

# Using the Cloud as a Platform for Evaluation and Data Preparation

Ivan Eggel, Roger Schaer and Henning Müller

**Abstract** This chapter gives a brief overview of the VISCERAL Registration System<sup>1</sup> that was used during all VISCERAL Benchmarks and released as open source on GitHub. The system can be accessed both by participants as well as administrators, reducing direct participant–organiser interaction and handling the documentation available for each of the benchmarks organized by VISCERAL. Also the upload of the VISCERAL usage and participation agreements is integrated as well as attribution of virtual machines that allow to participate in the VISCERAL challenges. In the second part a summary of the various steps in the continuous evaluation chain, mainly consisting of the submission, algorithm execution and storage as well as the evaluation of results is given. The final part consists of the cloud infrastructure details, describing the process of defining requirements, selecting a cloud solution provider, setting up the infrastructure and running benchmarks. The chapter concludes with a short experience report outlining encountered challenges and lessons learned.

## 1 Introduction

Over the past few years medical imaging data has been steadily growing at a fast pace. In 2013 for instance, the Geneva University hospitals produced around 300'000 images per day on average [8]. Working with increasingly big amounts of data has become difficult for researchers as the download of such big data would

---

Ivan Eggel, Roger Schaer and Henning Müller  
Institute for Information Systems, University of Applied Sciences Western Switzerland (HES–SO Valais), Sierre, Switzerland e-mail: [ivan.eggel@hevs.ch](mailto:ivan.eggel@hevs.ch), [roger.schaer@hevs.ch](mailto:roger.schaer@hevs.ch)

Henning Müller  
University Hospitals and University of Geneva, Geneva, Switzerland e-mail: [henning.mueller@hevs.ch](mailto:henning.mueller@hevs.ch)

<sup>1</sup> <http://visceral.eu:8080/register/Login.xhtml>

require a significant amount of time, especially in countries with slow internet connections.

In the context of the VISCERAL competitions where big data need to be shared with the participants, we decided to make use of a cloud infrastructure to host the data as well as to run the participants' code. On the one hand this prevents the download of the data, on the other hand participants are provided with equal-powered virtual machines in the cloud to run their code on, which makes the algorithms highly comparable in terms of performance. The evaluation infrastructure allows the competitions to be carried out efficiently and effectively, along with a continuous evaluation to take place beyond the competitions. In order to register and administer the participants but also to provide an interface between the participants and the cloud infrastructure, the VISCERAL Registration System has been developed.

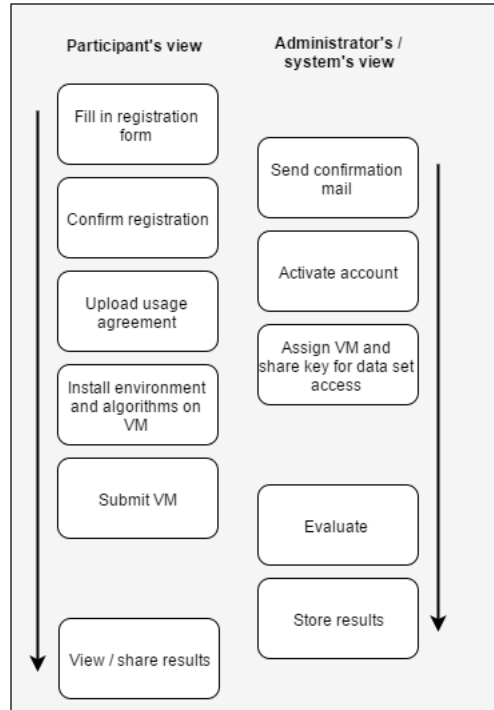
## 2 VISCERAL Registration System

The VISCERAL project [6] has as a main goal to create an evaluation infrastructure for medical imaging tasks such as segmentation [7], lesion detection and retrieval [5]. An important part of the project was to create an innovative infrastructure for evaluating research algorithms on large image data sets, and thus bringing the algorithms to the data instead of the data to the algorithms [3]. This is necessary when data grow large and image data have been identified as one of the main areas of large data sets [1].

In order for participants to have access to the cloud infrastructure provided by VISCERAL, participants have to register in the VISCERAL Registration System. This system's purpose however is not restricted to registration of participants but also has the role of participant management system and additionally provides an interface between the participant and the cloud infrastructure, which hosts virtual machines and storage for the data sets. Figure 1 offers a simplified overview of the system for all steps needed from the registration process until the ability to view the participant's results. The approach of using such an integrated system for running competitions is highly recommended, significantly reducing administrative overhead regarding organiser-participant interaction as well as manual cloud configuration by the organiser, particularly if there is a large number of registering research groups. Such a system can also be used for continuous evaluation, allowing participants to evaluate their algorithms at any time beyond the competitions.

### 2.1 Registration

In the registration form, participants are asked to fill in their contact information including affiliation and the benchmark in which they would like to participate. After receiving an email the participant needs to confirm their registration in order to



**Fig. 1** Registration and subsequent processes from participant's and administrator's / system's point of view.

obtain access to their personal dashboard. From there the VISCERAL usage agreement needs to be downloaded, printed and signed. An upload function allows for an upload of a scanned copy of the usage agreement which, upon approval by the organiser, grants access to the VISCERAL data set and the login credentials for a virtual machine (VM) in the cloud.

## 2.2 Participant Dashboard

After a successful registration and verification process the participants are given an extended view on their dashboard as shown in Figure 2, mainly providing:

- **Access details for VM and data set** A VM, depending on the OS platform, is accessed with a specific protocol (SSH for Linux, RemoteDesktop for Windows) and the credentials. In order for the participants to access the data set (read-only) a specific data key is provided.
- **Start/stop VM** Starting/Stopping a VM from the dashboard was implemented due to the fact that running a VM in the cloud causes financial costs, especially if it is never turned off during an extended period of time. Like this, participants who are not executing code are able to turn off their machines without requiring

a direct access to the cloud management system. It needs to be mentioned that during the first benchmark, several participants left their VMs active over many weeks without executing code, resulting in unnecessary costs. In order to partially resolve this problem, an automatic shut down of all VMs was introduced, scheduled every Friday evening, unless a participant excludes their VM from this shut down using the option in the dashboard.

- **Download of benchmark files** Benchmark files are files that provide additional information on a specific benchmark. This can represent information such as URLs of files in the data set that can be accessed from a VM, cloud usage guidelines or a data handling tutorial. The goal of these files is to give useful and clear information to the participant on how to use the system, the cloud and the data set, significantly reducing email exchange between the participant and organiser by preventing simple recurring questions.
- **Submit VM** After the installation of necessary libraries and algorithms inside the provided VM and a testing phase, the participant can submit their VM from their dashboard in order for the algorithm to be evaluated for its performance. Exact instructions on how to submit a VM and on what exactly must be provided in the VM, are provided in the form of a benchmark file.
- **View results** As soon as the evaluation has completed the participant is able to view the results in the dashboard by modality, body region, organ and configuration. Results explicitly granted to be published by participants are shown in the publicly visible LeaderBoard.

### *2.3 Management of Participants*

System administrators have access to the administration dashboard (Figure 3) that displays all registered users relative to a selected benchmark. In order to facilitate participant management, different colours highlight the participant's status. A grey background is used to indicate that a participant has registered but has not yet uploaded the VISCERAL usage agreement. A blue background suggests that a participant is waiting for administrator verification and account activation after the upload of the VISCERAL usage agreement. A yellow background is shown upon activation of the participant account, meaning that the participant is ready to be assigned a VM, whereas a green background indicates that all previous steps have been successfully carried out. It is also possible for an administrator to create new benchmarks as well as to manage existing ones (Benchmark Manager), e.g. by editing starting and ending dates. In order to administer files with additional information for each benchmark (benchmark files, Section 2.2) the File Manager is used. Besides that, administrators are also able to access and edit the information for the VM of each participant by consulting the VM Manager. Various tasks relative to the management of VMs, such as starting/stopping a VM and monitoring the current status of all VMs is done in this place. The LeaderBoard Manager is used for viewing/editing

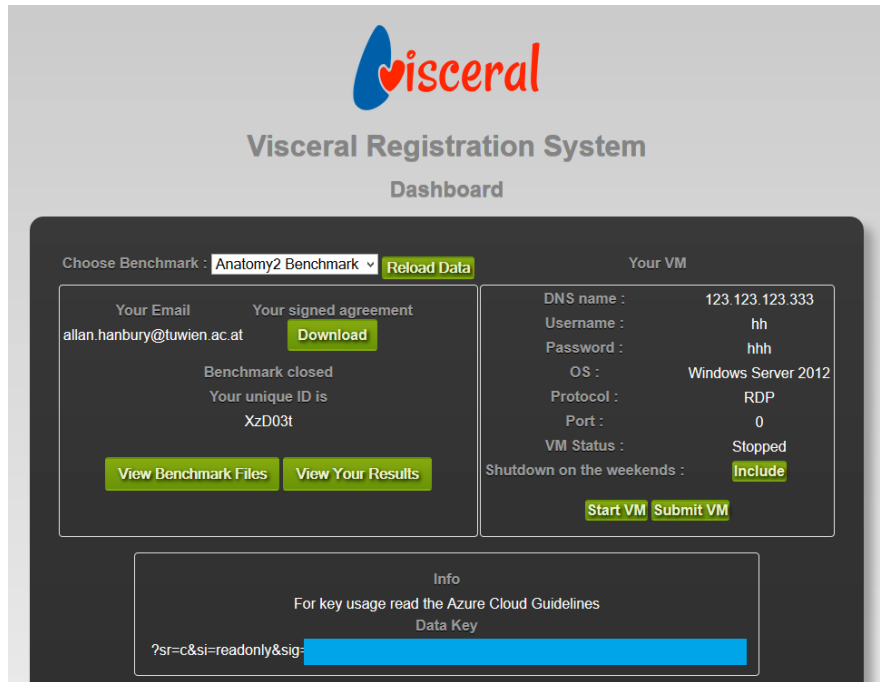


Fig. 2 VISCERAL Registration System participant dashboard

results for a specific organ that participants explicitly made available for the public (as described in Section 2.2).

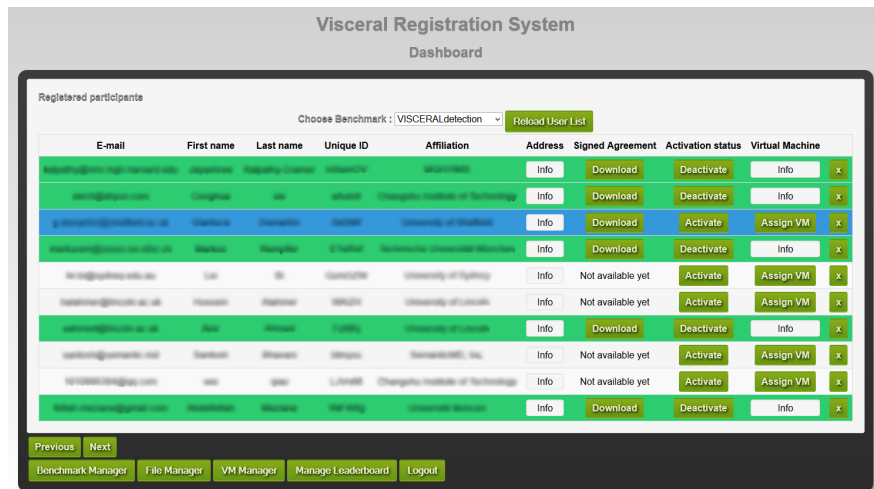


Fig. 3 VISCERAL Registration System administrator dashboard

## 2.4 Open Source Software Release

The Registration system was built with the JAVA EE<sup>2</sup> platform and Git<sup>3</sup> was used for the software management. On Github the project source is publicly available<sup>4</sup> under GNU General Public License for anyone to review and extend as they wish. Committing changes on the original codebase is not possible and requires the relevant privileges to be given. The aim in writing this code was to demonstrate the concept of cloud-based evaluation through having a working registration and administration system for the benchmarks. Due to this being the first version of the registration system that interacts so closely with the Microsoft Azure cloud, the code is only scarcely documented and contains many workarounds and solutions that should be improved in the future. The code is therefore not well suited for easy installation, nevertheless it has been made available so that the work in the VISCERAL project remains available for further development beyond the project.

## 3 Continuous Evaluation in the Cloud

This section mainly deals with the internals of the system interacting with the cloud after the participant has pressed the `Submit VM` button in the VISCERAL Registration System's participant dashboard (Section 2.2) A brief explanation of the different steps in the partly automated approach for the evaluation of segmentations on the test set generated by software submitted by participants is given. The high level of automation permits participants to submit their software multiple times to obtain results during a benchmark.

### 3.1 Submission

Before submitting a VM, the participant is asked to provide an executable in a specific directory, which takes a set of parameters defined by the organiser. The participant has to make sure that the executable properly calls their algorithms and is able to work with data in the cloud. In order to do so, participants have to accurately follow the instructions provided in the benchmark files. Clearer instructions generally mean that fewer problems occur when running the executable during the evaluation, resulting in less administrative overhead on the organisers' side.

---

<sup>2</sup> Java Platform, Enterprise Edition : <http://www.oracle.com/technetwork/java/javasee/overview/index.html>

<sup>3</sup> <https://git-scm.com/>

<sup>4</sup> <https://github.com/Visceral-Project/registration-system>

### ***3.2 Isolation of the VM***

In order to prevent participants from accessing and manipulating the VM after the submission, i.e. during the evaluation phase, a webservice is called from the VIS-CERAL Registration System as soon as a participant submits the VM. This webservice isolates the VM by creating a firewall rule in the cloud, blocking all remote access from outside the cloud. A second rule is created to explicitly allow certain ranges of IP addresses for the organisers. These rules are removed after the evaluation phase has terminated.

### ***3.3 Initial Test***

Letting participants run their own code on a VM can be error prone, as the first organized benchmark has proven. Submitted code often contains bugs or unhandled exceptions that make the evaluation fail. In order to prevent such situations in a limited way, the system tests the participant's executable prior to the final evaluation. For this test both a batch-script as well as a List of URLs of the test set files are downloaded to the VM. The script calls the participant's executable for a single test volume and ensures the match between output files and those expected by the participant. In case the test fails the VM is automatically shut down and returned to the participant in order to fix the faults present in their code.

### ***3.4 Executing Algorithms and Saving the Results***

After the initial test, the batch-script is called in order to execute the participant's executable for every volume contained in the test set as well as for each of the allowed configurations. A temporary drive in the VM is used in order to store the output files. The batch-scripts require to provide the test set URL list, the output directory, the participant ID and the benchmark as arguments.

In order to make results public and persistent, after the generation of each output file they are automatically uploaded to the cloud storage account and removed from the VM's temporary drive in order to ensure sufficient storage space for subsequent files. The process of storing the the output file to the cloud storage is performed with a secure webservice (HTTPS) connecting to the cloud provider's API. The files are stored in a folder dedicated to the participants' results inside the storage container.

### 3.5 *Evaluation of Results*

The results are evaluated using the EvaluateSegmentation<sup>5</sup> [2] software developed during the VISCERAL project. As soon as the output files are generated and stored in the cloud storage as described in Section 3.4, a script is called in order to evaluate and save the results in two steps:

- For each output file (segmentation), the EvaluateSegmentation tool is called in order to compare the segmentation with its corresponding ground truth. This results in an XML file with 20 evaluation metrics including the four metrics considered in this benchmark, i.e. the Dice, the adjusted Rand index, the interclass correlation, and the average distance.
- After this, the XML file is parsed and the metrics are inserted into a database in which each data set contains all information corresponding to a single metric value, e.g. metric id, participant id, volume, modality, organ, etc. This data is then displayed to the participant in the result dashboard or optionally in the LeaderBoard (Section 2.2).

## 4 Cloud-based Evaluation Infrastructure

This section details the technical and administrative aspects of setting up a cloud-based evaluation infrastructure, such as analysing the requirements, choosing a cloud provider, estimating costs, etc. The basic concept consists of storing large amounts of data in the cloud and providing participants in benchmarks with virtual machines (VMs) where they can access this data, install software and testing their algorithms for a given task (illustrated in Figure 4).

### 4.1 *Setting up a Cloud Environment*

Selecting and configuring a Cloud Environment requires the analysis of several points, which are detailed in this section. The analysis of requirements, as well as the evaluation of costs and logistical aspects are investigated.

#### 4.1.1 Requirements

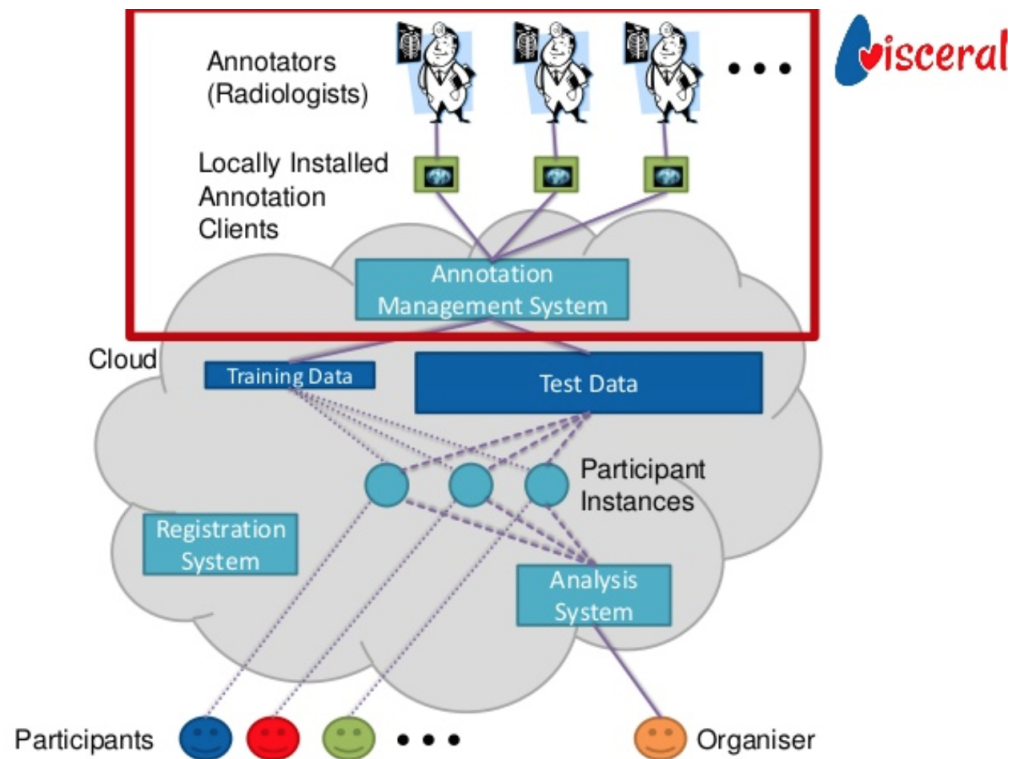
Cloud-based solution providers offer many products, including:

- data storage, both structured (database) and unstructured (files);
- computation with virtual machines;

---

<sup>5</sup> <https://github.com/codalab/EvaluateSegmentation>





**Fig. 4** Overview of the VISCERAL Cloud Framework. In the upper part (red rectangle), the process of data creation is described. Radiologists manually annotate images on locally installed clients and then submit their data to the annotation management system. From there, the training and testing sets are generated. Subsequently, participants who have registered and obtained a Virtual Machine can access their instance and optimize their algorithms and methods on the training data. Finally, the virtual machine is submitted by the participants and control is given to the organiser, who can then run the participant's executable on the testing set and perform the evaluation of the results while the participant has no access anymore.

- authentication and security mechanisms;
- application specific features:
  - distributed computing (e.g. Hadoop<sup>6</sup>);
  - high-performance computing
  - media services (e.g. video transcoding);
  - monitoring tools;
  - content caching;

The first step in setting up an environment is to determine which features are needed and to compare the availability, pricing and usage modalities of these features with

<sup>6</sup> <http://hadoop.apache.org>

different cloud providers. Another important step is to determine if there are any restrictions concerning the region of the world in which the data and services are hosted. Sample questions include:

- Can the data be hosted anywhere in the world, or only within a specific region (USA, Europe, ...)?
- If there is a region restriction, are all the required services available in this region?
- What are the costs of moving data between different regions?

Carefully reviewing the usage modalities of various cloud providers is an important step that can potentially impact the ease with which the infrastructure can be put into place. Once the required features are identified and a suitable provider is selected, the next step is planning the setup of the environment.

#### 4.1.2 Costs and Logistics

When planning the setup of a cloud environment, it is important to evaluate needs in terms of required resources, both to have a clear idea of the administrative workload (managing virtual machines, storage containers, access rights, etc.) and to estimate the costs of maintaining the infrastructure. All major cloud providers have cost calculating tools, making it easier to make an accurate approximation of monthly costs. Depending on the provider, different components can add to the total cost:

- Storage
  - Data stored (usually billed as Gigabytes per month)
  - Incoming / outgoing data traffic (usually billed per Gigabyte, incoming traffic is typically free)
  - Storage requests (PUT/COPY/POST/LIST/GET HTTP requests)
- Virtual Machines
  - Running virtual machines (usually billed by the hour)
  - Virtual machine attached storage
  - Data transfer to and from the virtual machines
  - Additional IP addresses

The costs also depends on the usage scenarios:

- Are data stored only for short periods and then removed, or do they need to be available for months or years?
- Are virtual machines required to be running 24/7 or are they used periodically for heavy computation and then turned off?
- Are Windows Virtual Machines required? (they are generally more expensive than Linux-based instances because of licensing costs)

Making cost projections for several months or a year can help in managing the resources more efficiently and making adjustments before the costs exceed expectations. Another aspect of the planning phase is to think about the resource man-

agement tasks involved. Any manual tasks can quickly become daunting when they need to be performed on a multitude of Virtual Machines. Properly configuring the base images used for future virtual machine instances can save much time and help in avoiding technical problems. Initial configuration tasks include:

- setting sensible values for password expiration and complexity requirements;
- disable unscheduled reboots on automatic update installation;
- configure the systems firewall if any ports need to be accessible from the outside;

## ***4.2 Setting up a Benchmark in the Cloud***

Once the cloud provider is selected and the infrastructure requirements are defined, a workflow for an evaluation benchmark needs to be created. This workflow includes at least the following elements:

- description of the different phases of the benchmark;
  - examples: data set creation, training phase, testing phase;
  - define what should happen in each phase and who is responsible for which task;
- required security measures;
  - geographic location of the data and infrastructure;
  - access control for participants and administrators: time restrictions for accessing the data, user rights, ...;
  - create security protocols: firewall software, antivirus, end-user agreement;
- creation of the required resources for the various phases;
  - storage containers for the data;
    - different containers for the phases (training, testing) are recommended. It makes locating and data management easier;
  - Virtual machines for computation;
    - creation of preconfigured machine templates (images) is recommended; it allows avoiding additional manual configuration on each machine after creation;
    - the variety of operating systems provided to the participants impacts the administrative workload involved in setting up the infrastructure; managing both Linux and Windows instances can make administrative tasks and automation more difficult, requiring at least two variants of all used scripts or tools;
- definition of data exchange protocols between the participants and the cloud infrastructure;
  - how can participants upload / download data to and from the cloud;

- are there additional data needed for the benchmark located outside the cloud (registration system, documentation...)?

### ***4.3 Cloud Setup for the VISCERAL Benchmarks***

The VISCERAL project was hosted in the Microsoft Azure cloud. The usage of a public cloud platform such as Microsoft Azure enabled virtually unlimited scalability, both in terms of storage space as well as computation power. The Microsoft Azure platform provides a framework for the creation and management of virtual machines and data storage containers, among a large offer of services. The platform's web management portal was used for the VISCERAL project to simplify the administrative tasks. A large amount of documentation and tools used for the different administrative tasks and technical aspects of the project are described on the Microsoft Azure website. Provision and management of VMs, as well as data storage, were the main cloud services used during the project. In the following paragraphs a brief description of these services is given.

#### **4.3.1 Storing Data Sets**

Initially, the full data set with both the medical data and additional annotations created by expert radiologists was uploaded to a cloud storage container. Other cloud storage containers were then created in each benchmark to store the training and test data sets, participant output files and evaluations. Time-restricted read-only access keys were distributed securely to the participants for accessing the training data sets. Participants had no access to the test set and subsequent evaluation results. Over the course of the project, new images and their annotations were added to the storage containers when required.

#### **4.3.2 Participant VMs**

In order to run the VISCERAL benchmarks, the participants needed access to the stored data and computing instances to execute their algorithms. Virtual machines running on the Microsoft Azure cloud infrastructure were pre-configured to run these tasks. Different templates were configured for 5 operating systems including both Windows and Linux. A virtual machine was provided to each participant, allowing them to access the training data set and upload their algorithms. Each VM has a temporary storage space where the participant output files are stored during the testing phase. This temporary data is deleted each time the VM is shut down. All the participant VM instances have the same computing specifications and capabilities. Participants can remotely access their VMs during the training phase. Moreover, they can install all the tools and libraries needed to run their algorithms. At this

stage they can optimize their approaches with the available training set. Specification guidelines were written by the administrators for each benchmark on the usage and permissions applying to the VMs. Through the participant dashboard in the VISCERAL registration and management system, participants received the private access credentials for their VM and had the option to start it or shut it down during the training phase. All the benchmark specifications and usage guidelines were also available in the dashboard.

#### ***4.4 Cloud Infrastructure Setup and Management Experience Report***

The VISCERAL project organized the first series of benchmarks with a large scale 3D radiology image data set using an innovative cloud-based evaluation approach. Having the data stored centrally yields legal, administrative and practical solutions to organizing benchmarks with large data sets:

- The data can be allocated in a private storage container that complies with the legal requirements from the data providers (also HIPAA compliant — Health Insurance Portability and Accountability Act).
- Better control over the project costs, since a cloud platform is flexible enough to increase or reduce the number of VMs and storage containers according to the shifting challenge needs.
- The scalability and storage capacity of a cloud platform is virtually unlimited. This feature opens up the possibility to run benchmarks on big data sets with a high number of participants.
- Different access permissions to the data are defined by the organizers with the option to make some information inaccessible to the participants (e.g. test set).
- The submitted algorithms can be evaluated by the administrators without the intervention from the participants. This generates an objective evaluation of the execution and results.

##### **4.4.1 Lessons Learned**

In retrospective, the following steps were highlighted due to their favourable influence in the benchmark success:

- Planning must consider every component of the platform and how to seamlessly connect them when running the benchmarks. Some early decisions in the project can have a decisive effect in the long term of the evaluation process.
- Clear specifications are paramount and should be defined in great detail from the beginning. Both organizers and participants need to rely on these specifications throughout the project.

- Continuous assessment of the infrastructure needs to be based on the participant feedback. The updates coming from these assessments need to be well documented and transparent to the participants.

#### 4.4.2 Current Challenges

Running benchmarks in the cloud is a significant paradigm shift [4] and requires an adaptation period, both from the benchmark organizers, as well as from the participant perspective. It can be challenging to move away from the classic model of providing the data to the participants (i.e. downloading data sets via traditional file transfer protocols such as FTP) and towards the new way of bringing the participants to the data (i.e. giving the participants data access through a virtual machine in the cloud). Outlined in this section are some pending challenges to consider for future benchmark organizers:

- **Narrow time frame for planning, setting up the infrastructure and running benchmarks** Access and usage of a cloud platform requires a learning period for most of the participants. Strict timelines for isolated benchmarks can limit the number of participants that go through the process of registering, training their algorithms and submitting their VMs. Having a continuous cycle of benchmarks (e.g. annual events) might attract more participants to adapt their algorithms in the cloud and eventually submit results for the benchmarks. This is not always possible because of the strict timelines to run the benchmarks and analyse the participants results for a finite number of competitions.
- **Freedom to adapt the cloud platform** Apart from data security and confidentiality considerations, using a public cloud environment can impact the level of customization available to administrators. The provided management tools need to be used as is, and little to no possibilities exist to adapt them to more specific needs. Scheduled maintenance operations can also make the platform unavailable or cause disruptions in the benchmark workflow. A private cloud may provide a more flexible environment to develop an evaluation framework. On the other hand, a public cloud simplifies setting up and maintaining the backbone of the platform and theoretically allows for limitless scalability. Choosing the right cloud computing option depends on the initial objectives of the benchmark.
- **Some of the components of the VISCERAL infrastructure were implemented outside the Azure cloud platform.** This was mainly due to the limited time frame to set up the needed infrastructure for running and preparing the benchmarks in the cloud. Technical limitations, such as reduced internet connection speed, as well as unfamiliarity of the users with the cloud environment were also hurdles in setting up the benchmarks. Having all the system components in the cloud would have allowed for a more streamlined benchmark organization process.
- **No uniform participant working environments** Managing different operating systems and heterogeneous participant prototyping languages and tools increased the workload of setting up the infrastructure : compilation of evaluation tool for

different platforms, handling OS differences in the automation process (Windows, Linux), VM maintenance, etc. Using a single family of OS could harmonize the infrastructure management tasks. However, this might result in less groups participating overall.

## 5 Conclusions

With a high number of participants in an evaluation benchmark, the administrative tasks for organisers represent a large amount of work. Using a system providing a high level of automation, the amount of work by organisers as well participants can be significantly reduced. In order to achieve this the VISCERAL Registration System, which is a platform that provides participant management and an interface between participants and organisers as well as between participants and the cloud, was developed. Participants in the VISCERAL benchmarks have used this system to register for the the various competitions but also in order to indirectly interact with the organisers and the cloud infrastructure. The main functionalities of the system included the handling of the registration process, the account activation, the management of usage agreements, the assignment and the submission of the VMs as well as the evaluation and the storage / provision of results. The development of such a system helped us to greatly reduce the time spent on administrative tasks such as e-mail exchange with participants and manual cloud interaction. Using a similar approach for running future competitions in the cloud can thus be highly recommended.

The use of a cloud-based infrastructure allowed straightforward scaling up of the VISCERAL benchmarks in terms of storage space and computation power (i.e. number of participants). Centralizing the data and providing standardized Virtual Machine instances to participants allowed us to streamline the management of the evaluation procedure. Certain crucial aspects that should be taken into consideration when setting up an Evaluation-as-a-Service platform were highlighted during the project, such as the importance of exhaustive planning and definition of clear specifications. Minimizing manual tasks and parts of the platform running outside the cloud infrastructure can help save large amounts of time, especially as the number of participants increases. Cloud-based evaluation platforms certainly represent the future and will become more used as researchers get familiar with this new paradigm.

## References

1. Riding the wave: How europe can gain from the rising tide of dcientific data. Submission to the European Comission, available online at <http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf> (October

- 2010), <http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf>
2. Aziz Taha, A., Hanbury, A.: Metrics for evaluating 3d medical image segmentation: analysis, selection, and tool. *BMC Medical Imaging* 15(1) (2015)
  3. Hanbury, A., Müller, H., Langs, G., Weber, M.A., Menze, B.H., Fernandez, T.S.: Bringing the algorithms to the data: cloud-based benchmarking for medical image analysis. In: CLEF conference. *Springer Lecture Notes in Computer Science* (2012)
  4. Hopfgartner, F., Hanbury, A., Müller, H., Kando, N., Mercer, S., Kalpathy-Cramer, J., Potthast, M., Gollub, T., Krithara, A., Lin, J., Balog, K., Eggel, I.: Report on the evaluation-as-a-service (eaas) expert workshop. *ACM SIGIR Forum* 49(1), 57–65 (2015)
  5. Jiménez-del-Toro, O.A., Hanbury, A., Langs, G., Foncubierta-Rodríguez, A., Müller, H.: Overview of the VISCERAL retrieval benchmark 2015. In: *Multimodal Retrieval in the Medical Domain (MRMD) 2015. Lecture Notes in Computer Science*, vol. 9059. Springer (2015)
  6. Langs, G., Hanbury, A., Menze, B., Müller, H.: Visceral: Towards large data in medical imaging — challenges and directions. In: Greenspan, H., Müller, H., Syeda-Mahmood, T. (eds.) *Medical Content-Based Retrieval for Clinical Decision Support*. pp. 92–98. *Lecture Notes in Computer Science*, Springer, Berlin, Heidelberg (2012)
  7. Jiménez del Toro, O.A., Goksel, O., Menze, B., Müller, H., Langs, G., Weber, M.A., Eggel, I., Gruenberg, K., Holzer, M., Kotsios-Kontokotsios, G., Krenn, M., Schaer, R., Taha, A.A., Winterstein, M., Hanbury, A.: VISCERAL – VISual Concept Extraction challenge in RAdioLogY: ISBI 2014 challenge organization. In: Goksel, O. (ed.) *Proceedings of the VISCERAL Challenge at ISBI*. pp. 6–15. No. 1194 in *CEUR Workshop Proceedings*, Beijing, China (May 2014)
  8. Widmer, A., Schaer, R., Markonis, D., Müller, H.: Gesture interaction for content-based medical image retrieval. In: *ICMR* (2014)