

# Workspace for Industry 4.0

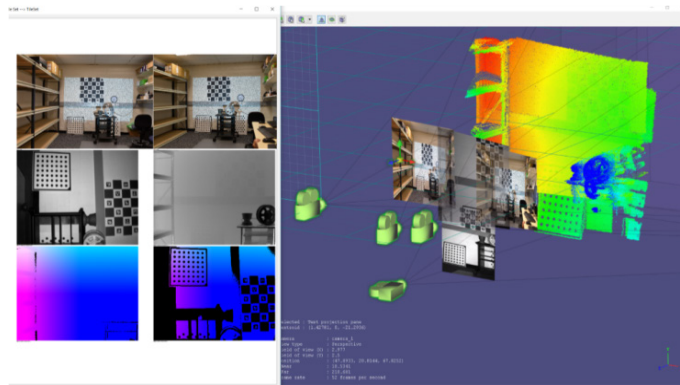
L. Hetherton, M. Bolger, Z. Zhang, N Oakes, D. Watkins and P.W. Cleary

CSIRO Data61, Australia

Email: [lachlan.hetherton@data61.csiro.au](mailto:lachlan.hetherton@data61.csiro.au)

**Abstract:** Compared to writing software from scratch, Scientific Workflow Systems (SWSs) provide a more convenient way to acquire data, customize inputs and combine algorithms in order to develop reproducible outputs. Workspace is an SWS that has been under continuous development at CSIRO since 2005 and which attempts to address key challenges faced by the research and commercial communities; researcher/developer productivity, reproducibility of results, collaboration between research teams and portability and interoperability between different computing environments and scientific domains.

One such domain is that of Industry 4.0, commonly referred to as the “fourth industrial revolution”. Industry 4.0 centers on the development of “smart factories” where cyber-physical systems and physical processes operate in concert to create a virtual copy (or “digital twin”) of the physical world in order to enable (or in many cases automate) decentralized decision making. In such smart factories, manufacturing processes often take the form of a workflow, where devices, data and algorithms are combined in different ways to produce outputs for decision making and automation purposes.



**Figure 1.** Part of the Workspace-based demonstrator - results of multi-sensor data collection. On the left, the raw image data from the sensors. On the right, DSLR images and point cloud data from time-of-flight sensors are shown in a 3D scene. Points are coloured by depth, where blue is closest and red is furthest away

In this paper, we explore how Workspace, in both design and implementation phases, can provide significant value when creating digital twins using arrays of sensors. An Industry 4.0 demonstrator application is used to highlight challenges and to show how a workflow approach can be used in applications representative of real-world needs. The demonstrator application is specifically focused around the inspection of manufactured components for quality-assurance purposes. In doing so, we will analyse how Workspace addresses the specific challenges that are presented by problems within the Industry 4.0 domain. These include the need to tailor the system to specific manufacturing processes at low cost, the ability to cope with large volumes of streaming data and changes in how the data is integrated and interpreted, developing the skills of operators to cope with the complicated processes involved while maintaining existing know-how, and the need to maintain the integrity of the disparate processes.

We provide a detailed study of those features and plugins of Workspace which are beneficial to the development of Industry 4.0 applications, such as distributed execution, OpenCV, PCL, ZeroMQ plugins, and the ability to create custom applications. These capabilities, along with the Workspace design principles of analyse, collaborate, commercialise, everywhere, align it well with problems in the Industry 4.0 space, allowing the creation of high performing implementations for use in industrial settings.

**Keywords:** *Workspace, industry 4.0, workflow, manufacturing, inspection*

## 1. INTRODUCTION

Industry 4.0 is a catch-all term for research and development in automation and data exchange technologies related to the creation of so-called “smart factories”. These make use of advancements in cyber-physical systems, robotics and autonomous systems, the Internet of Things (IoT), machine learning and artificial intelligence in order to build a digital representation of a factory’s state – a “digital twin”. This representation is used to inform and automate decision making in an efficient, decentralised manner. IDC estimates (Bigliani et al. 2019) that by 2020, 30% of the top 2,000 global companies will be using data from Digital Twins to improve product innovation success rates and organisational productivity, achieving gains of up to 25%. Industry 4.0 experts define four design principles (Hermann et al. 2016):

1. Interoperability – the ability of machines, devices, sensors and people to communicate with each other using IoT or Internet of People (IoP) technologies.
2. Information transparency – the ability to construct a digital copy (or digital twin) of the factory and its production process, accurately representing its state.
3. Technical assistance – supporting decision making by aggregating and visualising information effectively, as well as by automating tasks considered too unpleasant, exhausting or unsafe for humans.
4. Decentralised decisions – the ability of cyber-physical systems to make decisions on their own and to perform tasks as autonomously as possible.

Workspace aligns well with these principles. Firstly, its plugin architecture and performance characteristics enable interoperability between sensors, algorithms and robots. Secondly, its 2D image and 3D data processing and analysis capabilities have a track record of being used for digital twinning applications in other domains (Cohen et al. 2018; Cleary et al. 2015). Thirdly, its visualisation and custom GUI capabilities allow information to be tailored to technicians and management at appropriate levels of detail. Finally, its pluggable scheduling architecture enables data capture, data processing and decision making to be as asynchronous and decentralised as needed. Further evidence of Workspace’s alignment with the domain can be seen in the following list of specific challenges associated with developing Industry 4.0 software (Wolter et al. 2015):

- The need to combine multiple devices, algorithms and processes from multiple scientific domains and tailor them to each manufacturing process. Workspace workflows can be easily modified to adapt to the differing requirements, with operations or nested workflows substituted as needed. This reduces the cost of translating solutions between targets and maintaining them as new technologies are developed.
- Lack of adequate skills in manufacturing organisations to develop and operate the highly technical processes involved. The ability in Workspace to create custom GUI elements overlaying complicated workflow processes is critical to alleviating this challenge.
- Maintaining the integrity of production processes (e.g. critical machine-to-machine communication with very short, stable latencies): real-time where needed and asynchronous where extensive processing is required. The distributed processing capabilities and performance-conscious development approach of Workspace align well here. Workspace includes modules to assist with record-keep and provenance-tracking, to assist in tracking manufacturing quality and fault-finding in the case of system failure.
- The need to protect industrial know-how (e.g. trade secrets, device designs, control files for automation equipment). In many cases, this requires that specific devices or the system itself run in an offline environment, disconnected from outside interference. This aligns well with Workspace’s ability to run in completely disconnected environments, un-reliant on third party infrastructure.

In this paper we explore the development of a representative Industry 4.0 example which is focused on digital inspection in a parts manufacturing context. Manufacturing processes are increasingly global, both in terms of the scale of their target markets and the approach to their manufacture. As such it is critical to the value and quality of the end product that components produced at every stage are reliably manufactured to specification, often with tight tolerances. It has been shown that accurate measurements can have a 10-15% impact on production costs alone (Czichos 2011). For some engineering firms, this can equate to millions of dollars per unit (Starkie et al. 1995). This is especially evident when noting that the incorrect manufacture of a component not only necessitates the repetition of that particular production stage; depending on how far

down the line the defect is detected, entire downstream chains of production processes can be impacted at substantial cost. Consequences include direct re-work or replacement costs, delays in product completion, higher management and transport and logistics costs and in worst case scenarios where defects are not detected with system failure. In many cases the largest commercial impact is on the time to market for more customised or personalised items. Workspace includes plugins for the management of failure accountability and provenance, so it can provide solutions to help to keep track of the part-manufacturing process.

Many manufacturing inspection processes are either manual or semi-manual, with the standards in many industries often being strict and enforced by regulatory agencies. Some are made more complicated in that not all components can be easily measured due to their scale, location or that the environments involved are dangerous to humans. There are a range of approaches one could take to automating parts inspection depending upon the object being inspected and the criteria of the inspection: from simple 2D image analysis based on the output of optical sensors, through to the analysis of high resolution 3D laser scans or computed tomography (CT) volume scans. We explore how a stereo-photogrammetry based approach can be developed, using Workspace to integrate a variety of different sensors with point-cloud reconstruction algorithms to enable robust construction of a digital twin of a target object for the purpose of identifying manufacturing errors, informing decision making and/or automatically triggering corrective processes.

## 2. THE INSPECTION PROBLEM, CHALLENGES, REQUIREMENTS AND SOLUTION

The representative application that we consider here is one that generates a digital twin of a stationary object (part) for the purpose of determining the extent to which it deviates from manufacturing specification. This process can be performed by comparing 3D scans of the “as manufactured” part with the original computer aided design (CAD) model of the “as designed” part. One way of acquiring such a 3D scan can be using a dedicated 3D scanner in which case the scanner’s software API is exposed directly for use in Workspace. A significant drawback of this approach is that such scanners often need to be heavily-customised to the specifications of the object under measurement, which is not always possible due to the scale of the object, restrictive tolerance requirements and occlusion from objects in the surrounding environment (e.g. if it needs to be measured at-speed on a production line).

This workflow solution developed here attempts to overcome these limitations with an alternative approach: using a Workspace workflow to automate a stereo-photogrammetry process for constructing 3D point clouds from multiple stereo image pairs. This has the additional advantage of being a general process that is then applicable to a wide range of such applications. The workflow simultaneously captures data streamed from multiple sensors of different types placed in fixed positions, integrates the captured views and generates point-cloud datasets from them. There is no Workspace-inherent limitation to the number of sensors used. The point clouds are filtered and cleaned, then merged into a single, combined dataset and transformed into the coordinate system of the reference CAD model.

The sensor configuration (the types of sensors required, their locations and the underlying network topology) and the algorithm(s) used to fuse the data vary greatly depending on a number of factors:

- The physical properties of the object being inspected, such as its size and the material of which it is comprised.
- The conditions under which the object is being inspected, such as environmental lighting, or whether or not the object is in motion. In the latter case, high frequency data capture is then required. For simplicity, the object considered in this example is static.
- The accuracy requirements of the inspection, which are based largely on the manufacturing tolerances of the manufacturing processes involved in its creation.

Workspace allows us to create a workflow that incorporates these sensors and algorithms as interchangeable components with easy ability to adapt or evolve the processing methodology. In this way the workflow can be easily reconfigured to cater for changes in sensor configuration or the processing regimen at any point during or after the system’s development, thus avoiding the need to create a bespoke application for each specific usage scenario and reducing ongoing maintenance costs.

Two types of consumer-grade sensors are used here: consumer-grade digital single-lens reflex (DSLR optical) cameras and Microsoft Kinect time-of-flight sensors. Each has distinct advantages and disadvantages. The DSLR cameras are capable of capturing high-resolution imagery with a high degree of transverse accuracy. When combined in a stereo configuration, photogrammetry algorithms can be used to obtain depth information, though this is much less accurate. Conversely, time-of-flight sensors provide a lower degree of transverse accuracy but provide more accurate depth information and are capable of

capturing data at a higher frequency than by optical sensors. Combining the two sensor types in a fusion model allows limitations of each individual sensor type to be overcome.

A general cyber-physical solution needs to be capable of capturing data from a large, configurable streaming sensor network, at high levels of accuracy. The specific object under inspection in this demonstrator is an alloy wheel manufactured by T-Mag. The wheel is made of a magnesium alloy and is approximately 30 cm in diameter. Due to its relatively small size, only a small number of sensors are required to inspect it.

Data capture from each sensor needs to be triggered at approximately the same time. This requires minimising latency in the sensor network and account to be taken of the different triggering speeds of the different sensors. For example, data capture from an optical sensor is slower than from a time-of-flight sensor. Optical sensor images are of order of hundreds of megabytes, so large numbers of sensors generate significant data volumes that requires network bandwidth of order of Gb/s.

Stereo-photogrammetry 3D reconstruction usually requires high-quality sensor calibration in order to produce accurate results. Calibration data is both intrinsic (such as camera focal length) and extrinsic (such as its position in 3D space). These attributes allow construction of a digital twin of the full inspection environment including the sensors. This uses captured images of calibration boards in selected poses. To be operationally useful this process needs to be automated, streamlined and repeatable. Challenges include:

1. Experimentation being needed to capture the poses that minimise re-projection error in the extrinsic data. This can require extensive trial-and-error, user feedback or the capture of many poses.
2. The sensor space needs to be entirely “connected”, i.e. the graph that defines the sensor/pose visibility space needs to be connected which requires careful pose selection and knowledge of prior poses.
3. Scene disruption where sensors are perturbed by environmental changes, device wear-and-tear or human error.

A further consideration is the security of the sensor network. There are three clear reasons for this: firstly, captured data can contain images of commercially sensitive objects or designs; secondly, sensors may be operating in a space with operational personnel involving privacy protection, often enforced by workplace agreements; and thirdly, sensor data captures and other automated processes can be triggered over the network, opening up the possibility of attacks by malicious external agents. In this example, the sensor network operates on a virtual private network (VPN) accessible only to authorized staff.

### **3. WORKSPACE WORKFLOWS FOR DIGITAL PARTS INSPECTION**

#### **3.1. The high-level workflow solution**

At a high-level, the solution can be thought of as a workflow comprised of five distinct stages: 1) Calibration, 2) Data capture, 3) Reconstruction of the “as manufactured” dataset, 4) Comparison against the “as designed” dataset, and 5) Evaluation and comparison for meaningful errors.

##### ***Calibration***

We implemented a repeatable calibration process that scaled to the number and types of sensors in the network. Workspace makes it easy to scale operations, and quickly reconfigure existing workflows to new equipment types. Calibration is performed by a Workspace workflow that coordinates image capture from all sensors for the sequence of calibration board poses. It then uses this set of images to generate calibration model data. Both raw and calibrated data is stored in a database so that it can be easily accessed in the reconstruction phase. The workflow has a custom GUI which provides a simplified interface for use by operators conducting the calibration process.

This workflow leverages previously developed Workspace plugins, taking advantage of the high level of capability re-use that Workspace encourages:

- An OpenCV (Bradski 2000) plugin (CSIRO, OpenCV), which exposes critical computer vision algorithms and underlying mathematical functions to Workspace, such as image thresholding, brightness adjustment, matrix mathematics and feature identification.
- A database plugin which is used to store a representation of the system setup in a database, including sensor settings and calibration data. Housing this information in a non-relational database enables the solution to evolve in line with the sensor network configuration and data capture requirements.

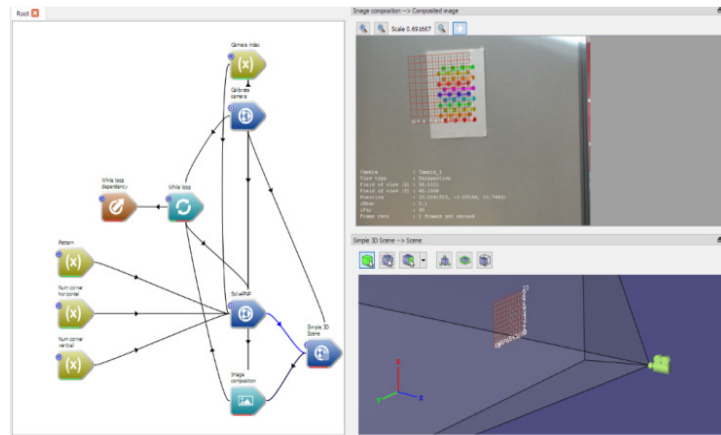


Figure 2. Workspace workflow for capturing calibration board poses. Inline visualisation widgets show pose estimation data.

**Data capture**

Streaming sensor data is captured with sensor trigger timing and data collection orchestrated by an overarching workflow. For continuous capture devices (such as time-of-flight sensors), data capture simply involves retrieving images from the device’s data stream, but for higher-resolution devices used here we required a trigger step to guarantee timing, followed by an explicit retrieval step.

To ensure that this process is capable of scaling to large sensor network sizes, the software solution makes use of Workspace’s distributed execution capabilities to execute data capture workflows on each computer to which sensors are attached. These run continuously, capturing and transmitting data at the request of the higher-level orchestration workflow running on a server. The orchestration workflow’s structure can differ depending upon the data capture requirements of the specific application for which it is intended. For example, in the case of the calibration orchestration workflow, calibration model computation algorithms are executed immediately after data is streamed back to the server from all sensors.

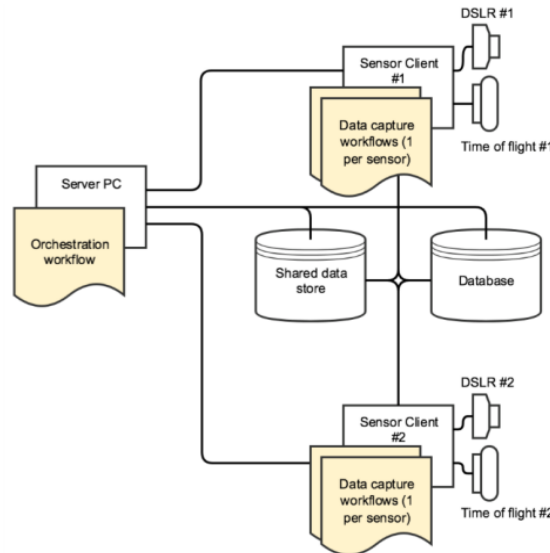
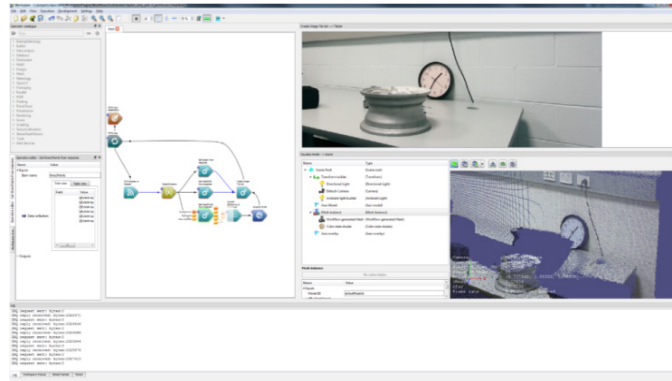


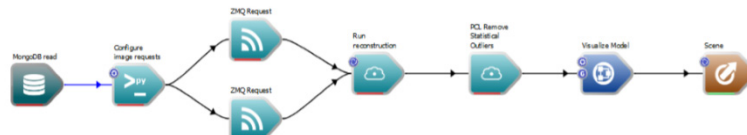
Figure 3. High-level system configuration of the overarching control workflow. Two data capture clients are connected to a server. Each client is connected to a DSLR and time-of-flight sensor. All computers have access to a shared data store (for output storage) and a database containing the network configuration.

Both the data capture and orchestration workflows depend on a number of specific Workspace plugins: 1) A Distributed Execution plugin, which ships out-of-the-box with Workspace. This is used to remotely schedule the data collection sub-workflows on the computers to which sensors are connected. 2) A ZeroMQ (IMatix 2014) plugin, which exposes request/reply functionality to Workspace in the form of specialised loop operations. These allow an executing workflow to be triggered in response to data being received from the message queue. This mechanism forms the communications backbone through which all data is piped and 3)

a plugin for each sensor type, many of which have their own proprietary Software Development Kits (SDKs), in this case for the MS Kinect and the Nikon DSLRs. Each plugin provides an operation that acquires data from a specific sensor either directly connected to the machine, or available on the same network.



**Figure 4.** Example data capture workflow executing on a client computer. On the right-hand-side, the wheel can be seen in both the raw image output and point-cloud output.



**Figure 5.** Example reconstruction workflow, showing an outliers removal step and reconstruction step.

***Reconstruction of the “as manufactured” dataset***

The process for the conversion of a captured multi-sensor dataset into a point cloud is performed within the orchestration workflow. The algorithm required will depend on the nature of the inspection task and the accuracy requirements of the manufacturing standard. Possible techniques include Structure from Motion (Mead et al. 2015) and the open-source 3D Underworld structured-light scanning system (Herakleous et al. 2014). Each of these techniques possess different strengths and weaknesses, suffice to say that a configuration that utilises the strengths of multiple sensor types is able to overcome the drawbacks of any one sensor type in producing a fused dataset (Haghighat et al. 2011). A significant advantage of using a workflow-driven approach is that these algorithms and sub-workflows can be easily replaced or substituted in order to adapt this generic workflow solution to the specific nature of each inspection task.

This part of the solution makes heavy use of the previously described OpenCV plugin to perform stereo-photogrammetry, as well as the Mesh and Point Cloud Library (Rusu et al. 2011) (PCL) plugins (CSIRO PCL), which each expose a variety of surface mesh and point cloud processing algorithms such as normal estimation, moving least squares, iterative closest point and organised surface reconstruction. These algorithms are used to clean and process the output point clouds.

***Comparison against the “as designed” dataset***

To perform the comparison, the reference CAD mesh needs to be read by the orchestration workflow and the fused point cloud dataset has to be aligned with its the coordinate system. Once aligned, an algorithm is used to compute an error value for each scanned point to indicate how close it is to the surface of the CAD mesh. The resulting error values can be visualised by colouring the reference CAD model using Workspace’s OpenGL shading capabilities that are supplied in the Rendering plugin. This plugin contains hierarchical point-cloud subdivision and level-of-detail algorithms to enable interactive visualisation of the very dense and high-resolution point clouds that make up the datasets, as well as surface geometry added to contextualize the output.

***Evaluation of the comparison data for meaningful errors***

A process is required to evaluate the comparison to determine meaningful errors and take appropriate corrective action. This can be done using error metrics such as L2 or L∞ norms with a detailed understanding of the object being inspected and its manufacturing tolerances. In a broader Industry 4.0 context, meaningful errors such as these could automatically be fed downstream into a situational awareness model of the larger system of which this manufacturing process is a part. The outputs of such a model could then be used to

inform higher-level decision making. Again the use of a workflow separates the processing choices from the software structure allowing this part of the process to be refined, adapted or customised at any point.

#### 4. CONCLUSIONS

A workflow-driven approach to developing Industry 4.0 applications is shown to be capable of alleviating many of the challenges presented by this domain. Workspace workflows are easily reconfigurable, elements can be added, interchanged, reconfigured or reapplied. Through the exploration of a Workspace-based solution for manufacturing parts inspection, key advantages were identified in relation to the Industry 4.0 guiding principles of interoperability, information transparency, technical assistance and decentralised decisions. The ability to reuse devices, algorithms and sub-workflows in different configurations to address different Industry 4.0 problems is certainly a significant advantage over traditional software development methodologies, as is Workspace's ability to abstract complicated workflows for less technical operators.

The demonstration case for which the Workspace solution explored involves the detailed, in-situ digital inspection of a manufactured component - being a 30 cm alloy wheel - and its comparison against the manufacturing specification. As a consequence of the generic workflow base and the ability to easily adapt/customise arising from a workflow approach, this solution is substantially customisable across the entire class of inspection problems ranging in size from mm's to 10's m's.

#### REFERENCES

- Bigliani, R., Carter, P., Dunbrack, L., Findling, S., Guzman, O., Hand, L., Hernandez, D., Jewell, J., Jyoti, R., Little, G., Ng, S., Rizza, M. N., O'Brien, A., Olvet, T., Parker, R., Prouty, K., Pucciarelli, J.C., Rajan, R., Seminara, J.P., van Vonno, J., Villate, R., Walker, M., Yates, M., Yorifuji, Y., Fitzgerald, S., (2018), IDC FutureScape: Worldwide Digital Transformation 2019 Predictions, Worldwide Digital Transformation 2019 Predictions
- Bradski, G. (2000) The OpenCV Library," *Dr. Dobb's J. Softw. Tools*.
- CSIRO, Workspace OpenCV Plugin.[Online]. Available: <https://github.com/csiro-workspace/opencvplugin>.
- CSIRO, Workspace PCL Plugin.[Online]. Available:<https://github.com/csiro-workspace/pointcloudplugin>.
- Czichos, H., Saito, T. and Smith, L. E. (2011) *Springer Handbook of Metrology and Testing*. Springer
- Deelman, E., Gannon, D. Shields, M., and Taylor, I., "Workflows and e-Science: An overview of workflow system features and capabilities," *Futur. Gener. Comput. Syst.*, vol. 25, no. 5, pp. 528–540, 2009.
- Gil, Y, Deelman, E., Ellisman, M., Fahringer, T., Fox, G., Gannon, D., Goble, C., Livny, M., Moreau, L., Myers, J. (2007) Examining the Challenges of Scientific Workflows," *IEEE Comput.*, vol. 40, no. 12, pp. 24–32
- Haghighat, M. B. A., Aghagolzadeh, A. and Seyedarabi, H. (2011) Multifocus image fusion for visual sensor networks in DCT domain," *Comput. Electr. Eng.*, vol. 37, no. 5, 789–797,
- Herakleous, K. and Poullis, C. (2014) 3DUNDERWORLD-SLS: An OpenSource Structured-Light Scanning System for Rapid Geometry Acquisition, *CoRR*, vol. abs/1406.6595.
- Hermann, M., Pentek, T., and Otto, B., "Design Principles for Industrie 4.0 Scenarios," in *49th Hawaii Int. Conf. Syst. Sci.*, 2016, pp. 3928–3937.
- IMatix, (2014) ZeroMQ,. [Online]. Available: <http://zeromq.org/>.
- Industry 4.0. [Online]. Available: <https://bit.ly/1PFdfwu>.
- Mead, S. R., Prakash, M., Magill, Bolger, C. M. and Thouret, J.-C. (2015) A Distributed Computing Workflow for Modelling Environmental Flows in Complex Terrain," in *Int. Symp. Environmental Software Syst.*, pp. 321–322.
- Rusu, R. B. and Cousins, S. (2011) 3D is here: Point Cloud Library (PCL)," in *IEEE Int. Conf. Robotics and Automation*,.
- Watkins, D., Thomas, D., Hetherington, L., Bolger, M., Cleary, P.W., (2017), Workspace- a Scientific Workflow System for Enabling Research Impact, *22<sup>nd</sup> International Congress on Modelling and Simulation*
- Wolter, M. I. *et al.*, (2015) Industrie 4.0 und die Folgen für Arbeitsmarkt und Wirtschaft,