

# Ontology Based Annotation of Contextualized Vital Signs

Albert Goldfain<sup>1,\*</sup>, Min Xu<sup>1</sup>, Jonathan Bona<sup>2</sup>, and Barry Smith<sup>2</sup>

<sup>1</sup> Blue Highway Inc., Syracuse, NY, USA

<sup>2</sup> University at Buffalo, Buffalo, NY, USA

## ABSTRACT

Representing the kinetic state of a patient (posture, motion, and activity) during vital sign measurement is an important part of continuous monitoring applications, especially remote monitoring applications. In *contextualized vital sign* representation, the measurement result is presented in conjunction with salient measurement context metadata. We present an automated annotation system for vital sign measurements that uses ontologies from the Open Biomedical Ontology Foundry (OBO Foundry) to represent the patient's kinetic state at the time of measurement. The annotation system is applied to data generated by a wearable personal status monitoring (PSM) device. We demonstrate how annotated PSM data can be queried for contextualized vital signs as well as sensor algorithm configuration parameters.

## 1 INTRODUCTION

Vital sign measurements are often obtained without close clinical supervision. In hospital settings, ambulatory patient monitoring devices are used to track vital signs when a patient is away from the bedside [1]. Telemedicine applications permit a patient to take readings from a location that is remote to their provider [2]. The availability of consumer-grade devices coupled with easy-to-use, web-based health portals has fueled the adoption of vital signs monitoring as part of the Quantified-Self movement [3]. Users can now independently collect various sorts of data for fitness, health, wellness, and disease prevention. What is often lost in these scenarios, relative to a clinically supervised encounter, is an interpretation of the user's context of measurement. As remote continuous vital signs monitoring becomes a reality, the quality of vital signs data will increasingly rely on accurately inferring and representing measurement context in an automated way.

We use the term *contextualized vital sign* for the aggregate of a vital sign and some non-trivial aspect of its measurement context. Paradigmatic contextualized vital signs include: *night-time* blood pressure, *post-operative* blood pressure, *resting* respiratory rate, *premenopausal* body temperature, and *reclining* heart rate. Such descriptions are often applied to snapshot (episodic) measurements, and efficiently recorded and transmitted.

This paper presents a representation of contextualized vital signs that uses ontologies from the Open Biomedical Ontology (OBO) Foundry. We then use this representation in an automated annotation system for a personal status monitoring (PSM) device data stream. We have developed the PSM system to classify motion, body position, and high-acceleration events (such as falls) alongside vital sign measurements. The specific example we use throughout is the

representation of a user's body position during a pulse rate measurement. However, the annotation system can scale to cover the entire suite of classifiers.

The utility of having ontologically annotated PSM data is manifested in several applications:

- Maintaining sensor configuration for each classification.
- Maintaining classification algorithm configuration.
- Training set construction from annotated PSM data for data mining and machine learning.
- Querying PSM results using annotations as criteria.
- Describing semantic alarms for continuous monitoring applications [11].

These are discussed below along with potential extensions to the system.

## 2 BACKGROUND

### 2.1 Personal Status Monitoring System

Accelerometers are the most prevalent sensors used for body-position classification applications [12]. The PSM device is a wearable multi-sensor system consisting of fourteen tri-axial accelerometers and multiple vital sign monitors, each of which is unobtrusive and noninvasive for the user. The accelerometers are mounted in such a way as to minimize noise and are positioned at the hips (2), knees (2), shins (2), shoulders (2), forearms (2), wrists (2), chest (1), and head (1). Four unsupervised classification algorithms are applied to PSM data in order to infer user motion, body position, device orientation, and fall events. Each of these classifications relies on either acceleration measured at each sensor or data derived from the combination of such measurements. For example, body position is inferred using a classifier that takes as input the relative angles between limbs (Euler angles) or (in simple cases) the tilt of a limb relative to the anatomical axes. When all of the accelerometers are used in the classification of body position, the result can be visualized as a rough skeletal wire-frame configuration. Only a subset of the accelerometers is typically required to accurately classify crude body positions such as "sitting", "standing", and "lying down". For clinical applications, one or two active accelerometers will suffice. Vital sign monitors include a heart rate and respiration rate monitor mounted on the chest. Figure 1 illustrates three different embodiments of the PSM device.

\* To whom correspondence should be addressed: [agoldfain@blue-highway.com](mailto:agoldfain@blue-highway.com)



Figure 1. Military, athletic, and clinical embodiments of the PSM

PSM data are transmitted wirelessly to a laptop computer running data capture software. This is the initial input to a workflow for motion, body position, and anomaly classification. PSM data can be stored off as: raw measurement values, features over windows of raw data in the time or frequency domain (mean, standard deviation, peak, energy, etc.), or as a time series of class labels.

This system is intended to be an end-to-end solution for capturing, storing, visualizing, and computing over contextualized vital signs. The annotation system described in this paper operates over a log of PSM data.

## 2.2 Annotation System Requirements

Our annotation system for PSM-generated contextualized vital signs needs to represent all of the following:

- Sensor configuration (e.g., sampling rate, anatomical location of sensors)
- Body position prediction algorithm configuration (e.g., algorithm parameters)
- Data types, unit labels, time stamps, and ontological types for body position measurement data and vital sign measurement data.
- Data provenance: The particular sensor outputs and algorithm inputs responsible for a particular prediction.
- Data redundancy: Awareness that measurements from different sensors may be related or of the same type.

## 2.3 Existing Resources

Ontologies implemented in OWL-DL are well suited for our task because they provide formal descriptions of the relevant entities and relationships. Data annotated using types from an OWL-DL ontology can be represented in various machine-readable formats for storage, transmission, presentation, and query.

To our knowledge, no single ontology implements a systematic treatment of contextualized vital signs. The resources to represent contextualized vital signs exist in available terminologies, controlled vocabularies, and clinical models. However, most of these are currently inadequate for our application because they attempt to enumerate only a few combinations of vital signs and postural terms, rather than allowing such terms to be built up via cross-products of a vital sign term and another term. Our approach will be to utilize OBO Foundry ontologies, since they are built orthogonally with little domain overlap and, thus, are better suited to accommodate cross-product terms. Below we first review some

of the resources outside of the OBO Foundry that are relevant to our application domain.

### 2.3.1 SNOMED-CT

The Systematized Nomenclature of Medicine – Clinical Terms (SNOMED-CT) clinical terminology includes a small hierarchy of terms involving body position<sup>1</sup>:

-Position of body and posture

-Body position

-Body position for procedure

-Body position for blood pressure measurement

-Body position for height measurement

SNOMED-CT provides no natural-language definitions to distinguish these body position terms from other terms like ‘posture’. The most specific terms in this hierarchy apply only to blood pressure and height measurements. If a new term is required, e.g., ‘body position for heart rate measurement’, it must be requested from the maintainers of the terminology. In this case, the user would have to temporarily annotate using the parent term ‘Body position for procedure’. A solution to this would be to use two separate terminologies, one that maintains measurement terms, and another that maintains body position terms, and to generate terms as cross-products between these two sources.

### 2.3.2 NCI-Thesaurus

The National Cancer Institute (NCI) Thesaurus includes the term ‘vital signs position’ as a subclass of ‘body position’. The CDISC Vital Signs Position of Subject Terminology<sup>2</sup> is part of the CDASH controlled terminology referenced in the NCI Thesaurus. The following three terms are included:

**Sitting:** The state or act of one who sits; the posture of one who occupies a seat.

**Standing:** The act of assuming or maintaining an erect upright position.

**Supine:** A posterior recumbent body position whereby the person lies on its [sic] back and faces upward.

These provide natural language definitions, but assume that the positions in question are observed directly by a clinician. No logical relations link these terms to the anatomical entities involved, or to body position measurement algorithms used in a wearable device like the PSM.

### 2.3.3 openEHR

Electronic health records form another healthcare informatics system where contextualized vital signs may be represented. The openEHR system provides an open standard for representing data in EHR systems, including a representation of common data elements in a structure called an archetype [4]. Archetypes include free text and a coded type system. Archetypes for several vital signs already exist in openEHR.

<sup>1</sup> <http://purl.bioontology.org/ontology/SNOMEDCT/397155001>

<sup>2</sup> <http://evs.nci.nih.gov/ftp1/CDISC/SDTM/CDASH%20Terminology.html#CL.C78431.VSPOS>

Along with a ‘data’ portion where the measurement is recorded, these archetypes contain a ‘state’ portion, representing the context of measurement. For example the ‘heart rate and rhythm’ archetype includes four body positions: ‘Lying’, ‘Sitting’, ‘Reclining’, and ‘Standing’, along with definitions and an assumed default value. Other vital sign archetypes include specialized body positions when they are relevant to clinical measurement contexts. For example, the blood pressure archetype’s state segment includes the position, “*Lying with tilt to the left*: Lying flat with some lateral tilt, usually angled toward the left side. Commonly required in the last trimester of pregnancy to relieve aortocaval compression.” Such free-text contextual descriptions are valuable, but to fully realize their value, they must be annotated and linked to machine-readable representations outside of the EHR itself.

### 3 ONTOLOGY FOR CONTEXTUALIZED VITAL SIGNS

We use OBO Foundry ontologies for our annotation system for several reasons: such ontologies are open-source, actively developed by domain experts, use stable IRIs to denote types, honor the distinction between individuals and universals, share the Basic Formal Ontology (BFO) as a common upper-ontology<sup>3</sup>, and share the OBO Relation Ontology (RO) as a common source for relations [7]. OBO Foundry ontologies are implemented in machine-readable formats (OWL-DL and OBO Format), and are developed to maximize reuse of terms and relations. OBO Foundry *reference* ontologies are general enough for use across several domains. These are in contrast to *application* ontologies, which import terms and relations from reference ontologies and define new application-specific terms and relations for the purposes of a given application.

We have developed the Ontology for Contextualized Vital Signs (OCVS)<sup>4</sup> as an application ontology for PSM data annotations. A central feature of OCVS is its use of external terms and relations from OBO Foundry ontologies when possible. These external terms are used to form cross-product definitions and description logic restrictions.

For example *standing pulse rate* can be defined using a necessary and sufficient DL-restriction using the Vital Sign Ontology (VSO) [8], the Ontology for Biomedical Investigations (OBI) [9], and the Experimental Conditions Ontology (XCO) [10]:

“The pulse rate of an organism in the standing position”  
 vso:‘pulse rate’ AND  
**inheres\_in** SOME (obi:organism AND  
**bearer\_of** SOME xco:‘standing position’)

The relations (in bold) are standard relations from the OBO Relation Ontology. OCVS does not have to redefine new terms in order to construct the definition of ‘standing pulse rate’, and the same cross-product template can be used for different vital signs (from VSO) and body positions

(from XCO). OCVS imports terms from the Unit Ontology<sup>5</sup> (UO) to represent measurement units. All terms are imported using the MIREOT mechanism [5]. Throughout the paper, the source ontology for a term will be indicated via its OBO prefix (e.g., obi:‘measurement datum’ is the term ‘measurement datum’ from the Ontology for Biomedical Investigations).

#### 3.1 Representing Measurements

These imported terms are combined with relations from the Relation Ontology to form the basic representations for PSM measurements. A PSM measurement datum consists of three acceleration magnitude measurements, three tilt measurements (relative to each device axis), a signal vector magnitude measurement (SVM), a signal magnitude area (SMA) measurement, and a time stamp representing an interval.

The time stamp represents the total running time in seconds from the beginning of the data acquisition session. Acceleration is given in g-units (1 g = 9.8 m/s<sup>2</sup>), which OCVS asserts to be a type of acceleration unit. The angle of tilt, relative to the acceleration along each axis  $a$ , is computed as follows:

$$\text{tilt}(a) = \text{asin}(a) \times \frac{180}{\pi} \quad (1)$$

This produces an angular measurement of stationary tilt in the range [-90, 90] degrees. SVM is computed with each reading as a function of all three acceleration components at a particular time  $(x(t), y(t), z(t))$ :

$$\text{SVM}(x(t), y(t), z(t)) = \sqrt{x(t)^2 + y(t)^2 + z(t)^2} \quad (2)$$

The SMA is a running total of the absolute sum of component-wise accelerations over a window of  $N$  readings:

$$\text{SMA}(x, y, z) = \frac{1}{N} \left( \sum_{i=0}^N |x(i)| + \sum_{i=0}^N |y(i)| + \sum_{i=0}^N |z(i)| \right) \quad (3)$$

At any given time, the SMA is a sum over a window containing the last  $N$  readings. In OCVS, we assert that each of these measurement data is a part of the ‘PSM measurement datum’ with the same timestamp.

There are multiple ways of measuring qualities such as tilt. OCVS includes term annotations indicating the formulas used to derive relevant measurement data, thus providing metadata for consumers of annotated data as to how each input parameter to the body position classifier was derived.

‘PSM Measurement Datum’ in OCVS is a defined class. Defined classes, like universals, correspond to OWL classes. OWL object properties are used to implement OBO Foundry relations, and OWL data properties are used to link particulars (OWL individuals) to data. The parts of a PSM measurement datum and relevant relations are shown in Figure 2.

A single PSM measurement datum is the input to the body position classification algorithm which has as output a

<sup>3</sup><http://www.ifomis.org/bfo/>

<sup>4</sup><http://www.awqbi.com/ontologies/ocvs.owl>

<sup>5</sup><http://code.google.com/p/unit-ontology/>

body position measurement datum. This is captured in OCVS using the representation in Figure 3.

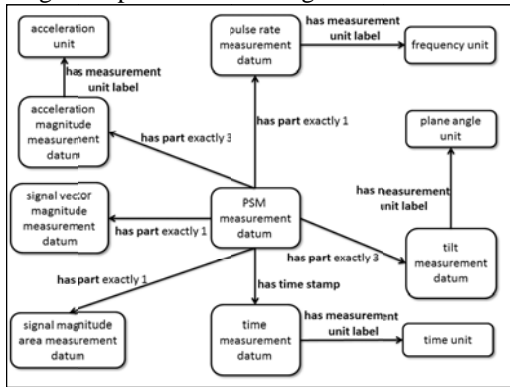


Figure 2. Parts of a Personal Status Monitor (PSM) measurement datum for a single accelerometer and single pulse rate sensor.

Currently, the ‘prone’, ‘sitting’, ‘standing’, and ‘supine’ position terms are imported from XCO. Various other clinically significant body positions have been identified (e.g., ‘decubitus position’, ‘Sims position’, ‘knee-chest position’). These are often specific to a medical procedure and beyond the scope of OCVS. Another application ontology could use OCVS’ representation scheme to represent these positions.

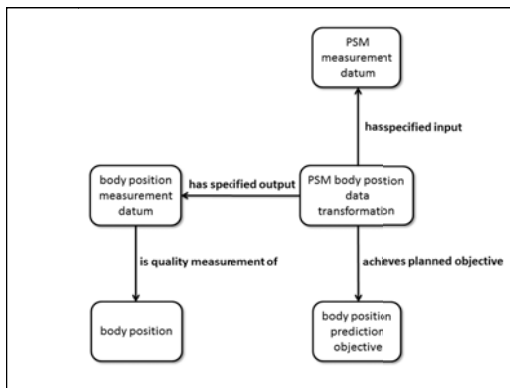


Figure 3. The body position classification algorithm is represented as an ‘obi: data transformation’

### 3.2 Accelerometer Configurations

Of particular interest for our application is a mapping of body positions to a minimal sensor set required to recognize those positions. For example, a single accelerometer positioned at the user’s sternum can tell whether the user is bending forward or backward along the sagittal plane, but cannot differentiate sitting from standing. Inversely, an accelerometer positioned at either of the user’s hips can differentiate sitting from standing, but cannot tell which way the user is leaning, since this must be measured at the upper body. If two accelerometers are used, one on a hip and another on the chest, then both measurement blind spots can be overcome. In fact, a new set of positions can be described in the two-accelerometer configuration that is a composition of positions recognized in either one-accelerometer configuration:

TABLE I. MAPPING BODY POSITIONS TO MINIMAL SENSOR SETS

Body Positions	Minimal Sensor Set Description
‘upright’, ‘bending forward’, ‘bending backward’	• accelerometer AND <b>has_anatomic_configuration</b> ONLY chest
‘sitting’, ‘standing’	• accelerometer AND <b>has_anatomic_configuration</b> ONLY (‘right hip’ OR ‘left hip’)
‘seated reclining’, ‘standing bending backward’, ‘sitting upright’, ‘standing upright’, ‘sitting bending forward (tucked/bracing)’, ‘standing bending forward’	• accelerometer AND <b>has_anatomic_configuration</b> ONLY chest • accelerometer AND <b>has_anatomic_configuration</b> ONLY (‘right hip’ OR ‘left hip’)

By explicitly encoding these constraints, we can check if the output class labels for our body position classification algorithm are appropriate given the active accelerometers before even attempting to invoke the algorithms.

In order to represent this information ontologically, we need to be precise about what it means for an accelerometer to have a particular anatomic configuration. We will assert that  $a$  **has\_anatomic\_configuration**  $c$  when:

1.  $a$  is a ‘tri-axial accelerometer’
2.  $c$  is an fma:‘anatomical entity’
3.  $a$  **adjacent\_to**  $c$
4. Each acceleration measured by  $a$ , along the device  $x$ ,  $y$ , or  $z$  axis, **is\_proxy\_for** the acceleration of **part\_of**  $c$ , along (a transformation of) the corresponding anatomical axis.

To capture the first condition, we assert that ‘Freescale MMA7660FC Triaxial Accelerometer’ (our chosen accelerometer type) is a ‘triaxial accelerometer’, and that ‘triaxial accelerometer’ is a obi:‘measurement device’. We use the Foundational Model of Anatomy (FMA) Ontology to represent anatomical entities. A small subset of FMA terms is imported into OCVS for the PSM application. The OBO **adjacent\_to** relation covers spatially proximal continuants. The fourth clause ensures that the measured values coming from PSM accelerometers are representative of the accelerations of the anatomical entities to which they are attached. OBI’s  $c1$  **is\_proxy\_for**  $c2$  relation holds between continuant instances  $c1$ ,  $c2$  when the measurement of  $c1$  is taken to determine what a measurement of  $c2$  would be (if  $c2$  were directly measurable). Since all the relevant anatomical entities for the PSM device are larger than the accelerometer, the measured value is only a proxy for that part of the anatomical entity that is closest to the sensor. The OCVS also imports FMA terms for each of the anatomical planes (‘coronal plane’, ‘median sagittal plane’, and ‘transverse plane’) and axes at their intersections.

However, since the anatomical axes are relative to the whole organism, the coordinate system must undergo a transformation to match up with the orientation of accelerometers at other anatomical locations. The relation is labeled **has\_anatomic\_configuration** to differentiate it from other device configuration parameters such as sampling rate.





to be integrated into the PSM platform without invalidating previously annotated (legacy) PSM data.

## 5 QUERYING ANNOTATED PSM DATA

We utilize the SPARQL to query annotated RDF-formatted PSM data. Annotated PSM data is queried locally using the ARQ command-line tool from the Apache Jena framework.

The following is part of a SPARQL query that returns all of the PSM measurement data measured using an accelerometer positioned at the sternum in which the inferred body position is ‘Bending Backward’:

```
SELECT DISTINCT ?psmmd
WHERE
{
  ?psmdevtype rdfs:label "PSM device"@en .
  ?cfg rdfs:label "Sternum"@en .
  ?psmmdt rdfs:label "PSM measurement datum"@en .
  ?d rdf:type ?psmdevtype .
  ?d part_of: ?psmdevpart .
  ?psmdevpart has_anatomic_configuration: ?cfg .
  ?psmmd measured_using: ?d .
  ?psmmd rdf:type ?psmmdt .
  ?alg has_specified_input: ?psmmdt .
  ?alg has_specified_output: ?bpmnd .
  ?bpmnd has_body_position_measurement: "Bending Backward"^^rdfs:Literal .
}
```

The query results are bound to `?psmmd` and represent PSM measurement data that satisfy the criteria in the WHERE clause. This query exemplifies several different search criteria we may apply to the annotated PSM data set. If we are interested in the details of the configuration (e.g., the devices used, their sampling rates, and their anatomical configurations), then we could expand the query on the results bound to `?psmdevpart`. If we are interested in the contextualized vital sign measurement value, we can expand the query on the results bound to `?psmmd` and examine its measurement data parts. If we want to obtain details about the algorithm configuration, we can examine `?alg`.

From a user interface perspective, it is easier to provide a web-based form from which queries can be constructed. We are implementing scripts to programmatically generate queries via the Graphite PHP Linked Data library<sup>7</sup>.

## 6 CONCLUSION

OCVS provides a representation of vital sign measurement context using the OBO Foundry ontologies. On the strength of cross-product definitions from orthogonal, independently developed ontologies, we are able to create descriptions of body positions, configurations, and queries in a compositional way. OCVS metadata captures enough domain knowledge to serve as a meaningful component of a pattern classification pipeline.

The semantic web standards used to build our annotation system enable decentralized development, storage, and query of resources. Further development on OCVS (or any of the OBO Foundry ontologies on which it relies) will not disrupt the data acquisition and classification routines.

OCVS is currently used to annotate continuous raw sensor measurement data. As such, the annotated PSM data is at the finest granularity. Currently, such data only need to be transmitted when an episodic reading is taken. In applications requiring more continuous transmission, OCVS-based

annotation can be applied to more coarse-grained data such as feature sets or sets of classifier outputs. A switch in data granularity will only require extension of the ontology rather than a switch of ontologies.

We believe that using OBO Foundry ontologies and semantic web standards can serve as the core knowledge representation for contextualized vital signs. Such a representation can be extended to perform further contextualization (e.g., disease-based contextualization) depending on the requirements of the particular application.

## ACKNOWLEDGEMENTS

We would like to acknowledge the useful comments of three anonymous reviewers.

## REFERENCES

- [1] Davenport, D. M., Deb, B., and Ross, F. J. (2009), Wireless Propagation and Coexistence of Medical Body Sensor Networks for Ambulatory Patient Monitoring, *Proc. Body Sensor Networks (BSN2009)*, IEEE Press, 41–45.
- [2] Pattichis, C. S. (2002), Wireless telemedicine systems: an overview, *Antennas and Propagation*, **44**(2): 143–153.
- [3] Swan, M. (2009), Emerging Patient-Driven Health Care Models: An Examination of Health Social Networks, Consumer Personalized Medicine and Quantified Self-Tracking, *International Journal of Environmental Research and Public Health*, **6**(2): 492–525.
- [4] Beale, T. (2003), Archetypes and the EHR, *Advanced Health Telematics and Telemedicine*, IOS Press, 238–244.
- [5] Courtot, M., Gibson, F., Lister, A. L., Malone, J., Schober, D., Brinkman, R. R., and Rutenber, A. (2011), MIREOT: The minimum information to reference an external ontology term, *Journal of Applied Ontology*, **6**(1): 23–33.
- [6] Smith, B., Ashburner, M., Rosse, C., Bard, J., Bug, W., Ceusters, W., Goldberg, L. J., Eilbeck, K., Ireland, A., Mungall, C. J., OBI Consortium, Leontis, N., Rocca-Serra, P., Rutenber, A., Sansone, S., Scheuermann, R. H., Shah, N., Whetzel, P. L., and Lewis, S. (2007), The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration, *Nature Biotechnology*, **25**: 1251–1255.
- [7] Smith, B., Ceusters, W., Klagges, B., Köhler, J., Kumar, A., Lomax, J., Mungall, C., Neuhaus, F., Rector, A. L., and Rosse, C. (2005) Relations in biomedical ontologies, *Genome Biology* **6**(5): R46.
- [8] Goldfain, A., Smith, B., Arabandi, S., Brochhausen, M., and Hogan, W. R. (2011), Vital Sign Ontology, *Proceedings of the Workshop on Bio-Ontologies*, ISMB: 71–74.
- [9] Brinkman, R. R., Courtot, M., Derom, D., Fostel, J. M., He, Y., Lord, P., Malone, J., Parkinson, H., Peters, B., Rocca-Serra, P., Rutenber, A., Sansone, S., Soldatova, L. N., Stoeckert, C. J., Turner, J. A., Zheng, J., and OBI Consortium (2010), Modeling biomedical experimental processes with OBI, *Journal of Biomedical Semantics* **1**(Suppl 1): S7.
- [10] Shimoyama, M., Nigam, R., McIntosh, L. S., Nagarajan, R., Rice, T., Rao, D. C., and Dwinell, M. R. (2012), Three ontologies to define phenotype measurement data, *Frontiers in Genetics* **3**(87).
- [11] Goldfain, A., Chowdhury, A. R., Xu, M., DelloStritto, J., and Bona, J. (2011) “Semantic Alarms in Medical Device Networks”, *Proceedings of Third Joint Workshop on High Confidence Medical Device Software and Systems/Medical Device Plug-and-Play Interoperability*.
- [12] Mannini, A. and Sabatini, A. M. (2010), Machine Learning Methods for Classifying Human Physical Activity from On-Body Accelerometers, *Sensors* **10**: 1154–1175.

<sup>7</sup> <http://graphite.ecs.soton.ac.uk/>