
https://www.sisostds.org/DigitalLibrary.aspx?EntryId=32979

The SISO Standards Product will be published on the Approved Standards webpage here:
https://www.sisostds.org/ProductsPublications/Standards/SISOStandards.aspx

The associated XML-schema data files will be published on the SISO Product Data Files webpage here:
https://www.sisostds.org/Schemas.aspx

4. Product description

Human Performance Markup Language (HPML) is an XML-Schema-based language intended to cover all meaningful aspects of human performance measurement in various training and operational environments. The HPML hierarchy enables the representation of both generic concepts (e.g., measurements and assessments) and mission specific concepts (e.g., instances of measurements and instances of assessments) necessary for capturing the experiences associated with human performance and human behavior. Specifically, it is an XML based language designed to express performance measurement concepts in a format that is both machine and human readable. It enables the explicit combination and transformation of performance data into performance measurements and assessments. This allows measures to be constructed independent of any specific training or operational system. At a basic level, the performance measurement instructions defined in HPML can be used to specify the system data elements to be collected, the calculations to use to process the data, and when to produce performance measurement results.

At a high level, HPML is broken up into many different sub schemas that represent the different parts of the overall HPML schema. Each part of the schema has different dependencies that work together to calculate measures and assessments on a given data source. Figure 1 provides a schema dependencies diagram that shows the links between each of the main groups within the schema.

The schema is separated into six distinct groups, 1) HPML, 2) Computation, 3) Results, 4) Assessments, 5) Measures, and 6) Instances and Periods. These groups make up the core components of HPML and can be added to or expanded with additional links in the schemas. Each group is described in more detail below:

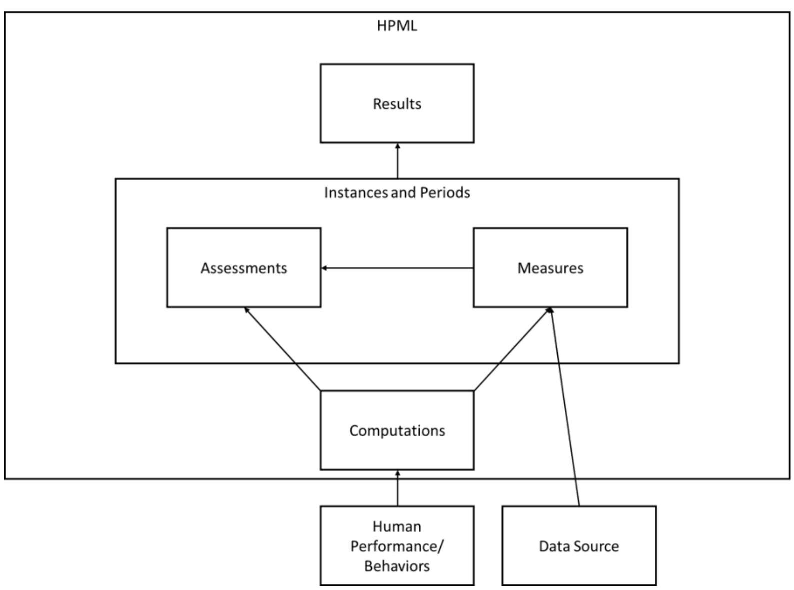

Figure 1. HPML high level conceptual architecture

HPML. The HPML group refers to shared and top level elements (e.g., MeasureDefinition element) and attributes that all the other schemas must reference. This schema organizes all sub schemas that make up the standard HPML Core.

Computation. The Computation group refers to the schemas that represent the algorithms, triggers, and other computational components of HPML. These components can be combined with both the Assessments and Measures groups.

Measures. The Measures group refers to schemas representing the linking of data sources and computation to produce measurement outputs from a given data source.

Assessments. The Assessments group refers to schemas representing the assessment of a given measurement's values either by category (Expert, Novice, etc.) or Value (100%, 99%, 75%, 10.3, etc.).

Results. The Results group refers to schemas representing the output of both measures and assessments, detailing the information produced by specific measure points throughout an entire measurement period.

Instances and Periods. The Instances and Periods group refers to schemas representing the creation and use of measures and assessments for a given context. This schema defines the instantiation of HPML elements at specific points in time, or specific locations within space. Every element in this schema has a time and/or location component. Whether that time is very short or spans several years, whether the location is a small area or a line on a map, the schema refers to when and where data should be computed so that measures and results can be linked to specific places and times.

While the segregation of HPML into multiple separate, but dependent, schemas may produce an added level of complexity, ultimately, it provides the language with the added benefit of extensibility, which is critical in the ever changing technological landscape of training environments. The modularity of HPML's design allows for the replacement of sub-schemas with new additions while still maintaining the core capability. Under this design, a developer could extend the schema to allow for more functionality, or to test new components for a specific context, while keeping the base schema intact. A full description of the HPML schema can be found in the HPML user guide posted to the HPML SISO Study Group.. The User Guide includes descriptions of elements of the schemas (e.g., definitions of what the elements and attributes mean) as well as example measures utilizing these schemas.

5. Identification of the community to which product applies

Any M&S community involved in the training of human participants can benefit from the use of HPML. HPML provides a method for defining how to measure and assess humans or other systems against data provided from any Model or Simulation. In the near term, the community that will derive the largest benefit will be the training simulation community.

6. Problem(s) and/or issue(s) that the proposed product addresses

There is no standard method of expressing human performance measures. Currently performance measurement approaches rely on proprietary measurement built into individual simulations. This approach makes it difficult to move measurement between simulations and easily make modifications to how performance will be calculated or assessed. Also, there is a need for transparency in the calculations so that it is clear how measurements were calculated and assessed. Typically, this is not possible because many simulation based measures are locked within the Simulation software and not readily available to the end user.

a) State the specific need or requirement for this product in the community (1-2 sentences)

Within training environments there is a need for an open and standard way to measure and assess human and system performance against varying data sources. Such an approach should be a clear and descriptive way to understand the methods by which measurement and assessment are obtained on specific data points.

b) Provide a detailed discussion of the specific need or requirement for this product in the community.

The amount of data we generate grows daily. According to Eric Schmidt, former CEO of Google,

There were 5 exabytes of information created between the dawn of civilization through 2003, but that much information is now created every 2 days. (Schmidt, 2010 at Google's Atmosphere convention.)

This is true, in microcosm, in the training domain. Simulator repositories, computer-based instruction systems, and educational data management systems are full of video and audio files and archived responses to a wide variety of scenario situations and computer and machine-scored tests.

These data might be used to improve training in many ways, ranging from student feedback to adaptive training to evaluating different approaches to instruction. Unfortunately, in large part this has not happened. A primary reason is that the raw data need to be transformed in order to be useful.

This need is particularly salient in the ever evolving context of Live, Virtual, and Constructive (LVC) training. The addition of virtual and constructive elements to live training provides a level of control, convenience, and cost savings that makes live training a more practical solution. On the flip side, the addition of live elements to virtual and constructive training adds a level of realism that cannot be otherwise achieved. There is even the possibility of dynamically mixing these three kinds of elements. In the future, for example, when a maintenance problem develops with an actual aircraft, a virtual or even constructive entity might be substituted during training [1], and when there is trouble with the distributed simulation network that prevents live elements from seeing virtual or constructive elements, an actual aircraft might be substituted instead to enable the continuation of a training exercise.

Standardized modern simulation network protocols such as the High Level Architecture (HLA), Distributed Interactive Simulation (DIS), and Test and Training Enabling Architecture (TENA) make much of this possible, because they track and report on entities in a similar fashion whether they are live, virtual, or constructive. Because LVC training, especially LVC after-action reviews, are enhanced by objective, automatically computed performance measurements, it follows that a standard that can track and measure the performance of all three kinds of entities—that is, all the entities in the training battlespace, even as they change from live to virtual to constructive—is an important component of training of the future [2].

This in turn implies a need to go beyond device-specific measurement for LVC training. Measurements that are embedded in actual platforms, simulators, or even in behavior specifications for constructive entities, might not be transportable from medium to medium unless they agree on a standard for expressing and capturing performance measurement and assessment.

HPML holds the promise to be such a standard. It is agnostic with respect to the kinds of data used to compute measurements and assessments. It also externalizes measurement and assessment specifications, so that they can be used with any data source. As long as there is a way to map data from live platforms to simulators to constructive entities, HPML will be able to specify measurements that can seamlessly adapt to LVC environments as the environments themselves change.

HPML was designed to provide a means by which the transformation of raw data can be specified and the results recorded [3]. Two particularly important transformations are measurements, which are direct computations on the raw data, and assessments, which are computations on the measurements and the data that provide interpretation for measurement. To help clarify this distinction, consider the following examples. A student's answers on an exam are the raw data, percent correct on the exam is a measurement, and conversion of percent correct into a letter grade is an assessment. A pilot's glideslope position as they land the plane is the raw data, computation of the deviation from ideal glideslope during landing is the measurement, and categorization of the size of that error as acceptable or unacceptable is the assessment. A leader's electronic communications with her distributed team are the raw data, the patterns inherent in those communications are the measurements, and the extent to which her communication patterns match the communication patterns of known high-performance leaders is the assessment. And so on.

Measures and assessments make it considerably more likely that training professionals can use available data to provide more effective training experiences [4]. Historically, getting such measurements and assessments has meant making changes to the simulation equipment, training environment, or computer-based instruction system. Because this can be time-consuming and expensive, HPML is designed to adapt to a wide range of data sources, and to provide a way to specify measurements and assessments that is external to specific training equipment. This also enables applications in the training environment to share measurement and assessment specifications and results, thus encouraging a healthy ecosystem of training applications that will result in improved training experiences.

c) Have you investigated similar products in the community to ensure no overlap exists?

We have reviewed all existing Study Groups, Product Groups in SISO and determined that nothing overlaps sufficiently with HPML. We have also reviewed standards relating to the logging of human activity (i.e., Activity streams , Contextualized Attention Metadata [CAM], Tutor Message Format [TMF], Experience API [xAPI], and eXtensible Event Stream) and found that none of them provide a specific nor sufficient coverage of human performance within simulation or learning environments.

7. Indication of the maturity of the product

a) Detailed description on how the problem/issue will be solved (approach)

The problem revolves around the need for a configuration/document that can be read by both humans and machines and can be used to measure and assess the performance of humans operating in various simulation environments. The solution that we propose is the adoption of an XML schema called HPML (Human Performance Markup Language) as a universal approach to capturing human performance in simulation environments.

The schema will create a soft link between specific data points and measurement allowing for easier transition between different simulations with different data structures. We call these Data Requests. These elements in the schema allow you to query data from any data source like a Structured Query Language (SQL) statement. The Data Requests can then be linked to a measurement. If the data source changes only the Data Request must change - not the entire measurement.

This standard schema outlines how a measurement should be calculated using a format that is generic enough to allow people to create and implement new algorithms, but strict enough to provide a recognizable structure that people can use to understand the intent of the measurement or assessment. Standardization facilitates interoperability. With the consistent use of a standard schema, measures and assessments become transferable and usable in multiple domains (e.g., different weapon platforms) and interpretable by different simulators and users within the modeling and simulation community. If people cannot understand how a measurement or assessment is derived, then we will lose the ability to be interoperable.

The schema also adds ways of representing the output of these measurements and assessments. This is important for traceability as it allows you to link the results back to the measures and assessments that produced them. We have created both guidelines on how this schema and resulting documents can be validated, and user guides describing best practices for using the schema and resulting documents.

b) Brief discussion on the maturity of the proposed product.

HPML has been used internally at Aptima for over 7 years and has evolved over those 7 years through review and implementation. The current version of HPML has been used in practice in many platforms and simulations. An HLA BOM was developed for the NASMP FOM called the HPM BOM, this BOM allows the passing of HPML via HLA. A comprehensive User Guide detailing the library of HPML schemas and schema elements has been produced and is available for reference.

c) Brief discussion on alternative approaches to the proposed product

The bottom line is that without a standard, uniform measurement across domains and platforms cannot exist. This leaves few viable alternatives other than continuing with the status quo. Detailed documentation of measurements for each platform is a good first step. It will allow people to better understand how measurement in a simulation works and how the results were gathered. However, this action must be repeated for each instance of measurement and for each platform into which measurement is introduced. The amount of labor required to achieve such documentation would likely be prohibitively time consuming.

xAPI is an alternatively that solves part of the problem outlined in this nomination. xAPI provides a standard format for passing measure results between learning management systems (LMSs) and other performance systems. However, it does not provide a standard way for defining and calculating human performance. Thus, while it can help pass data from multiple sources to a single integrated location, it does not pass any of the contextual information necessary to allow for interpretation of the results. Those exposed to the results from multiple sources may have only limited information from which to make assessments about the performance of their trainees.

d) Provide examples of prototypes of the proposed product or reasons why this product will not be prototyped.

Figure 2, below, provides an example of a simple performance measure implemented in HPML. In this example, the *DataRequest* identifies the object and its specific attributes that are required as input to a measurement. In this case, the *DataRequest* is requesting the **X,Y and Z** components of the **WorldLocation** attribute of the **Aircraft**, which has an **Id** equal to a value that will be supplied later as a parameter. The *SystemMeasureTemplate (part of the Measures group)* in our example uses the **WorldLocation** attribute of the **Aircraft** object to determine if the aircraft is currently above a prescribed altitude. The maximum altitude in this measure is defined as a *constant* of 4000 ft. in the *SystemMeasureTemplate.* A *Parameter* is also passed into the *SystemMeasureTemplate*, specifying the specific **Aircraft** from which to retrieve the **Altitude**. The use of a *Parameter* here allows the *SystemMeasureTemplate* to be reused for different aircraft. At the core of the *SystemMeasureTemplate* are two nested *MeasureComputations*. The inner most computation uses the *DataRequest* (**WorldLocation.X,Y,Z**) and passes the parameter defined in the template to the data request. The inner most computation also uses a function library plugin to compute the Altitude. The result of this computation is then used as an operand for the outer *MeasureComputation*, which compares the altitude value to a *constant* that points to the maximum altitude *ref* and determines if the aircraft's altitude is greater than the prescribed maximum altitude. The *SystemMeasureDefinition* finalizes the specification contained within the *SystemMeasureTemplate*. More specifically, the *SystemMeasureDefinition* defines any parameters that have been exposed, in this case the marking identifier of the aircraft.


```
<DataRequests>
    <DataRequest
        id="61300000-0000-0000-0000-000000000003"
        type="ENTITY"
        dataSourceRef="Source"
        from="Aircraft"
        select="WorldLocation.X,
            WorldLocation.Y,
            WorldLocation.Z"
        where="[Marking] == {Marking}">
        <Parameters>
        <Parameter name="Marking" type="String"/>
        </Parameters>
    </DataRequest>
</DataRequests>
<Measures>
    <SystemMeasureTemplate
        id="61300000-0000-0000-0000-000000000004"
        oneOff="true"
        name="CrossedAltitudeThreshold">
        <Description>
Measure to determine if aircraft broke the altitude threshold
        </Description>
        <Constants>
<Constant name="MaxAltitude" type="Int32" value="4000" units="feet"/>
        </Constants>
        <Parameters>
<Parameter name="Aircraft_Marking" type="String"/>
        </Parameters>
<MeasureComputation operator="Greater_Than" functionName="Math">
<MeasureComputation operator="default" functionName="Altitude">
                <Input
                type="datarequest"
                    ref="61300000-0000-0000-0000-000000000003">
                <Input
                type="parameter"
                ref="Aircraft_Marking"
                name="Marking"/>
            </Input>
    </MeasureComputation>
    <Input type="constant" ref="MaxAltitude"/>
    </MeasureComputation>
    </SystemMeasureTemplate>
</Measures>
<MeasureDefinitions>
    <SystemMeasureDefinition
        name="CrossedAltitudeThreshold"
            id="7F12964C-1DF7-E3FE-5FCF-E5870A1879A4"
            systemMeasureTemplateId="61300000-0000-0000-0000-000000000004"
      visibility="VISIBLE">
         <Input
       inputType="VALUE"
           name="Aircraft_Marking"
       type="String"
           value="AC1"/>
  </SystemMeasureDefinition>
</MeasureDefinitions>
```

Figure 1. Simple performance measure implemented in HPML

e) What impact will the proposed product have on the M&S community?

The incorporation of HPML as a standard will have several positive benefits to the M&S community including, briefly: measurement reuse, increased availability of the right data, increased understanding of trainee strengths and weaknesses, and reduced workload for instructors and system operators. HPML use will provide a way for the M&S community to create interoperable performance systems in simulation and modeling systems. It will enable users to leverage existing HPML performance measures in new domains and platforms with minimal modification reducing. The status quo customization for every implementation is costly and inefficient. Reuse will eliminate this, especially when reapplying measures to similar platforms (e.g., fifth generation fighters). The standard language of HPML being human readable enables community stakeholders to review and understand what measures are capturing, ensuring that the right data are being used. With increased accessibility to measures and the consequential reduced workload, more and richer measures can be applied to better assess trainees and more precisely diagnose proficiency gaps.

f) What impact will this proposed product have on the SISO community?

This will open a new avenue for SISO, which has historically focused solely on Simulation and Modeling systems and not on the humans using those systems for training, or understanding purposes.

g) What is the impact to the community on the LACK of this proposed product?

Human measurement is here to stay and will only grow in the future. Simulation systems will need a way to measure performance and end users will want a way to reuse performance measures defined across multiple simulation and modeling systems. This proposed product has had a long period of review and good backing from multiple domains. It provides a solution to a pervasive problem that will continue as the focus on human performance measurement remains forefront.

h) What are the domain implications for this proposed product?

This will work across multiple domains and be geared specifically at bringing Human performance to all domains. We hope to deliver a clear and standard way to define Human Performance across all domains.

i) State which SIW conference track takes an active interest in the development of this proposed product.

M&S Specialty (SPEC) Track

8. Planned compliance testing

Specification of compliance tests will be included in the standard, but there is no planned compliance service.

9. Schedule of product development milestones including reviews and reports

Approval phase

Study group final report and product nomination	August 2015
Approval of product nomination by SAC	April 2016

2015-2016: Technical specification of HPML

HPML PDG Kickoff	May 2016
Review and commenting on HPML 1.1	May - July 2016
First draft for PDG discussion	July-Aug 2016
HPML First draft approval	Aug - Sept 2016

Review and commenting on HPML 1.2	Oct - Nov 2016
Second draft for PDG discussion	December 2016
Second draft approval	Jan - Feb 2017
HPML User Guide update and Schema document	
User Guide and Schema for PDG discussion	December 2016
User Guide and Schema draft approval	Jan-Feb 2017
SAC / EXCOM processing	
HPML PDG provide Circulation Package to SAC	March 2017
SAC conduct Initial Ballot	April 2017
HPML PDG conduct Balloter Comment Resolution	2017 Spring TIM
HPML PDG document Comment Resolution in Tracking System	2017 Spring TIM
SAC conduct Recirculation Ballot (if needed)	July 2017
HPML PDG conduct Balloter Comment Resolution (if needed)	2017 Fall SIW – Sep 2017
HPML PDG document Comment Resolution in Tracking System (if needed)	Sep 2017
HPML PDG provide Product Approval Package to SAC for processing	Oct 2017
EXCOM approval	Dec 2017
SAC publish approved SISO Standard	Jan 2018
HPML PDG provide SAC Terms of Reference for Product Support Group (PSG)	Feb 2018

10. Candidate volunteers for the product development effort

Name	Organization	Email Address	Phone Numbers
Michael Tolland	Aptima, Inc.	mtolland@aptima.com	781.496.2458
Mike Knapp	Aptima, Inc.	knapp@aptima.com	321.710.3201
Courtney Dean	Aptima, Inc.	cdean@aptima.com	781.496.2455
Webb Stacy	Aptima, Inc.	wstacy@aptima.com	781.496.2437
Beth Atkinson	NAWC-TSD	beth.atkinson@navy.mil	407.380.4773
Eric Watz	Lumir	eric.watz.ctr@us.af.mil	937.938.4068
Alex Walker	MedStar - SiTEL	alex.walker@email.sitel.org	202.364.5180 ext. 9248
Marc Banghart	Wyle, Inc	marc.banghart@wyle.com	904-644-6605
Robert Pratten	Transmedia Storyteller, Ltd	robert.pratten@gmail.com	415-287-4150

11. Product review cycle

Product(s) should be reviewed in accordance with SISO policy and procedures. Most importantly, the product(s) should be reviewed as often as technology changes and/or new information becomes available.

Acronyms

BOM	Base Object Model
CAM	Contextualized Attention Metadata
CEO	Chief Executive Officer
DIS	Distributed Interactive Simulation
FOM	Federation Object Model
HLA	High Level Architecture
HPM	Human Performance Management
HPML	Human Performance Markup Language
LMS	Learning Management System
LVC	Live, Virtual, and Constructive
M&S	Modeling and Simulation
NASMP	Naval Aviation Simulation Master Plan
PDG	Product Development Group
SAC	Standards Activity Committee
SISO	Simulation Interoperability Standards Organization
SPEC	Specialty
SQL	Structured Query Language
TENA	Test and Training Enabling Architecture
TMF	Tutor Message Format
xAPI	Experience Application Program Interface
XML	Extensible Markup Language

References

Author Name	Document Title	Event Name	Date Published	Name of Publication
Roessingh, J. J. M., & Verhaaf, G. G.	Training Effectiveness of embedded training in a (multi-) fighter environment-a discussion paper	NATO HFM-169 Research Workshop on the Human Dimensions in Embedded Virtual Simulation	2010	proceedings of the NATO HFM-169 Research Workshop on the Human Dimensions in Embedded Virtual Simulation
Schreiber, B. T.	Transforming training: A perspective on the need for and payoffs from common standards		2013	Educational Publishing Foundation
Stacy, E. W., Merket, D., Freeman, J., Wiese, E., & Jackson, C.	A Language for Rapidly Creating Performance Measures in Simulators	2005 Interservice/Industry Training, Simulation & Education Conference	2005	In Proceedings of the 2005 Interservice/Industry Training, Simulation & Education Conference
MacMillan, J., Entin, E. B., Morley, R., & Bennett Jr, W.	Measuring team performance in complex and dynamic military environments:		2013	Military Psychology

Author Name	Document Title	Event Name	Date Published	Name of Publication
	The SPOTLITE method.			
Poepppelman, T., Hruska, M., Long, R., Amburn, C.	Interoperable Performance Assessment for Individuals and Teams Using Experience API	Second Annual GIFT Users Symposium	2015	Proceedings of the Second Annual GIFT Users Symposium