THE PERSONALISED 'DIGITAL ATHLETE': An evolving vision for the capture, modelling and simulation, of on-field athletic performance.

Jacqueline Alderson & William Johnson

School of Sport Science, Exercise & Health, University of Western Australia, Perth, Australia

Technological advances in the areas of three-dimensional (3D) body scanning, in-vivo imaging and novel forms of motion capture and data analytics (e.g. deep learning neural networks) are rapidly bridging the lab versus field-based nexus that has historically plagued the applied sport biomechanist. Similarly, exponential advances in hardware and computer processing power has witnessed the emergence of the personalised 'digital athlete', an overarching vision that facilitates, via the integration of multiple technologies, real-time biomechanical data collection, modelling and reporting for immediate biofeedback.

KEY WORDS: digital athlete, deep learning, real-time variable estimation

INTRODUCTION: Globally, nearly twenty three billion dollars will be spent on wearable technologies in 2016 and it is estimated that by the year 2021 this market will increase to $171 billion (Reportlinkercom, 2016). In the year 2020, 411 million wearable units will be sold in the US alone, with 70% of these being wrist-based devices such as smartwatches and fitness trackers (Paul Lamkin, 2016). Despite the marketing hype surrounding the potential scale of wearable devices transforming the sport science landscape, and the staggering amount of information that can be collected and stored using such devices, the one-dimensionality of the information provided is currently of limited downstream use to the sport biomechanist. These limitations aside, the rapid pace of technological improvement, particularly in the area of computational and neuro-musculoskeletal modelling, has witnessed successful lab-based real-time data kinematic and kinetic data capture with immediate biofeedback for some time (Crowell et al., 2010; Mullineaux et al., 2012). In recent years, researchers have successfully extended the technology to include real-time estimations of joint contact and muscle forces (Pizzolato, Lloyd, Sartori & Reggiani, 2014). With continued advances in; passive and active imaging, multi-sensor integration, advanced historical data mining, scalable real-time processing architectures, and non-linear data science analytics techniques (e.g. deep learning), it is clear that the laboratory versus field nexus that has hampered the sport biomechanists' ability to collect, model and visualise on-field/in-game data in real-time, will soon be bridged. This paper outlines an approach currently being developed by a team of biomechanists and data scientists to achieve the personalised 'digital athlete' vision.

METHODS: The personalised 'digital athlete' schema comprises two workbench phases that rely on advanced big data structure architectures. These architectures are capable of auto-classifying data and facilitate elegant, bidirectional data transfer, thereby enabling real-time biomechanical modelling and bio-feedback.

Phase 1: Creating structurally databases and training sets

A series of input modules, comprising both raw and processed/modelled data, feed into a single traditional relational database management system (RDBMS) (Figure 1). These modules may be fully integrated horizontally (i.e. a single participant may have informed all modules) or partially integrated (i.e. a single participant informs a single, or a partial subset, of modules).

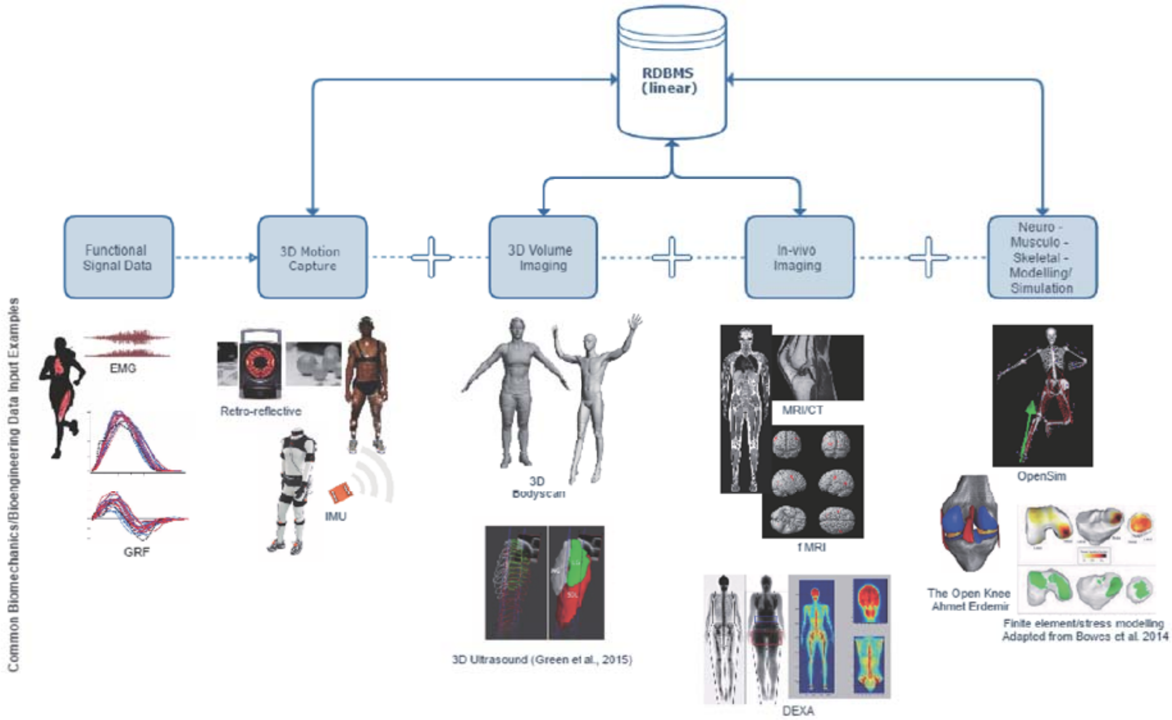

Figure 1: Independent data modules feeding a traditional linear relational database management system (RDBMS).

Within the complete personalised 'digital athlete' vision it is expected that the input data will extend beyond traditional kinematic and kinetic variables, with emerging technologies also playing a critical role in personalised component of the 'digital athlete' framework. As an example, we have developed a novel approach for estimating athlete-specific body segment inertial parameters (BSIPs), by integrating mass distribution gathered via a Dual X-Ray Absorptiometry (DXA) scanner, with body volume/shape information collected using a raste-stereographic 3D surface scanner. Whilst previous attempts to derive BSIPs from DXA and 3D scans assumed rigid body segments, our technique re-projected the DXA 2D mass data into the 3D scan while considering a deformable (non-rigid) transformation when aligning and registering the two scans. This approach resulted in more accurate BSIP estimation compared to traditional estimation techniques and was has now been adopted as standard with our lab data collction protocols (Rossi et al., 2012; El-Sallam et al., 2013).

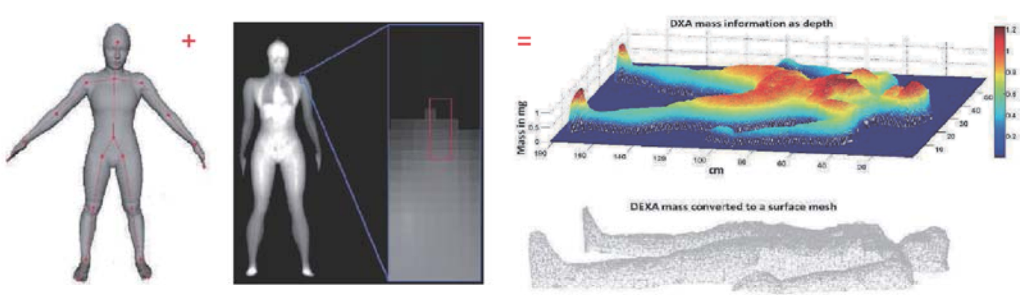

Figure 2: BSIP estimation workflow showing 3D scan merged with a DXA scan (Rossi et al., 2012; El-Sallam et al.,2013)

While many research groups and centres may have large databases of concurrently collected coordinate 3D (c3d) motion files, GRF and EMG data, very few have the capacity or funding resources to fulfil the data input criteria for all modules (see Figure 1 for examples), and this is especially true for studies involving large sample sizes. Subsequently, advanced data science techniques can be employed to mine, prepare and structure all data (legacy and newly acquired) within a single database management system that maintains the integrated relationship between datasets. Importantly, once a minimum number of samples are obtained within each module (~>25,000), advanced non-linear data science techniques can be used to estimate missing data (i.e. create a complete athlete profile across all modules when only partial data are available).

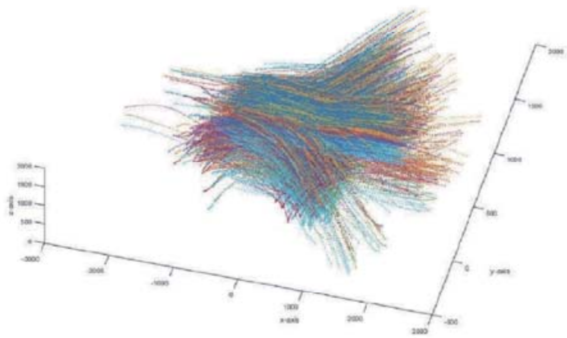

Figure 3: e.g. motion capture trajectories

Phase 2: Deep Learning (Neural Network) Schema (DNN)

This phase requires the appropriately prepared RDBMS (*Phase 1*) to provide the necessary inputs into a front-end training set required for use in a DNN schema. Figure 4 depicts a representative workbench for estimating one set of missing data - *estimation of ground reaction forces and moments (GRFMs) using only motion capture trajectories* as the input criteria. Although this is a case example of only one biomechanical variable's relationship to another, it is representative of the architecture of the entire 'digital athlete' workbench which effectively comprises multiple DNNs. Inputs required for the DNN training set are provided by the RDBMS. In this example these inputs are 1) motion capture trajectories from running, walking and sidestepping lab based trials, and their associated 2) GRFMs. The raw data are fed into the training set architecture and are required to pass through a series of decision gates and reformatting pipelines. The aligned datasets are then translated into lower dimensional space for importing into the multi-layer non-linear DNN. The created DNN can then be used to estimate GRFMs (output) from any motion capture trajectories provided (input).

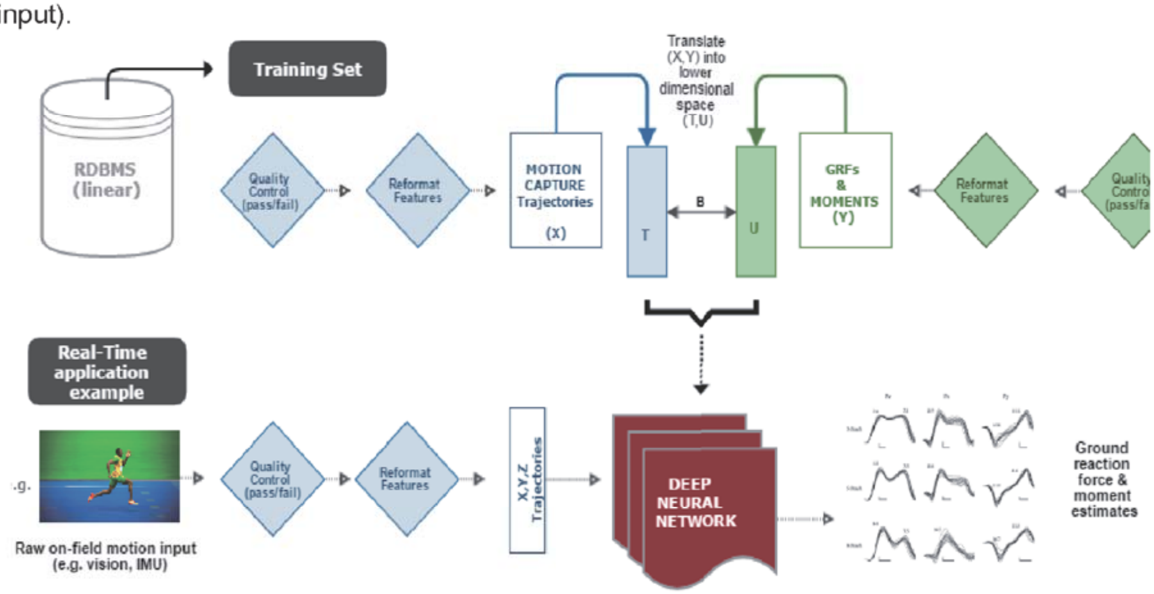

Figure 4: Schematic of a deep learning neural net training set and a real-time application example showing how it can be used to estimate unknown data.

CONCLUSION: This paper outlines a developing philosophy that compiles legacy and newly acquired data in an advanced 'big data' architectures that facilitate the creation of a personalised 'digital athlete'. This personalised digital fingerprint will be driven by subject-specificity, yet will be founded on large and extensible population datasets by which missing individual data can be estimated. The wide adoption of this type of approach to data collection and analytics will require a big picture lens adoption by the sports biomechanics community. Away from the traditional repeated experimental designs of the past, to one involving large and constantly expanding data sets, non-linear computation and a collaborative team that involves data scientists and information and communications technology specialists.

REFERENCES:

Bowes, M.A., McLure, S.W., Wolstenholme, C.B., Vincent, G.R., Williams, S. & Conaghan, P.G. (2014). Bone marrow lesions are spatially associated with denuded cartilage in osteoarthritic knees: data from the osteoarthritis initiative. *Osteoarthritis and Cartilage*; vol. 22, Supplement, Page S261.

Crowell, H.P., Milner, C.E., Hamill, J., Davis, I.S. (2010). Reducing impact loading during running with the use of real-time visual feedback. *Journal of Orthopaedic and Sports Physical therapy*;40(4),206-13.

El-Sallam, A., Bennamoun, M., Honda, K., Lyttle, A., Alderson, J. 2015, Towards a fully automatic markerless motion analysis system for the estimation of body joint kinematics with application to sport analysis. International Conference on Computer Graphics Theory and Applications, GRAPP 2015 - 10th International Conference on Computer Graphics Theory and Applications; *VISIGRAPP Proceedings*, p. 58-69.

Mourao Rossi, M., El-Sallam Abd, A., Benjanuvatra, N., Lyttle, A., Blanksby, B. & Bennamoun, M. (2012). A Novel Approach to Calculate Body Segments Inertial Parameters from DXA and 3D Scanners Data. *Proceedings of the 4th International Conference on Computational Methods (ICCM2012)*. Australia, Vol. 1, p.1-9.

Mullineaux DR, Underwood SM, Shapiro R, Hall JW. (2012). Real-time biomechanical biofeedback effects on top-level rifle shooters. *Applied Ergonomics*; 43(1),109-14.

Paul Lamkin. (2016). Forbes. Retrieved 9 June, 2016, from http://www.forbes.com/sites/paullamkin/2016/02/17/wearable-tech-market-to-be-worth-34-billion-by-2020/

Pizzolato, C, Lloyd, D. G., Sartori, M., Reggiani, M. (2014).CEINMS: an OpenSim neuromusculoskeletal toolbox to examine how different neural solutions affect muscle forces and joint moment estimates. *9th Australasian Biomechanics Conference (ABC9)*. Wollongong, NSW, Australia, 2014

Reportlinkercom. (2016). Reportlinkercom. Retrieved 9 June, 2016, from http://www.reportlinker.com/p02042684-summary/Wearable-Computing-Technologies-Applications-and-Global-Markets.html

Green, DJ., Panizzolo, F.A., Lloyd,D.G., Rubenson, J & Maiorana, A.J. (2016). Soleus Muscle as a Surrogate for Health Status in Human Heart Failure. *Exercise and Sport Sciences reviews;* 44(1),45-50.

Acknowledgements

Special thanks to UWA colleagues Prof. Ajmal Mian and Dr Cyril Jon Donnelly, PhD students Koji Honda, Marcel Rossi, and external collaborators; Prof. David Lloyd, , Dr Amar Ell-Sallam and Dr Andrew Lyttle, for their past and on-going contribution to the content of this paper.