Number 113 April 2018

# ERCIM NEWS

www.ercim.eu

## Special theme:

# Smart Farming

Also in this issue:

Research and Society:

## Research Evaluation

*ERCIM News is the magazine of ERCIM. Published quarterly, it reports on joint actions of the ERCIM partners, and aims to reflect the contribution made by ERCIM to the European Community in Information Technology and Applied Mathematics. Through short articles and news items, it provides a forum for the exchange of information between the institutes and also with the wider scientific community. This issue has a circulation of about 6,000 printed copies and is also available online.*

*ERCIM News is published by ERCIM EEIG*
*BP 93, F-06902 Sophia Antipolis Cedex, France*
*Tel: +33 4 9238 5010, E-mail: contact@ercim.eu*
*Director: Philipp Hoschka, ISSN 0926-4981*

***Contributions***
*Contributions should be submitted to the local editor of your country*



***ERCIM News online edition***
*http://ercim-news.ercim.eu/*

***Next issue***
*October 2017, Special theme: Digital Humanities*

***Subscription***
*Subscribe to ERCIM News by sending an email to en-subscriptions@ercim.eu or by filling out the form at the ERCIM News website: http://ercim-news.ercim.eu/*



## Contents

### JOINT ERCIM ACTIONS



### RESEARCH AND SOCIETY

This section on "Research Evaluation was edited in cooperation with Informatics Europe by the guest editors Hélène Kirchner (Inria) and Fabrizio Sebastiani (ISTI-CNR)



### SPECIAL THEME

The special theme section "Smart Farming" has been coordinated by Mirco Boschetti (IREA-CNR) and Erwin Schoitsch (AIT)

## RESEARCH AND INNOVATION

This section features news about research activities and innovative developments from European research institutes



## EVENTS, IN BRIEF

Report



Announcements

Call for Nominations

# 2018 Cor Baayen Young Researcher Award

***The Cor Baayen Young Researcher Award is awarded each year to a promising young researcher in computer science or applied mathematics. The award was created in 1995 to honour the first ERCIM President. The award carries a prize of € 5000.***

### Eligibility

- Nominees must have carried out their work in one of the 'ERCIM countries': Austria, Cyprus, Finland, France, Germany, Greece, Hungary, Italy, Luxembourg, Norway, Poland, Portugal, Sweden and The Netherlands. Nominees must have been awarded their PhD (or equivalent) after 30 April 2015.
- A person can only be nominated once for the Cor Baayen Award.

The Cor Baayen award is a young researcher prize: The selection panel will therefore consider the quality of the PhD thesis and all the achievements done up to the nomination date.

The nominees must have performed their research during at least one year in an institution located in one of the countries mentioned above.

### Submitting a nomination

ominations have to be made by a staff member of an ERCIM member institute. Self nominations are not accepted. Each institute can forward two nominations to the final evaluation. If there are more than two nominations coming from an ERCIM member institute, it is the responsibility of this institute to select the maximum two nominees.

Nominations must be submitted through the Cor Baayen Award Nomination Form and provide a URL to the candidate's full text PhD thesis as well as the candidate's full text best papers or other main scientific achievements (max. 5).

**Deadline for submitting a nomination: 21 May 2018**

### Selection

The selection of the Cor Baayen Young Researcher Award winner is the responsibility of the ERCIM Human Capital Task Group, who might consult expert opinion in reaching their decision.

### Further Information

Can be obtained from your ERCIM representative or from the Cor Baayen Young Researcher Award coordinator Claude Kirchner, Inria (claude.kirchner@inria.fr).

The Cor Baayen award winner will be invited to present her or his contribution highlights at a joint award ceremony between ERCIM and Informatics Europe during ECSS 2018 in Gothenburg (Sweden), 08-10 October 2018.

### About Cor Baayen Cor Baayen

The Cor Baayen Young Researcher Award is named after the first president of ERCIM and the ERCIM 'president d'honneur'. Cor Baayen played an important role in its foundation. Cor Baayen was scientific director of the Centrum voor Wiskunde en Informatica (CWI) in the Netherlands, from 1980 to 1994.

Baayen joined the institute in 1959 as a researcher of the department of pure mathematics, where he became group leader in 1965. As scientific director, Cor Baayen convinced the government to include CWI in a program that stimulated Dutch computer science research. He initiated or stimulated several new research areas including cryptography, computer algebra and performance analysis.

# HORIZON 2020 Project Management

A European project can be a richly rewarding tool for pushing your research or innovation activities to the state-of-the-art and beyond. Through ERCIM, our member institutes have participated in more than 80 projects funded by the European Commission in the ICT domain, by carrying out joint research activities while the ERCIM Office successfully manages the complexity of the project administration, finances and outreach.

The ERCIM Office has recognized expertise in a full range of services,including:

- Identification of funding opportunities
- Recruitment of project partners (within ERCIM and through our networks)
- Proposal writing and project negotiation
- Contractual and consortium management
- Communications and systems support
- Organization of attractive events, from team meetings to large-scale workshops and conferences
- Support for the dissemination of results.

**How does it work in practice?**

Contact the ERCIM Office to present your project idea and a panel of experts within the ERCIM Science Task Group will review your idea and provide recommendations. Based on this feedback, the ERCIM Office will decide whether to commit to help producing your proposal. Note that having at least one ERCIM member involved is mandatory for the ERCIM Office to engage in a project.

If the ERCIM Office expresses its interest to participate, it will assist the project consortium as described above, either as project coordinator or project partner.

**Please contact:**

Peter Kunz, ERCIM Office
peter.kunz@ercim.eu

# ERCIM "Alain Bensoussan" Fellowship Programme

*ERCIM offers fellowships for PhD holders from all over the world. Topics cover most disciplines in Computer Science, Information Technology, and Applied Mathematics. Fellowships are of 12 months duration, spent in one ERCIM member institute. Fellowships are proposed according to the needs of the member institutes and the available funding.*

> "Many thanks to ERCIM fellowship program for providing me the exceptionally positive experience that not only helped me to improve my scientific knowledge but also helped me in many valuable International collaborations.
>
> **Basant AGARWAL**
> Former ERCIM Fellow

> "Success is based on choice, having and maintaining a motivation worth fighting for. I truly appreciate ERCIM Alain Bensoussan Fellowship Program that brought me the position of a Sr. Assistant Professor. I can think of no better postdoctoral fellowship other than ERCIM as it encourages fellows to pursue their own ideas and develop broad collaborations.
>
> **Monalisa MANDAL**
> Former ERCIM Fellow

### Who can apply?

The Fellowships are available for PhD holders from all over the world.

### What is the duration?

Fellowships are of a 12 months duration, spent in one of the ERCIM institutes.

### How to apply?

Only online applications are accepted.

### Which topics/disciplines?

The topics focus on the ERCIM institutes' scientific fields of competence.

### Where are the fellows hosted?

Only ERCIM members can host fellows. For information about the possible hosting institutes, see "participating ERCIM members". For more details, see the homepages of the individual members, or contact the national contact for the ERCIM Fellowship Programme for a specific institute. When an ERCIM member is a consortium the hosting institute might be any of the consortium's members. When an ERCIM Member is a funding organisation, the hosting institute might be any of their affiliates. Fellowships are proposed according to the needs of the member institutes and the available funding.

### What are the conditions?

- have obtained a PhD degree during the last 8 years (prior to the application deadline) or be in the last year of the thesis work with an outstanding academic record
- be fluent in English
- be discharged or get deferment from military service
- the fellowship is restricted to two terms (one reselection possible)
- have completed the PhD before starting the grant (a proof will be requested before your arrival).

We encourage not only researchers from academic institutions to apply, but also scientists working in industry.

In order to encourage mobility:
- a member institute will not be eligible to host a candidate of the same nationality.
- a candidate cannot be hosted by a member institute, if by the start of the fellowship, he or she has already been working for this institute (including phd or postdoc studies) for a total of 6 months or more, during the last 3 years.

**Application deadlines for the next rounds:**

**30 April and 30 September 2018**

**More information:**
http://fellowship.ercim.eu/

**Please contact:**
Emma Lière
ERCIM Fellowship coordinator
fp-info@ercim.eu

# Digitalisation of Seas and Oceans is the New Gold of the Future

by Sara Garavelli (Trust-IT)

“Supporting Blue Growth with innovative applications based on EU e-infrastructures” was the motto of the final event of the BlueBRIDGE project [L1], held on 14-15 February 2018 with participation of over 50 stakeholders, including policy makers, marine and maritime data practitioners, representatives from research institutions, international organisations, governments and EU-funded initiatives.

Blue Growth is the long-term strategy to support sustainable growth in the marine and maritime sectors as a whole. At the same time, Blue Growth has significant data needs that are only partially met. The BlueBRIDGE H2020 project that began in 2015 has worked hard to address these data gaps by developing innovative solutions in the form of Virtual Research Environments (VREs). The final event proposed a rich selection of BlueBRIDGE data services to these future adopters as well as to inventory potential interest for further collaboration with private and public stakeholders.

### How does BlueBRIDGE support Blue Growth? Is there a concrete demand for global repositories of merged public data?

The main mission of BlueBRIDGE is the exploitation of data e-infrastructures to facilitate communities of scientists, data practitioners, SME innovators and educators from various domains to address Blue Growth challenges related to the knowledge production chain. The major barriers that these communities currently encounter are related to i) the difficulty of accessing heterogeneous data that are spread through multiple remote sources and ii) the complexity of the technologies available to process these data. The barriers make a data-centred research process quite slow, if not impossible to pursue.

The solutions developed by BlueBRIDGE aim at removing these barriers. BlueBRIDGE is serving the Blue Growth sector by having developed a set of Virtual Research Environments (VREs), collaborative applications that data managers and policy makers can use to access and analyse data in an easy way. All of the BlueBRIDGE VREs rely on a European hybrid data e-infrastructure, D4Science (www.d4science.org).

### The BlueBRIDGE Services at a glance

FFacilities offered by BlueBRIDGE include services that render the usage of computing resources transparent for the end user, from data production and representation to the discovery of and generation of complex data products. The underlying BlueBRIDGE infrastructure automatically collects information on the processes performed, facilitating the repeatability and reproducibility of experiments and the reuse of results. These features were particularly important in recent studies conducted to estimate the spread of the pufferfish in the Mediterranean Sea, and to develop a general purpose approach which can also be applied to other species, as well as in the climate change area.

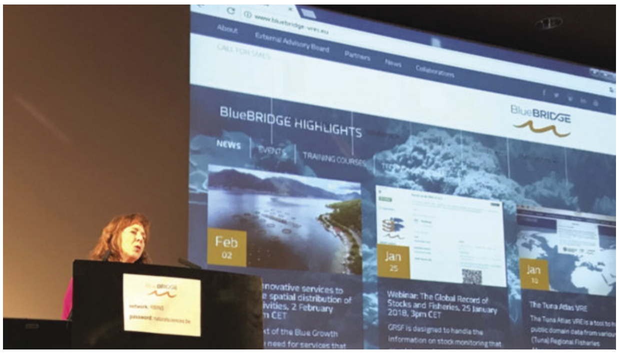

*Donatella Castelli, Director of Research at CNR-ISTI Pisa and BlueBRIDGE Coordinator opening the event.*

### Is there a concrete demand for such global data repositories?

In a panel session during the workshop, the participants explained that it is important to build such repositories because the aggregated repository can generate added value data. And, this doesn’t mean that the single data source will lose its value! On the contrary, the aggregate data can highlight even better the mandate and the purpose for which the individual data source was created.

The panel also highlighted three fundamental aspects for the successful implementation of global repositories:

- Data providers contributing to merged repositories must be involved in the discussions from the beginning to truly understand the benefits and also to save costs that arise from the adoption of new standards. Marc Taconet (Food and Agriculture Association of the United Nations) also reported an important outcome from the recently held FIRMS meeting: three data providers contributing to the implementation of the Global Record of Stocks and Fisheries Knowledge Database (GRSF) confirmed the importance of having such a dialogue, because the exercise of comparing their data allowed them to improve their individual datasets.
- A policy framework supporting the implementation is fundamental. In particular, there is the need to establish a mechanism to discuss standards during the implementation process (because some of the contributors could reconsider what they have done and agree to modify some of their standards) and the need to establish a clear mechanism for content management so that when there are conflicting data, a discussion can ensue as to whether the data should be published or not.
- The underlying infrastructure must respect the ownership of the data providers by reporting citation of the sources, and it has to be flexible enough to quickly react to the changes required by the providers and the aggregator.




*Sieglinde Gruber, DG Research and Innovation, European Commission.*


*Panel on global repositories of merged public data.*

### We can only understand the seas and the oceans if we approach them as ecosystems

"Digitalisation of seas and oceans is the new gold of the future, but we can only be successful if we re-use existing data and applications, and if we look at the earth as a system of land, ocean and space," said Sieglinde Gruber, Head of the Marine Resources Unit, DG Research and Innovation of the European Commission.

Gruber recognised how BlueBRIDGE was able to work in that direction by bringing together different stakeholders, reusing existing technologies and datasets, and trying to leverage as much as possible on previous investments. In particular, the wide scope of BlueBRIDGE was appreciated in terms of the variety of stakeholders reached and its use of marine data combined with satellite data to address the data challenges of the "earth" ecosystem.

### How BlueBRIDGE is contributing to different Blue Growth data management challenges

*BlueBRIDGE supporting stock assessment*

Together with ICCAT, the International Commission for the Conservation of Atlantic Tunas, BlueBRIDGE is performing stock assessment of eastern Bluefin tuna. The VRE developed by the project enables a reproduction of the work done to update the projections analysis of the Eastern Bluefin tuna stock assessment. The analysis is used for providing scientific advice to define the quota and to try new parameterizations of the model for future work. The online approach offered by the VRE has been welcomed by ICCAT for Eastern Bluefin stock assessment.

*BlueBRIDGE supporting maritime spatial planning*

A Protected Area Impact Maps VRE, developed by GRID-Arendal improves the understanding of the spatial distribution of existing and planned marine protected areas and their coverage of seafloor features. The VRE is the result of a joint collaboration with the European Commission Joint Research Centre (JRC). This is the second time the two teams join their forces: by combining a wide range of open source tools, a new machine with increased functional capacity will ease data processing workflow, sharing it as a versioned resource. This progress will significantly facilitate access and replicability of codes, environmental settings, parameters and maps on the status of marine protected areas in African, Caribbean and Pacific (ACP) countries. Biodiversity conservation specialists working in BIOPAMA, The Biodiversity and Protected Areas Management Programme, should benefit enormously.

BlueBRIDGE developed services that collect and combine Environmental Observation (EO) data with aquaculture data presenting the three VREs to support a computing intensive ontology driven feature analysis of Synthetic Aperture Radar (SAR) and multispectral optical imagery (using Sentinel-1 and -2 data, and Very High Resolution optical imagery), where the results are displayed on maps for human reviewers. The first two VREs are specialised in recognizing offshore fish farming in Greece and Malta, whilst the third one is specialized in identifying coastal aquaculture ponds in Indonesia.

*BlueBRIDGE Supporting aquaculture*

Using the services provided by BlueBRIDGE, aquafarmers can estimate the performance of their production exploiting state of the art Machine Learning methods based on the real historical production data. Furthermore, they are able to make accurate production plans, future investment plans by exploiting the geoanalytics platform and techno-economic analysis combining production, financial and environmental data. In this way, they can make correct and timely decisions to strengthen their aquaculture's position against competition.

*BlueBRIDGE supporting education*

Efforts made by BlueBRIDGE and other initiatives address training required to access datasets or executing data-intensive models which at the moment are quite costly and effort consuming. The BlueBRIDGE VREs have demonstrated to be a good instrument to support this type of training and they are considered particularly useful in developing countries where internet connection and technologies competences are not well developed.

For more details, all presentations of the BlueBRIDGE final event can be found at [L2]

The BlueBRIDGE project is coordinated by ISTI-CNR, while ERCIM EEIG provides administrative and financial support.

**Links:**

[L1] https://www.bluebridge-vres.eu | @BlueBridgeVREs
[L2] https://kwz.me/hbV

**Please contact:**

Donatella Castelli, ISTI-CNR
donatella.castelli@isti.cnr.it


Illustration: John McKiernan / Source: http://whyopenresearch.org

# Research Evaluation in Informatics

This section was edited in cooperation with Informatics Europe. Guest editors: Hélène Kirchner (Inria) and Fabrizio Sebastiani (ISTI-CNR)

***Research evaluation, as applied to individual researchers, university departments, or research centres, plays a crucial role in recognising and supporting research that can lead to advances in knowledge and benefits to society.***

Evaluation can have a tremendously positive effect in improving research quality and productivity. At the same time, the effect of following wrong criteria or practices in research evaluation can have seriously negative long-term effects, potentially on several generations of researchers. These negative effects range from demotivating researchers who, despite performing good-quality research, get evaluated negatively, to wrongly promoting unworthy research endeavours at the expense of worthy ones.

To achieve the positive effects of research evaluation, its specific goals must be clearly formulated upfront, and it must be performed in a transparent way that is aligned with these goals. The evaluation should follow established principles and practical criteria, known and shared by evaluators, researchers, and the broader scientific community.

Research evaluation should indeed take into account the specificities of Informatics as a new science and technology, and its rapid and pervasive evolution. Informatics Europe recently published a report on this topic entitled “Informatics Research Evaluation”. The report focuses on research evaluation performed to assess individual researchers, typically for promotion or hiring. The recommendations and conclusions are outlined in the first article of this section.

The Informatics Europe report also raises several questions on the publication culture and its evolution, on the importance of artifacts, and on impact assessment for this scientific domain. Several contributions in this special section are devoted to exploring these questions.

### How to evaluate the quality and impact of publications?

Dino Mandrioli (Politecnico di Milano) addresses the conferences vs. journals controversy. He explains why focusing on conference publications, rather than on journal articles, may be harmful for research quality and evaluation in informatics, and he gives serious arguments in favour of journals. He argues that conferences should go back to their original, authentic goal, i.e., circulation and discussion of ideas; journals, in his view, then remain the best and natural (although not the only) medium for publication of research results, and research evaluations should treat them accordingly. The recent trend towards the tight coupling of conferences and journals, with conference papers appearing in a journal, or conferences incorporating journal-first papers into their program, could reconcile the divergent views.

Stefano Mizzaro (University of Udine) argues in favour of an alternative to traditional peer review that relies on crowd-sourcing, i.e., that exploits data and information from fellow researchers who read, spread, comment on, and cite papers. The idea is to compute a quality index of papers based on the post-publication comments they receive; in turn, this allows the computation of quality indexes for researchers as authors and researchers as readers, via algorithms in which these indexes mutually reinforce each other. Although it is unclear whether the quality of the scientific literature will be improved by crowd-sourcing peer review, proposals such as this warrant further study. The development of public archives and the concept of overlay publication might offer an opportunity to experiment with this idea.

### How to evaluate software, artifacts and outreach?

Alain Girault and Laura Grigori (Inria) address the growing importance of software in academic research and, as a consequence, the necessity of taking software development activities into account when evaluating researchers. They support this idea by describing the software evaluation procedure at Inria that allows researchers to characterise their own software along different criteria, in the context of recruitment or team evaluation. Software evaluation goes together with the ability to reproduce experimental results described in a publication, and opens the way to a new publication model which, in addition to open access, also includes open data and open software, simultaneously released.

### How to take into account open science criteria?

Laurent Romary (Inria) addresses the link between open access and research assessment and argues that, despite some fears, open access offers the potential of a wealth of information for the strategic management of research, typically for identifying experts or emerging topics in a given field. He points out challenges related to the possible fragmentation of the publication corpus and to the actual quality of the available information, but argues that if all the proper conditions are fulfilled for a trusted information corpus of scientific publications, we can actually foresee the basis for an open transparent process for research assessment. This opens a wide variety of usages that we have to invent and deal with in an ethical way. He suggests for instance a visibility index related to the dissemination effort of a researcher or a research institution. He concludes with a reminder that open access on research data and tools ultimately exist to help researchers themselves.

### How to measure scholarly impact?

Giovanni Abramo (IASI-CNR) revisits the concept of scholarly impact, i.e., the footprint that a research paper leaves on scientific research performed after its publication. He remarks that, assuming that the amount of citations obtained by a certain paper is a good proxy of its scholarly impact, this impact can be measured only at the end of the "life cycle" of the paper, i.e., after it has ceased to influence current research (and thus ceased to be cited). This is too late for the practical needs of those who need to evaluate research (e.g., a researcher, for promotion or hiring); this means that there is a need of "early indicators of impact", i.e., measures capable of predicting whether and how much a paper is going to be cited in the future. Giovanni Abramo recalls current efforts, essentially based on (a) using as features the early citation counts plus the impact factor of the journal in which the paper was published, and (b) using late citation counts as benchmarks for the accuracy of these predictors, but stresses in conclusion that more research is needed in this area.

Many more questions on research evaluation could be raised and good practices exchanged, but we should be aware that each of us, as soon as we are involved in any evaluation process, has the opportunity to encourage and promote research evaluation criteria which should be based primarily on quality and impact.

**Please contact:**
Hélène Kirchner, Inria, France
helene.kirchner@inria.fr

Fabrizio Sebastiani, ISTI-CNR, Italy
fabrizio.sebastiani@isti.cnr.it

# Recommendations for Informatics Research Evaluation

***In March 2018, Informatics Europe published a report [1] focusing on the main principles and criteria that should be followed when individual researchers in informatics (computer science) are evaluated for their research activity, addressing the specificities of this area. This subsumes evaluation of a specific piece of research and can often be generalised to university departments or research centres, since their research performance is largely determined by their individuals.***

This report confirms the findings of the 2008 Informatics Europe report (Research Evaluation for Computer Science, Informatics Europe Report, Eds. Bertrand Meyer, Christine Choppy, Jan van Leeuwen and Jørgen Staunstrup) on the subject, and at the same time incorporates a number of new observations concerning the growing emphasis on collaborative, transparent, reproducible, and accessible research. The conclusions of the report are summarised in nine recommendations to people involved in evaluation committees and funding agencies:

1. Informatics is an original discipline combining mathematics, science, and engineering. Researcher evaluation must adapt to its specificity.

2. A distinctive feature of publication in informatics is the importance of highly selective conferences. Journals have complementary advantages but do not necessarily carry more prestige. Publication models that couple conferences and journals, where the papers of a conference are published directly in a journal, are a growing trend that may bridge the current gap between these two forms of publishing.

3. Open archives and overlay journals are recent innovations in the informatics publication culture that offer improved tracking in evaluation.

4. To assess impact, artifacts such as software can be as important as publications. The evaluation of such artifacts, which is now performed by many conferences (often in the form of software competitions), should be encouraged and accepted as a standard component of research assessment. Advances that lead to commercial exploitation or adoption by industry or standard bodies also represent an important indicator of impact.

5. Open science and its research evaluation practices are highly relevant to informatics. Informatics has played a key enabling role in the open science revolution and should remain at its forefront.

6. Numerical measurements (such as citation and publication counts) must never be used as the sole evaluation instrument. They must be filtered through human interpretation, specifically to avoid errors, and complemented by peer

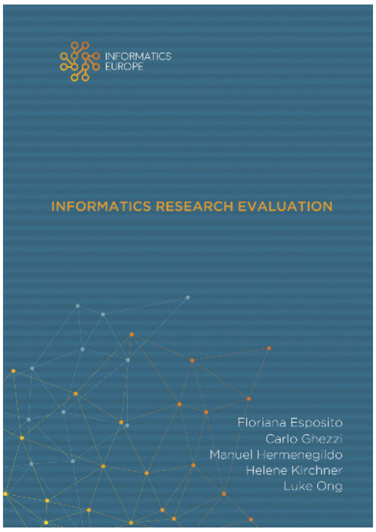
review and assessment of outputs other than publications. In particular, numerical measurements must not be used to compare researchers across scientific disciplines, including across subfields of informatics.

7. The order in which a publication in informatics lists authors is generally not significant and differs across subfields. In the absence of specific indications, it should not serve as a factor in the evaluation of researchers.

8. In assessing publications and citations, the use of public archives should be favoured. When using ranking and benchmarking services provided by for-profit companies, the respect of open access criteria is mandatory. Journal-based or journal-biased ranking services are inadequate for most of informatics and must not be used.

9. Any evaluation, especially quantitative, must be based on clear, published criteria. Furthermore, assessment criteria must themselves undergo assessment and revision.

These recommendations are consistent with a recent statement of three national Academies (Statement by three national academies (Académie des Sciences, Leopoldina and Royal Society) on good practice in the evaluation of researchers and research programmes) that also provides recommendations on evaluator selection, overload and training.




**Links:**
The Informatics Europe report is now available on its web side in its final form:
http://www.informatics-europe.org/publications.html.

The direct link to download is:
http://www.informatics-europe.org/component/phocadownload/category/10-reports.html?download=76:research-evaluation-2018

**Reference:**
[1] Floriana Esposito, Carlo Ghezzi, Manuel Hermenegildo, Hélène Kirchner and Luke Ong (eds): "Informatics Research Evaluation, Informatics Europe Report, 2018

# On the Conferences vs. Journals Controversy

by Dino Mandrioli (Politecnico di Milano)

***Conferences are harmful for research evaluation.***

The document [1] addresses the critical issue of informatics research evaluation in a complete and thorough way, touching all the main facets of the problem. I substantially agree with most of the arguments and proposals presented therein. In this note, however, I discuss the only point of serious disagreement, i.e., the role assigned to conferences for the evaluation of research in computer science. I admit that it is a fact that in our community – probably the only scientific community that exhibits this specificity – conferences often overwhelm journals in the research assessment procedure, but I believe that this anomaly is seriously flawed, dangerous, and should be contrasted – not be accepted or, even worse, encouraged.

There are two main arguments to support preferring conferences over journals as the main publication medium:
1. Journal publishing takes too long
2. Some conferences are even more selective than many journals.

My main counterpoints to the above arguments are:
1.a) It is true that reviewing (nowadays not publishing!) journal papers often takes a long time, but for highly technical papers this is necessary. Research advances quickly but its evaluation needs its own time!
1.b) Some highly prestigious journals (e.g., JACM, TOPLAS) have adopted effective policies to dramatically reduce the average response time: typically, a preliminary quick review states whether the paper is readable and interesting enough to deserve a thorough further reviewing effort; if a paper passes this phase, the author knows that acceptance will depend mostly on technical soundness and can be more patient. See also the symmetric point 2.b) about conference reviewing.
1.c) From a more general, "social", point of view, this claim is a symptom of the typical "time to market" syndrome which often causes serious damage even in the industrial world: research needs serious evaluation before "going to the market" (remember the "cold fusion" phenomenon).

2.a) Top journals are highly selective too, sometimes even more than top conferences; furthermore, often people try to "submit anyway" to conferences even if the paper is not yet mature (e.g., due to the deadline); or they submit several papers maybe hoping for some "good luck" for some of them. Many authors are more conservative and careful when submitting to high-level journals.
2.b) Being highly selective does not guarantee high quality. The conference and journal reviewing processes are necessarily different: having many papers to review with hard deadlines, even for serious committee members, leads to superficial reading where generic interest, "sympathy for the topic" (or even for the author), read-

ability, and, in some cases, distributing the load to several sub-reviewers, overwhelms thorough and comparative evaluation. The typical conference bidding process tends to produce clusters of strongly related topics, thus producing the risk of self-referencing environments, just the opposite of the real goal of cross-circulating new ideas. Furthermore, in program committee (PC) meetings, whether in person or electronic, the prevailing opinion is often the one of the member who "speaks the loudest" ("over my dead body").

The 2014 "NIPS experiment" has been mentioned a lot lately (see e.g., [2]), where the PC was split into two independent subcommittees which agreed only on the top and bottom evaluations and whose opinions varied substantially on about 60% of the papers in-between. This fact was interpreted in various ways; in some cases it was even used as an argument in favour of keeping top conferences highly selective: "If we pick up just the top 10%, we are sure that we reward just the best papers". I have little doubt that the bottom 10% warrant rejection, but this approach risks rejecting many potentially important papers in the grey area – some of which are probably even more interesting than those unanimously accepted. The literature exhibits plenty of examples of pioneering papers that suffered from several rejections before the community appreciated their value. Important, novel contributions are often controversial and misunderstood; it is easier to gain general appreciation by means of results in fields that are already familiar to a wide audience. Of course, the same problem may occur even with journal publishing, but in this scenario it can be mitigated by a careful selection of reviewers, by the opportunity for rebuttal, and by …more time and thought for the final decision.

2.c) The tight page limits compel authors to submit "ε-papers": small, maybe relevant, advances can be more easily published than well-developed, completed, long-term research. Technical details are omitted or confined to appendices, which are rarely checked.

In conclusion, I would like conferences to be relegated to their original, authentic goal, i.e., circulation and discussion of ideas. Journals remain the best, natural, certainly not only, medium for research evaluation.

**References:**

[1] F. Esposito et al.: "Informatics Research Evaluation", An Informatics Europe Report, Draft Version: 20-10-2017

[2] M. Vardi, "Divination by Program Committee", CACM, Vol. 60, N. 9, p. 7, 2017.

**Please contact:**
Dino Mandrioli, Politecnico di Milano, Italy
+39-02-2399-3522
dino.mandrioli@polimi.it
http://home.deib.polimi.it/mandriol/SitoInglese/perswebsiteengl.html

# Evaluation in Academic Publishing: Crowdsourcing Peer Review?

by Stefano Mizzaro (University of Udine)

***In academic publishing, peer review is the practice adopted to evaluate the quality of papers before publication. Each submission undergoes a thorough review by some peers, who decide whether the paper is worthy of publication. The pre-publication filter by peer review is important to guarantee that scientific literature is of high quality. However, although peer review is a well known, widely adopted, and accepted practice, it is not immune to criticism, and its appropriateness has been challenged by many. One recent example is the paper by Tomkins et al. [1].***

Some such opponents go beyond discussion and criticisms, and propose concrete alternatives. Such proposals essentially aim to exploit data and information that are available but are not currently used for evaluation purposes. Indeed, after publication, papers are (hopefully) read by fellow researchers, who usually form an opinion about their quality. However, all these opinions remain private, or are communicated among a few collaborators in an informal way, or perhaps manifest themselves (although in a rather implicit way) in citations. In extreme terms, we might say they are lost (at least, for the purposes of research evaluation). At the very least, the research community fails to leverage precious data and information that could help research evaluation.

Some alternative proposals, including Readersourcing [3] and TrueReview [2] make use of this data. Both rely on the opinions of readers, who are asked directly to judge the papers they have read. Reader assessments are then stored, in an open way, as numerical scores. This would then allow a numerical quality score to be computed for papers, authors, and readers too, by means of an algorithm capable of estimating the quality of the assessments. In other words (see the original publications for all the details), readers are invited to submit a score on the quality of the papers they have read; all the scores concerning a specific paper may then be aggregated (and weighted according to the quality of the readers) in order to compute an overall quality score for the paper; the scores of all papers by a certain author may in turn be aggregated in order to compute a quality index for researchers as authors; and the quality of the expressed assessments may be used to compute an overall quality index for researchers as readers.

These proposals, and similar ones, utilise crowdsourcing [3]: the activity of outsourcing, to large crowds of unknown people, tasks that are usually performed by a few experts. Although this approach might seem unfeasible – and it has had its share of criticism - there exist many successful examples of crowdsourcing. Wikipedia, for instance, is a free, good-quality crowdsourced encyclopaedia; marketplaces for crowdsourcing exist and are flourishing [L1], [L2]; and crowdsourcing platforms specialising in research and development activities are available [L3] .

Indeed, crowdsourcing peer review would also address the issue of the scarcity of reviewers, which seems to be a growing problem, as discussed [3]. Although it is not clear whether crowdsourcing peer review will allow the quality of the scientific literature to remain high, this is an interesting research question that, in my opinion, is worth addressing.

**Links:**
[L1] www.mturk.com
[L2] www.crowdflower.com
[L3] www.innocentive.com

**References:**
[1] A. Tomkins, M. Zhang, W. D. Heavlin: "Reviewer bias in single- versus double-blind peer review", PNAS, 114(48):12708–12713. November 28, 2017, http://www.pnas.org/content/114/48/12708.
[2] L de Alfaro, M; Faella: "TrueReview: A Proposal for Post-Publication Peer Review" (white paper), technical report UCSC-SOE-16-13, 2016.
[3] S. Mizzaro: "Readersourcing - A manifesto",Journal of the American Society for Information Science and Technology, 63(8):1666-1672, 2012, Wiley.

**Please contact:**
Stefano Mizzaro, University of Udine, mizzaro@uniud.it

# Taking Software Development Activities into Account when Evaluating Researchers

by Alain Girault and Laura Grigori (Inria)

***Software is becoming increasingly important in academic research and consequently, software development activities should be taken into account when evaluating researchers (be it individually or teams of researchers). In 2011, a dedicated working group of the Inria Evaluation Committee addressed the issue of software evaluation and proposed criteria that allow researchers to characterise their own software. In the context of recruitment or team evaluation, this self-assessment exercise allows the evaluators to estimate the importance of software development in the activities of the authors, to measure the impact of the software within its target community, and to understand its current state and its future evolution.***

In computer science and applied mathematics, software is crucial for several reasons:
- It can be used as a proof of concept to prototype an original idea.
- It can be distributed to an entire community, thus serving as a dissemination tool (quite often there can even be some competition between teams worldwide).
- It can be used as an experimentation platform (e.g., the Grid5000 management software).
- It can be delivered to industry and allow a faster transfer of academic results to industry.

Competitions between international teams address key challenges and aim at establishing new records with respect to a given criterion (e.g., size of the solved problems, speed or accuracy of the results). This is the case in several domains of computer science, including the model-checking, computer vision, high-performance computing, and the 3D meshing communities. In this context, it is mandatory to devise and implement new algorithms and to make the corresponding software publicly available. This allows the new algorithms to be tested by competitors.

A concrete proof of the growing importance of software is that several journals and conferences accept "software artifacts" to support a published paper, which are formally tested and evaluated. A recent trend focuses on the ability to reproduce the experimental section included in an article. Allowing this reproducibility is an important step forward towards reproducible research: given the same data and the same software, one has to derive the same results. Some journals already provide a "replication certificate" for the papers the results of which have been reproduced. We are convinced that this is an adequate step, and we support a publication model that, in addition to open access, also includes open data and open software, released simultaneously.

There are two different evaluation processes at Inria (in general, not only for software):
- The evaluation of the Inria teams that takes place every four years.
- The evaluation of the Inria researchers when they apply for various promotions or recruitments (in particular when they are hired as junior researchers and later when they apply for a promotion as senior researchers or higher ranks).

Taking software into account when evaluating a researcher or a research team at Inria has been common practice for over 15 years. Reflections on this topic started in the early 2000s within the Evaluation Committee of Inria, and a working group was created in 2011 under the guidance of Gérard Berry, which produced a document that allows the researchers and teams to evaluate their own software development [1]. The working group recommended the use of several qualitative and quantitative criteria to assess a software. This self-assessment process has been used at Inria for seven years and has been demonstrated to work satisfactorily. The criteria are then used by the evaluators to estimate the importance of software development in the activities of the authors of the software, to measure the impact of the software in a specific community, to understand its current state and its future evolution. More precisely, the criteria are:
1. *Audience*: ranging from personal usage to a massive usage.
2. *Software originality*: ranging from implementing known ideas to fully original ideas.
3. *Software maturity*: ranging from a software used for demonstration purposes only to a high assurance software certified by an evaluation agency (e.g., DO178 in civil avionics).
4. *Evolution and maintenance*: ranging from no particular plan to a fully organised support including, e.g., a user group or a forum where users can ask questions.

5. *Software distribution and licencing*: ranging from internal use within the authors' team to external packaging and distribution (e.g., as a Linux, Matlab, or R toolboxes or packages, just to take examples).

For each criterion, the software authors choose between the four or five available levels, spanning the extreme options listed above, the one that best characterises their software. Those levels are not used to rank the software, the higher levels are not necessarily the best. Indeed, the evaluators appreciate not only the size or the maturity of the software, but also the originality of the implemented ideas and algorithms.

In addition, the authors mention their own contribution along the different axes that matter for the production of software: design and architecture, coding and debugging, maintenance and support, and finally project management. For each of these criteria, an individual's contribution can range from occasional contribution to being (one of) the principal contributor(s). This second set of criteria is essential for highlighting the respective contribution of the different researchers to a particular software. As mentioned above, higher levels are not necessarily the best: to provide a concrete example, for the same software there can be the main architect who provided the core algorithms that form the basis of the software, and the main developer who coded the software and decided the architecture and the test and debug plan. Both roles are key to the success of the software.

In an additional succinct section, the researchers can add any information they consider relevant, such as the application targeted by the software, the user's community (research, education, industry), its impact, a short description of the differences with respect to other state-of-the-art software with same functionalities, and so on. This section should also include the programming language, the number of lines of the code, and a url from which the software can be downloaded. The url is very important so that the evaluators can download and test the software themselves. The criteria listed in this report are self-assessment criteria, so being able to download and test the software is essential to validate the self-assessment performed by the software authors. One trend that we are seeing at the moment is that source files are stored on repository systems like git, so the evaluators can monitor precisely the actual contribution of each author to a particular software.

In conclusion, we would like to emphasise that software development plays a crucial role at Inria, it definitely constitutes a key part of the activity of our research teams, both in the applied mathematics and computer science fields. As such, being able to evaluate the software produced by the teams and the researchers is of paramount importance. The criteria for self-assessment reported in [1] have been used for the past seven years and have proved to work very well.

**Reference:**
[1] Inria Evaluation Committee, "Criteria for Software Self-Assessment", 2011.

**Please contact:**
Alain Girault and Laura Grigori, Inria
alain.girault@inria.fr, laura.grigori@inria.fr

# Open Access and Research Assessment: Why Would We Link the Two?

by Laurent Romary (Inria)

***Despite some concerns within the scientific community, open access has the potential to offer a wealth of information for the strategic management of research.***

Recent years have seen an increase in initiatives intended to foster open access, i.e. the free dissemination and re-use of scientific publications, at institutional, national and European levels. The European Union is taking a strong stance on the issue, making open access mandatory in its H2020 program, and funding the OpenAire initiative to coordinate publication repositories all over the continent. Several countries have also taken important legislative initiatives with the "Zweitveröffentlichung" principle in Germany and its equivalent in the recent "Loi pour une République Numérique" in France. Moreover, several institutions, such as Inria (FR), CentraleSupelec (FR) and the University of Liège (BE) have set up a "deposit mandate", whereby annual reporting is entirely based upon the scholarly publications that are made public in designated open publication repositories (e.g., HAL in France).

Whilst this clearly represents major progress for the fast and free dissemination of scientific knowledge, there are concerns within the scientific community that their activity and results may be scrutinised in an inappropriate way, and in particular that researchers' assessments may come to be based upon numerical impact measures computed from the material in publication repositories. Scientists are also uncertain with regards to the possible poorer quality of material put online at an early stage, as well as with the risk of being plagiarized, even if on the contray open access is the best rampart against plagiarism. It is vital that we begin to balance these concerns by acquiring an understanding of the ways in which science and scientists may benefit from an ambitious open access policy, and that we communicate these benefits to the scientific community. For instance, a thorough coverage of the publication corpus of an institution, or of an entire country, could provide a means to analyse co-publication patterns, and in doing so could help identify collaboration strengths or weaknesses that could be fostered or improved. Text and data mining techniques can also provide mechanisms to identify expertise in a given research landscape or emerging topics within a given field. At the end of day, open scientific information could be used in a whole range of ways that we have to invent and deal with in an ethical way, in compliance with the DORA principles, for instance.

Right now we have even more pressing challenges, though, including: a) the possible fragmentation of the publication corpus and b) the actual quality of the available information. The first issue has to do with our capacity to have a coherent


*Illustration 1: Loi pour une république numérique.*

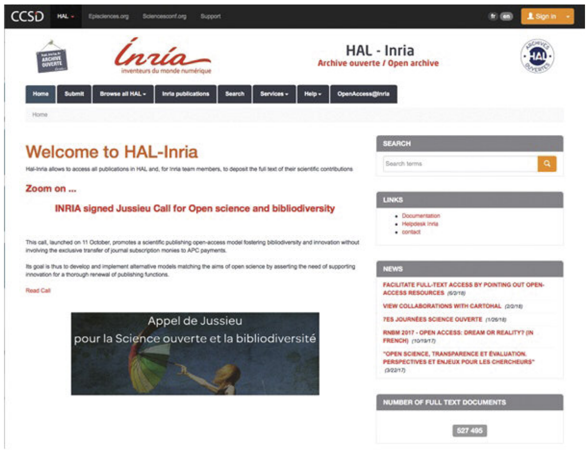



*Illustration 2: The HAL-Inria publication portal.*

picture of the publication landscape and could be based upon the existence of a network of interoperable publication repositories where content, as well as access or download information, can be used, re-used and above all mined under open licences. Relying on private third parties (under the "gold open access model") carries with it the risk that we may be left with diminished re-use possibilities for our own publications, which may end up spread across various publishers' servers.

The second issue is even more central when considering publication information as the basis for the assessment and design of scientific policies. We need to ensure that the documentation associated with publication, comprising author identification, affiliations and precise publication information, is recorded and curated in a way that guarantees a trusted background for further analyses. For instance, it is important to maintain proper (open) authorities for research institutions upon which affiliation information can be based. In this domain we need to determine how much we can rely on third party initiatives such as ORCID where publishers may have too strong a voice.

If all the conditions are met for a trusted information corpus related to scientific publication, this could conceivably form the basis for an open, transparent process for research assessment exercises. In fact, this is exactly what the French assessment institution Hcéres is proposing, with their announcement in early 2018 that future assessment campaigns would rely entirely on the French national publication repository HAL to disseminate a) reference publications, as selected by the research organisations under assessment, as the sole source of information for reviewers and b) the self-assessment reports themselves, so that this information is made available to the wider research community and the public.

In conclusion, if we are serious about the issue of open access to publications - and probably to research data too - we could imagine setting up a visibility index that would relate to the proportion of openly available and re-usable content in a public repository in relation to the actual research output of an individual, team laboratory or institution. While not necessarily useful for "assessing" scientific outputs, such an index could encourage scholars to feel concerned with the wider dissemination of their own research.

At the end of the day, if open access policies are to be used for any form of assessment, its goal should be to assist the researchers themselves, providing them with appropriate data and tools to reflect on their own research trends, practices and, possibly, impact [1] [2]. This implies, in addition to the factors that we have touched on above, that openness also applies to research data, expressed in formats compliant with international standards, and tools, developed and distributed according to open source principles.

**References:**

[1] S. Harnad: "Open access scientometrics and the UK Research Assessment Exercise", Scientometrics, 79(1), 147-156, 2008.
[2] L. Romary, C. Armbruster: "Beyond institutional repositories", International Journal of Digital Library Systems, 1 (1), pp.44-61, 2010. 〈hal-00399881〉

**Please contact:**

Laurent Romary, Inria, France
laurent.romary@inria.fr

# Revisiting the Scientometric Conceptualisation of Impact and its Measurement

by Giovanni Abramo (IASI-CNR and The University of Waikato)

***"Impact", jointly with "citation", is probably the most frequent term in scientometric literature. It recurs so often that its definition is more and more taken for granted, hardly ever spelled out. Impact is often defined by means of the indicators employed to measure it, rather than as a concept in itself. The obsessive pursuit of measurement resembles a case of gold fever. One apparent aspect of this is to "first measure, and then ask whether it counts". The aim of this short note is to revisit the conceptualization of impact, and relevant measurement indicator, which could serve in opening a discussion.***

To better appreciate the role of scientometrics in assessing impact, we need to observe its position within the overall research value chain. Governments and private organizations invest in research to accomplish scientific and technical advancement, which has the potential of contributing to socio-economic progress. Taking for granted that the measurement of "social" impact is beyond the scope of scientometrics, what kind of impact is measured by scientometrics? To answer this question we have to consider scientific activity as "information processing": the science system consumes, transforms, produces, and exchanges "information". Scientometrics studies and analyzes, through the reference lists, the exchanges of information encoded in published papers. Scientists collect and analyze prior knowledge encoded in verbal forms and add value to it for producing new knowledge, nearly always encoded in papers made accessible to other scientists. For the production of new knowledge "to have an impact" it has to be used by other scientists. We can state that evaluative scientometrics is aimed at measuring/predicting the impact of research on future scientific advancement. We can conclude that scientometrics concerns "scientific impact" or "scholarly impact".

But how to measure it? We frequently hear that the proper basis of research evaluation would be for experts to review the work of their colleagues. In fact, reviewers cannot assess impact, at the most being able to predict it, by examining the characteristics of the publication under evaluation.

In scientometrics citation is the natural indicator of impact, as it certifies the use of the cited publication towards the scientific advancement encoded in the citing publication. This position derives from the Mertonian or normative theory [1], according to which scientists cite papers to recognize their influence, being aware that exceptions (uncitedness, undercitation, and overcitation) occur. Although a reviewer might judge certain characteristics of a publication, he/she cannot certify its "use". Also, no other indicator certifies it. Journal impact indicators reflect the distribution of citations of all hosted publications, not the individual ones. Also altmetrics (i.e. Mendeley reader counts) do not certify the proper use of a publication. But this does not mean that journal impact and/or altmetrics cannot be useful in assessing the scholarly impact of a publication.

Ideally, when the life of a publication is over, i.e. when it is no longer cited, the accrued citations reflect the publication's impact. Unfortunately policy-makers and managers, hoping to make informed decisions, cannot wait the many years (or decades) for citation life-cycles to end, and so to conduct research assessments.

The very challenge is: how to predict future impact? In literature the answer is not univocal, but what should not be in discussion is that late citation counts (as proxy of long-term impact) serve as the benchmark for determining the best indicator (and its predictive power), in function of the citation time window. Initially, scholars faced the problem of how long the citation time window should be for the early citations to be considered as an accurate proxy of later citations [2]. They have also investigated the possibility of increasing the accuracy of impact prediction by combining early citation counts with journal impact indicators [3].

The identification of alternative indicators or combinations, which could vary in type and weight across disciplines, as a function of the citation time window, and early-citations accrued, should be conducted by comparing their prediction of impact to the benchmark of late citation counts. This is the actual challenge for our community.

The multitude of assumptions, conventions, limitations, and caveats of evaluative citation analysis still apply; we just hope to have stimulated a proper discussion for a definitive convergence on the meaning and measurement of a fundamental concept of the scientometric science. Differently, every assessment exercise becomes metaphysical.

**References:**

[1] R. K. Merton: "Priorities in scientific discovery", in R. K. Merton (Ed.), The sociology of science: Theoretical and empirical investigations (pp. 286–324), Chicago: University of Chicago Press, 1973.

[2] G. Abramo, T Cicero, C.A. D'Angelo: "Assessing the varying level of impact measurement accuracy as a function of the citation window length", Journal of Informetrics, 5(4), 659-667, 2011.

[3] G. Abramo, C.A. D'Angelo, G. Felici: "Predicting long-term publication impact through a combination of early citations and journal impact factor", working paper, 2017.

**Please contact:**

Giovanni Abramo, IASI-CNR Italy and The University of Waikato, New Zealand
+39 06 72597362, giovanni.abramo@uniroma2.it

Introduction to the Special Theme

# Smart Farming

by Mirco Boschetti (IREA-CNR) and Erwin Schoitsch (AIT)

"Smart Anything, Everywhere" is the new hype around IoT, internet of things, combined with intelligence, autonomy and connectivity. Smart systems are today's drivers of innovation; in all areas of industry and society, highly automated, intelligent systems are taking over tasks, services – and maybe one day, control of our lives.

One of the most diverse, and for many ICT and technology professionals, unfamiliar, areas of application is farming: here the whole life-cycle from soil to fork, from livestock and the field to end-customer, is covered.

With the global population predicted to exceed nine billion by 2040, food production is a major challenge, which will be further exacerbated by climate change, reduced water supplies in many regions, and the environmental impacts of intensive plant and livestock production. The Food and Agricultural Organisation of the UN (FAO) advises that digital technologies be adopted to help increase productivity, to address the food security risk faced in some regions of the world [2]. Studies indicate that if each region of the world can maximise its efficiency of agricultural land management practices, (depending on local climate and environmental conditions) there should be adequate food supplies to support the population.

In Europe, digital technologies could help farmers face other more specific challenges, such as profitability, environmental footprint and sustainability of their businesses.

ICT and agricultural technology advances of recent decades are converging into smart farming, which encompasses precision farming, which covers site specific crop management responding to inter and intra-field variability in crops. This will impact all areas of agriculture, including: efficiency of crop and livestock management; increased production at less cost; conservation of resources such as water and energy; reducing the amount and environmental impacts of fertiliser and pesticides whilst maximising their effectiveness by optimising the timing, location and quantity of application ("minimum environmental footprint"); and preserving and guaranteeing food quality and safety throughout the life cycle and the food value chain in a workflow-like manner. Smart farming is a playground that hosts various stakeholders and technologies: natural science disciplines, including biology, agronomy, meteorology and soil science devoted to crop management and resource preservation, breeding and genetics aiming to develop plants that are more resistant to stressors and require less input; different specialists of livestock and plant health management; a huge variety of sensors from IoT in situ to measure (micro) environmental variables, crop status and animal welfare, and remote sensors such as drones and satellite data to monitor farm and territorial scale conditions; data science to interconnect and exploit existing OGC (Open Geospatial Consortium) data and weather forecast from modelling; ICT solutions for big data handling and interpretation, and mathematical algorithms to pass from measure to decision; actuator technologies and automated machinery or robots to transform decision into action as well as cloud-based data and information systems on a local and more regional basis, for co-operatives to support each other and for large agricultural enterprises to provide a range of services.

From "real-time" to long-term optimisation, all methods and technologies are relevant in this context to achieve the common goal of sustainable and productive farming at high standards of food safety, quality and traceability.

Smart farming solutions can also contribute where labour is the bottleneck of the system. This occurs in developed countries where there is a shortage of labour resources in the agricultural sector coupled with a huge need for more intensive production to satisfy food demand. Labour is also an issue in developed countries for many farming activities, particularly for small farms or marginal farming with fields that are dispersed or difficult to access. Specialised workers are needed for specific agricultural production in areas that are difficult to manipulate, e.g., remote or steep pastures in alpine areas, or for high demand crops (e.g., vineyards) and organic production. In all of these contexts labour can be automated by (almost) autonomous machines and other means within the precision farming paradigm. Last but not least, security issues have to be tackled to achieve integrity, guarantee availability of data and maintain communication, and to preserve privacy and IPRs of producers, users and data providers.

The European Commission, DG Connect, well aware of the rising importance of internet of things, has put considerable effort into digitising our industry and the internet of things, and in 2015 initiated AIOTI, Alliance for Internet of Things Innovation, as a European portal and initiative to support the creation of large-scale pilots (LSP projects) in this area. AIOTI provided 12 documents in different areas of interest, including the working group (WG06) on "Smart farming and Food Safety" [1], which was an important input for the Horizon 2020 Call on IoT large-scale pilots.

One of the first areas identified by the EC as being worthy of funding and support was "Smart Farming and Food Security" (Pilot 2). The description of this LSP was:

- Allow monitoring and control of plant and animal products from farm to fork
- Help farmers' decision making with regards to inputs and management processes
- Design architectures to "program" objects for optimal behaviour
- Enable consumers to access traceability information throughout the whole food chain.

One LSP on "Smart Farming and Food Security" was granted, called IoF2020, Internet of Food and Farm 2020 (30 million euros over four years), with 73 partners, coordinated by Wageningen University. In the second half of 2018, IoF2020 will have an open call of five million euros for new partners [3].

In the meantime, AIOTI has become an independent association with almost 200 members (Sept. 2016), covering

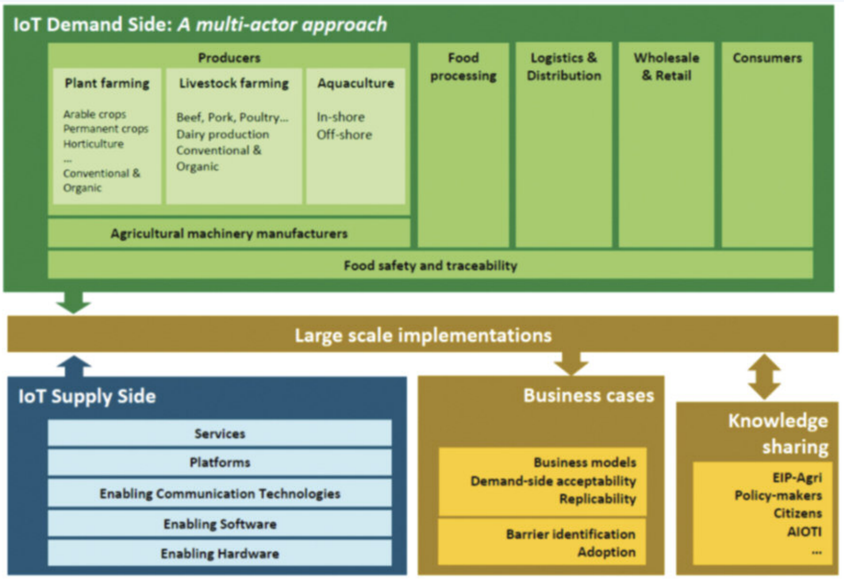



*Figure 1: Stakeholders of the Smart Farming Multi-Actor Approach (from AIOTI, WG06, [1])*

many sectors in 13 working groups [1]. AIOTI WG 06, Smart farming and Food Safety, focuses on the following key objectives, linked to other WG topics as well, because the issues addressed by smart farming are manifold and quite diverse [1]:

- *Focus on efficiencies across the 'from farm to fork' chain:* cropping, livestock farming, food processing and food distribution are all parts of the value chain to deliver products to the final consumer. This includes reducing food loss by improved production and distribution.
- *Focus on (livestock) farming and environment.* The environmental impact of the livestock sector is large and includes greenhouse emissions, waste, manure and inefficient feeding. Data-driven smart farming and synergies in the distribution chain can contribute to more sustainable production.
- *Focus on agriculture and water:* Irrigated agriculture currently accounts for 70 percent of world water withdrawals or pollution, and particularly in regions of southern Europe, dramatic improvements in agricultural water use are necessary.

Besides the big initiatives mentioned above, there are already a large number of national and European projects active in this area; a selection of these can be found referenced in the articles of this issue, but there are many more, such as those referenced in the AIOTI publications and the European Cordis web site (e.g., AfarCloud, an Horizon 2020 ECSEL JU smart farming project). The multiversity of subtopics and issues is quite well represented in the articles submitted and presented in this special theme.

*Precision farming* is looked at by several articles from different viewpoints:

- Building an IoT infrastructure for precision farming,
- Achieving sustainable soil management in conditions of high spatial variability,
- Data information and decision support systems to support farmers and other stakeholders as a service (monitoring prescription maps in one case, building a knowledge centre based on IoT and big data technologies or by applying mathematical methods - models and algorithms - for complex decision making to a range of farming types in two other cases), or
- by describing holistic approaches as concept (from automated machinery to the cloud) or implementation (by pioneering farmers in Tuscany, under support of a university and a regional agricultural consortium).

*Livestock management* is described from a dairy management point of view (insemination decisions, for instance) and the contactless monitoring of animal growth parameters by optical methods for health and productivity reasons, as decision support to farmers.

Field monitoring is performed by:

- drone swarms to detect and fight different kinds of weeds,
- swarms of intra-canopy sensors, which collect data for disease prediction models,
- remote sensors on board of satellite and aerial platform, which are exploited in an agricultural related application, to acquire images as a fundamental source for preservation of natural habitats (bio-diversity conservation).

*Data science* ("big data"), reasoning and decision support for the agricultural sector and the research community is the topic of three articles:

- Forecasting allocation of biomass crops (as opposed to food crops) for planning non-food farming activities,
- virtual research environments
- methods for massive high throughput plant phenotyping for the agriculture and food research communities.

Other challenges discussed are security issues (analysis of security threats and risks of IoT services for smart farming, data integrity and availability), a new method (3D X-ray imaging) for rapid food quality inspection, and an engineering solution to exploit biogas residues as a source of nitrogen fertiliser combining energy production and reduction of agricultural pollution.

Of course, topics overlap. If we look at technologies rather than farming sectors and applications, we find contributions related to sensors, remote sensing, modelling, machine learning, IoT, ICT methods, cloud and big data related.

This demonstrates the diversity of challenges and issues to be tackled in smart farming and food safety (security), and highlights the importance of adopting a holistic approach when it comes to practical implementations.

**Links:**
[L1] https://aioti.eu
[L2] https://kwz.me/hbD
[L3] https://kwz.me/hbF
[L4] https://www.iof2020.eu/about

**Please contact:**
Erwin Schoitsch, AIT Austrian Institute of Technology, Austria,
erwin.schoitsch@ait.ac.at

Mirco Boschetti, IREA-CNR, Italy,
boschetti.m@irea.cnr.it

# AmI Garden: Building an IoT Infrastructure for Precision Agriculture

by Manousos Bouloukakis, Christos Stratakis and Constantine Stephanidis (ICS-FORTH)

***A unique implementation of precision agriculture in an experimental greenhouse in Crete, Greece, utilises software driven automations to boost greenhouse vegetable cultivation. The system is based on agricultural scenarios for each individual species that grows inside the greenhouse, consisting of the environmental parameters that each vegetable needs for an optimum life cycle, as well as machine learning techniques that allow the system to predict and provide information about possible hazards.***

The Food and Agriculture Organisation (FAO) predicts that the Earth's population will reach 9.6 billion by 2050, which will require a 70 % increase in food production. At the same time, two main challenges to livestock and plant production include: climate change, which is the main factor contributing to altered life cycles of plants and animals; and plant diseases, which affect the majority of food crops, reducing food production and quality. Research results have shown production losses (42% average, by comparing primary and secondary yield losses resulting from a variety of negative factors, such as plant diseases, irrigation and fertilisation misplanning, etc) on the six most important food crops worldwide. Research on precision agriculture technologies and techniques in the context of smart farming has the potential to provide technological solutions to address some of these issues.

ICS-FORTH has recently initiated AmI-Garden, a smart farming project in the framework of its Ambient Intelligence Research Programme [L1]. A small experimental IoT greenhouse has been constructed and equipped with polycarbonate cover sheets and all the necessary infrastructure and hardware (automatic window-roof opening/closing, sliding door, fan installation for heating/cooling, vegetable breeding lamps etc.). Inside the greenhouse, a network of wireless sensors is used to measure environmental conditions and parameters, such as air/soil temperature and moisture, sunlight level, soil conductivity, quality and level of chemical ions in irrigation water, etc. The sensors communicate through IoT gateways to the greenhouse's data centre for storage and post-processing.

The system comes with pre-installed agricultural scenarios, a set of activity flows based on environmental conditions that are ideal for each plant species and are monitored in the greenhouse as explained above. The scenarios currently contain parameters to predict common diseases of the plants, as well as unexpected changes in the greenhouse's microclimate. For example, the irrigation process is built as an agricultural scenario using data from current plant status and past data in order to establish the optimal amount of water to irrigate. The parameters of this scenario are based on specific plant breed and environmental variables.

The intelligence behind the scenarios is based on critical limits and thresholds to create cultivation rules. On top of this rule based process, event-driven activation of various automations in the greenhouse is provided, for example, automatic humidity/temperature control, soil fertilisation (hydro fusion) and precise irrigation. Various sets of raw data are produced and ingested into the system, as the life cycle of each one of the plants evolves, in order to be used as the main input for the system's actuations based on the agricultural treatment scenarios.

AmI Garden applies machine learning techniques to data streams such as temperature, moisture, humidity etc. Data interpretation and data mining tech-

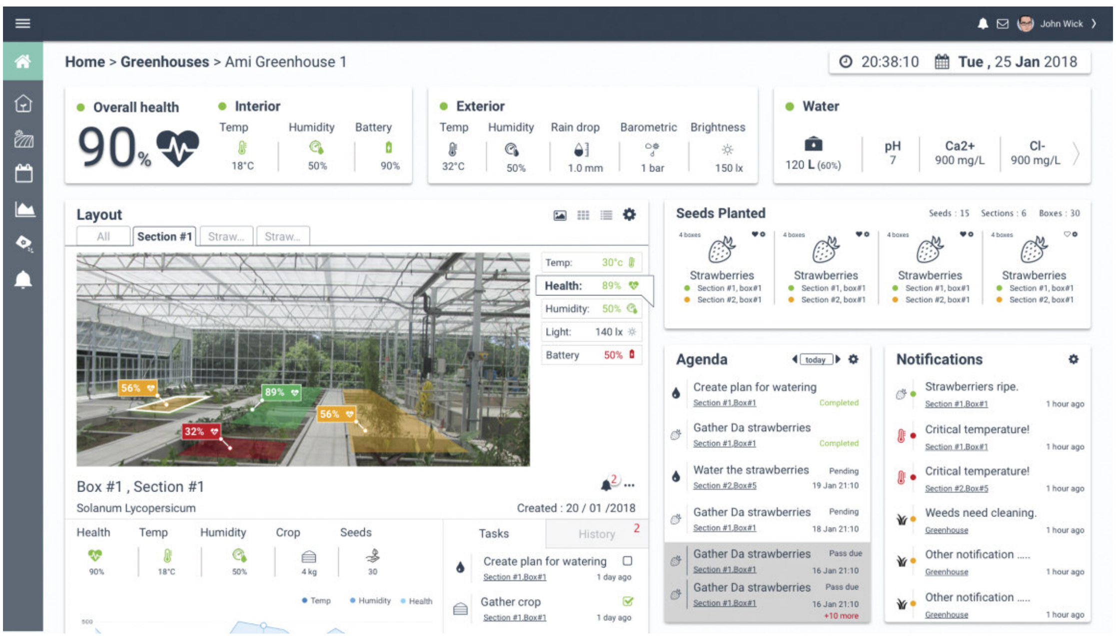



*Figure 1: Greenhouse Dashboard.*

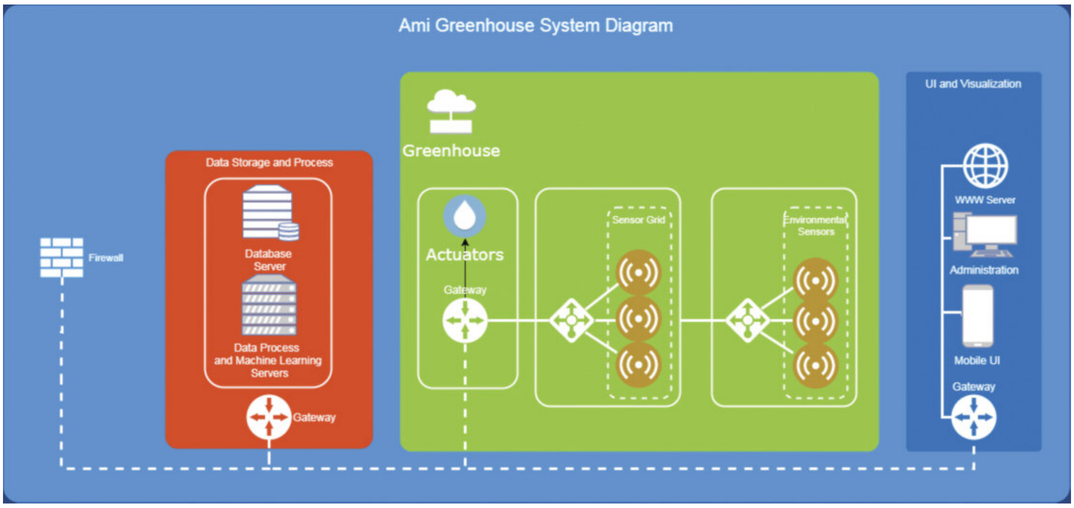



*Figure 2: Ami Garden Diagram.*

niques are employed to detect patterns and identify potential problems or opportunities for better plant growth (for example, specific values for moisture over a long time period can cause harm). Problem diagnosis techniques are employed to recognise characteristic symptoms (sensor communication/calibration problems, abnormal leaf wetness, water quality issues etc.) in order to develop and confirm hypotheses about possible causes.

AmI Garden's goal is to optimise cultivation processes and the use of resources as well as the efficient use of land. IoT technologies can increase production but can also enhance the quality of agricultural products. Currently, we are in the phase of collecting, storing, processing and visualising large amounts of field measurements in order to have enough data to ingest into the process engine. After that, data modelling will be used in order to further improve the predefined agricultural scenarios and train the machine learning algorithms.

The results of the initial tests that are currently being conducted, will be expanded on a large (open air) land field and larger greenhouses in order to verify the system's scalability and ensure a strong deployment potential.

**Link:**
[L1] http://ami.ics.forth.gr/

**References:**
[1] S. Wolfert, et al.: "Big Data in Smart Farming – A review", Agricultural Systems, 153, 69-80. doi:10.1016/j.agsy.2017.01.023
[2] Agriculture. Special Issue "The Impact of Plant Disease on Food Security", 2012. Retrieved from https://kwz.me/hbG
[3] U. Cortes., et al.: "Artificial Intelligence and Environmental Decision Support Systems", Applied Intelligence 13, 77–91, 2000.

**Please contact:**
Constantine Stephanidis
ICS-FORTH, Greece
+30 2810 391741
cs@ics.forth.gr

# Smart Farming: From Automated Machinery to the Cloud

by Christoph Schmittner (AIT), Christian Hirsch (TU Wien) and Ma Zhendong (AVL List GmbH)

***Smart Farming – the application of IoT and Industry4.0 technologies within the agricultural domain – has the potential to increase yield and efficiency while decreasing environmental impact. A European project will address the challenges necessary to reach these goals. The results will be demonstrated in multiple real world agricultural demonstrations.***

The agricultural domain is facing new and increased challenges, such as an increasing labor shortage due to the movement from rural areas to cities and the declining attractiveness of agriculture jobs. Simultaneously, there are pressures to increase productivity and cost-effectiveness to supply a rising world population while minimising the use of chemicals and adapting to changing climatic conditions [1]. Future agriculture needs to adopt and expand production processes, technologies and tools – a challenge that will be met through research and development activities from all domains. There are already first approaches: precision farming and automation have already increased productivity and quality while reducing the need for manual labor. The use of high-tech, e.g., digitally-controlled farm implements, and even unmanned aerial vehicles for monitoring and forecasting is also on the rise. Most current solutions address singular tasks. An important aspect for the future is the interoperability and configurability of these systems.

"Smart Farming" applies and combines these technologies with approaches from industry4.0 and smart mobility to address the challenges and develop a holistic system. AFarCloud is an ECSEL JU Project, starting in the second half of 2018 which will enable and extend the usage of smart farming by combining experts from industry4.0, smart mobility and Internet of Things (IoT) with agricultural solution provider and end user.

IoT sensor networks in the field will detect pests and environmental conditions and a range of other factors affecting plant health. Communicating this information via secured edge nodes with the AfarCloud middleware, and storing them in a backend, enables a fast reaction and adaptation and simultaneously charters and monitors long term trends and compares production and conditions across years. Automated vehicles compensate for the shrinking labor force and enable precision farming by closing the feedback loop between sensing, control and actuation. Automated farming vehicles, working in combination with a distributed sensor network can apply pesticides and fertiliser based on the measured needs of small patches of the complete agricultural production unit. To enable the future of farming, Afarcloud will work on a distributed framework for autonomous farming that will allow the integration and cooperation of agriculture cyber physical systems in real-time in order to increase efficiency, productivity, food quality and reduce farm labor costs. An important aspect of this approach is the holistic view of farming and food production, including production, energy, food quality and services. Taking the complete value chain, from farming equipment to the end-product into account enables a complete approach.

Automated farming vehicles should become usable by farmers and cooperate to enable new applications as basic vehicles can cooperate to address more complex tasks without human interaction. The framework will enable the seamless interaction of multiple devices and systems towards a cooperative system of systems.

A prerequisite is the safety and security of all involved systems [2]. Automated farming vehicles face challenges that only partially overlap with smart mobility, but safe operation and secure communication and cooperation are a necessity for the practical application in both domains. Additionally, sensor networks in the agricultural domain must cope with greater demands on their robustness against environmental and mechanical damages while ensuring the same level of security as sensor networks in less challenging areas. Food security is a critical area to which increased connectivity and automation introduce multiple challenges. There exists the potential for both malicious attackers and human misuse or failures to threaten food supply.

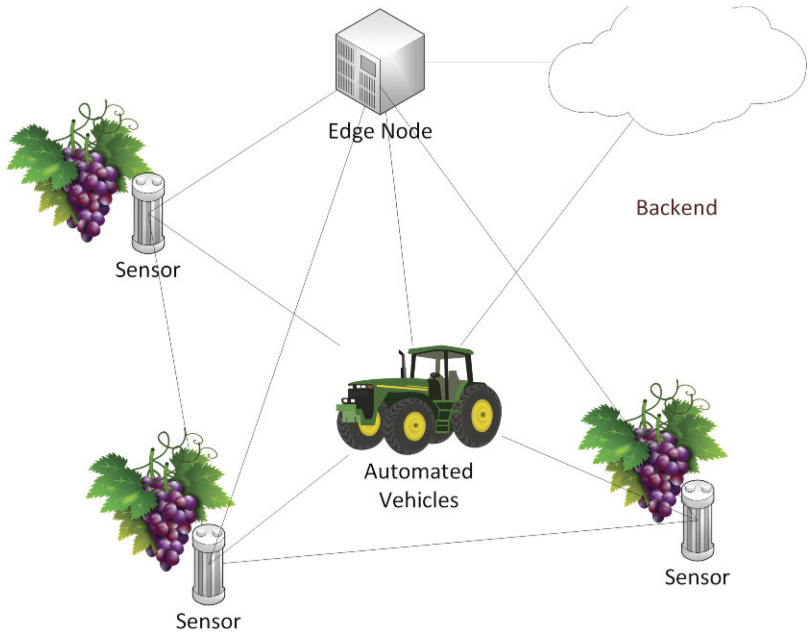



*Figure 1: Example UseCase. An overview of an intended scenario.*

AfarCloud will build upon and extend existing IoT frameworks and standards from multiple domains to develop a holistic, safe and secure agricultural framework [3]. For example, the Data Distribution Service for Real-Time Systems (DDS) standard by OMG will be used to enable reliable, scalable and real-time data exchanges using a publish/subscribe pattern and automotive safety and security standards will be utilised for safety and security of automated farming machines.

AFarCloud outcomes will strengthen the European agricultural domain and support food security by applying ICT-based solutions to farming. Partners throughout the complete value chain, including all key actors needed for the development, demonstration and future market uptake of the precision farming framework are working on the project and supporting a leadership role for Europe in the smart farming market. The results of the project will be demonstrated and evaluated in multiple real-world agricultural demonstration sites, ranging from livestock farming to crop farming.

Acknowledgment
"AfarCloud is partially funded by the EC Horizon 2020 Programme, ECSEL JU, and the partners National Funding Authorities under grant agreement # 783221"

**Link:**
https://www.ait.ac.at/en/research-fields/dependable-systems-engineering/

**References:**

[1] G. Feder, A. Willett, and W. Zijp: "Agricultural extension: Generic challenges and the ingredients for solutions," in Knowledge generation and technical change, Springer, 2001, pp. 313–353.

[2] C. Schmittner, Z. Ma, and T. Gruber, "Standardization Challenges for Safety and Security of Connected, Automated and Intelligent Vehicles," presented at the The 3rd International Conference on Connected Vehicles & Expo (ICCVE 2014), Wien, 2014.

[3] R. Khan, S. U. Khan, R. Zaheer, and S. Khan, "Future Internet: The Internet of Things Architecture, Possible Applications and Key Challenges," 2012, pp. 257–260.

**Please contact:**

Christoph Schmittner, Erwin Schoitsch
Austrian Institute of Technology, Austria
christoph.schmittner@ait.ac.at,
erwin.schoitsch@ait.ac.at

# Adopting Precision Agriculture Strategies to Survey the Effect of Tillage Systems on Crop Growth Under High Spatial Variability Soil Conditions

by Marco Sozzi, Donato Cillis, Francesco Marinello and Luigi Sartori (University of Padova, Italy)

***Precision agriculture tools can be of benefit to low-intensity tillage systems. This study investigates spatial variability in a conservative tillage system using proximal sensors and satellites.***

Efficient and reliable methods for measuring spatial variation in soil properties are fundamental to precision farming. Over the last decade, geophysical sensors based on non-destructive measurement of soil electrical conductivity (or its inverse, resistivity) have been extensively used in precision agriculture. Different potential patterns provide information on subsurface and its electrical properties. The electrical resistivity of the soil can be considered as a proxy for the variability of a soil's physical features including texture, type and moisture. At the same time, satellite and airborne sensors offer an increasing amount of information about soil properties and ground cover. In particular, vegetation indices, based on remotely-sensed spectral reflectance in the near-infrared and visible bands, have been widely used for monitoring vegetation [1]. The study investigates the potential to monitor the evolution of the vegetation index to modulate the agricultural operations on soft wheat (Triticum æstivum L.).

In the present work a survey on soil and vegetation variability, coupling a multi-depth automatic resistivity profiler (ARP©, Geocarta, France) [2] and Normalized Difference Vegetation Index (NDVI) as derived from Sentinel-2 imagery data. The analysis was based on a relationship between resistivity and NDVI index. The study was carried out in the 2015-2016 season in a farm located in Caorle (VE) – Veneto, Italy (45.63°N 12.95°E). Wheat cultivation was managed with three different soil tillage systems, characterised by different tillage intensity: conventional (CT), minimum (MT) and no-tillage (NT). Each plot was 1.5 ha in size. The ARP system was pulled through the field to collect data at three different depths simultaneously (0-0.5 m, 0-1 m and 0-2 m). Data were real-time referenced by differential global positioning system (DGPS). On the other hand, NDVI vegetation index was assessed using Sentinel-2 mission's satellite images. To obtain information during the representative stages of the crop cycle, five satellite images, from 20 January 2016 to 30 March 2016, were considered. Images were chosen selecting only low noise scenes, excluding those affected by clouds or fog. Soil spatial variability was analysed processing three depth levels ARP data using the statistical software MZA (Management Zone Analyst, University of Missouri-Columbia [3]), which implements a fuzzy c-means unsupervised clustering algorithm that assigns field information into like classes. To perform the analysis, NDVI and ARP data were re-sampled building a 10×10 m reference grid. Finally, soil resistivity and vegetation index maps were correlated dividing ARP values into classes representing step values of resistivity of 5 Ohm·m, falling in the reference management zone managed by different tillage systems.

As a result, four homogeneous classes were identified after the MZA analysis: three of them cover all the tillage sys-

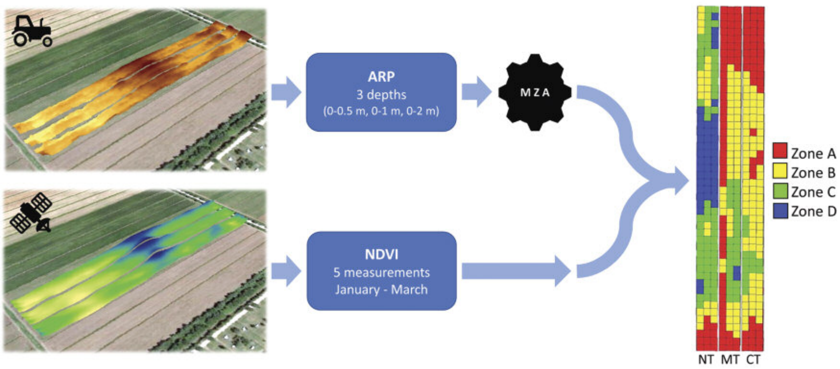



*Figure 1: Summary of the applied process.*

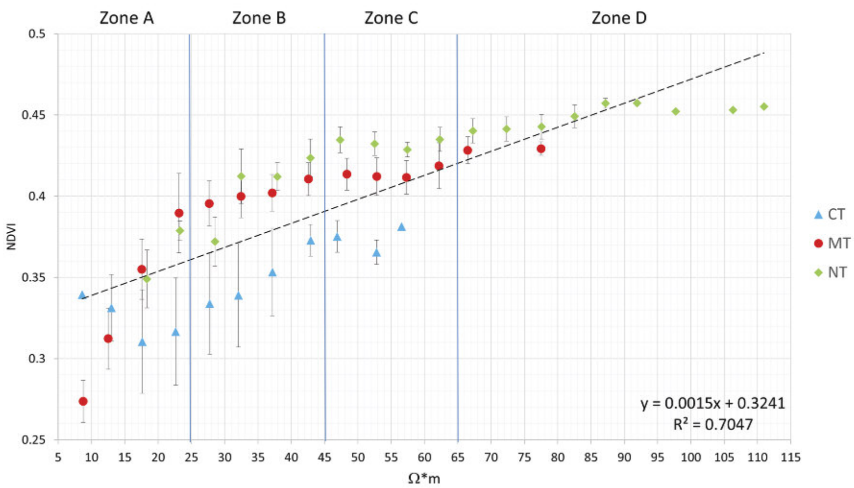



*Figure 2: Correlation between ARP and NDVI for tillage system and management zone.*

tems, while the fourth zone affects only the side of the field managed with MT and NT. Such homogenous zones (namely A, B, C and D) are characterised by increasing ARP values. Soil resistivity showed high variability in terms of soil features, in fact, a wide range of resistivity was observed. High resistivity levels characterise sandy soil with high water drainage (zone D), while low resistivity levels are typical of clay soil with high salinity (zone A) that can occur in a stagnant condition if affected by soil compaction. Besides, to increase the confidence and the level of information on the experimental area, the average of the five dates' NDVI data was taken into consideration. A clear correlation can be detected between soil features and vegetation vigour, as reported in Figure 2. Increases in ARP value correspond to high vegetation vigour, expressed by NDVI. In addition, it is possible to observe the NDVI temporal trend of the single homogeneous zones under different soil tillage systems. Correspondingly, the contribution of the soil tillage technique on NDVI index and the influence on soil features can be investigated. CT shows lower NDVI than MT and NT for all homogeneous zones. This is caused by its features: indeed, CT tends to cool soil faster than MT and NT. In addition, continuous inversion of soil layer leads to a decrease in soil fertility and nutrients availability. Finally, the higher number of passages distinguishing CT for the seedbed preparation could affect soil compaction leading to a stress condition for subsequent plant growth, especially in the field characterised by clay soil. Conversely, MT and NT do not involve inversion of soil layer and require a lower number of passages across the field. Besides, their features lead to a mitigation of soil cooling phenomena allowing fast seed germination. Zones with low resistivity are characterised by high variability of the NDVI measurements, due to the cooling phenomena at the beginning of the crop cycle.

This study has been performed under the framework of the AGRICARE project - Introducing Innovative Precision Farming Techniques in Agriculture to Decrease Carbon Emissions, supported by the European Commission – DG Environment through the LIFE+ programme; under contract number LIFE13 ENV / IT / 000583.

**References:**

[1] P. Gong et al.: "Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM+ data", Int. J. Remote Sens., 2013.
[2] D. Cillis, et al.: "Field-scale electrical resistivity profiling mapping for delineating soil condition in a nitrate vulnerable zone", Appl. Soil Ecol., 2017.
[3] J. J. Fridgen, et al.: "Management Zone Analyst (MZA)", Agron. J., 2004.

**Please contact:**

Marco Sozzi, TeSAF Department, University of Padova, e-mail: marco.sozzi@phd.unipd.it

# Agrodat: A Knowledge Centre and Decision Support System for Precision Farming Based on IoT and Big Data Technologies

by Róbert Lovas (MTA SZTAKI), Krisztián Koplányi (eNET) and Gábor Élő (Széchenyi István University)

***Since 2014 the Agrodat project and its collaborating partners have been working on introducing new, cost-effective sensor technologies and advanced ICT solutions for the Hungarian agriculture sector. Currently, 50 farmers and other targeted user communities have access to data analytics services with automatic warnings in case of hazardous conditions.***

Agricultural production always bears a high degree of uncertainty because yield can fluctuate significantly depending on weather conditions, pests, irrigation and fertilisation, but the market and political factors also considerably influence the profitability of farming. With the help of precision farming, as a special method of crop management, many production factors can be precisely tracked, thereby reducing the risk. Additionally, the cost of producing the crop and the environmental impacts might be significantly reduced, since this new approach enables farmers to make decisions about variables such as: which areas of land/crop within a field can be managed with reduced levels of fertiliser, chemicals, and irrigation water, depending on particular factors, like the yield potential of the crop in the given area.

In general, to implement precision farming and related methods, a vast amount of data about the cultivated fields must be collected from various sources, and analysed. Nationwide projects, such as Agrodat from Hungary, require even more sophisticated IoT and big data solutions starting from the sensor level, through to the applied communication network, and ending with the advanced data analytics facilities and knowledge centre.

The AgroDat project [L1] (launched in January 2014) is a collaboration between HP Enterprise, eNET Internet Research and Consulting Ltd., Institute of Computer Science and Control of the Hungarian Academy of Sciences (MTA SZTAKI), and Széchenyi István University. The project's aim was to create a knowledge centre for precision agriculture based on both local sensor data and semi- or unstructured data from international repositories.

By December 2017, at the end of the project, nearly 1,000 complex sensor pillars were deployed at selected locations covering more than 3,000 hectares


*Figure 1: High-level concept of Agrodat: data sources + Big Data centre + decision support system with knowledge base/multilingual search engine.*

on the properties of 50 farmers in Hungary. The sensor pillars, which we developed, have a modular structure and are equipped with anti-theft solutions (GPS, tilt sensor) as well as facilities to measure: wind speed and direction; precipitation; air temperature (at 20cm/2m) and relative humidity; solar radiation and reflection; leaf wetness; and soil moisture, temperature and conductivity at various levels. The IoT communication network is based on the 4G mobile network, and M2M communication enabled SIM cards. [1]

In order to provide an experimental platform for developers, researchers, and other involved parties, a new big data centre was established in 2015 and operated by MTA SZTAKI with hierarchical storage and advanced databases (e.g., Cassandra and HPE Vertica); GPGPU cluster for processing raw data; Apache Hadoop and Spark clusters for analytics, etc.

An OpenStack on-premise cloud is responsible for providing an elastic and flexible framework for the research platform to facilitate the development of higher-level new software layers including the cloud orchestrator tool for complex services based on Occopus [2], the IoT back-end for data collection [3], the Flowbster [4] cloud-oriented workflow system to process large scientific data sets , and the various modules of decision support / visualisation systems.

The aggregation of related scientific and other semi- or non-structured data from international repositories leverages partly on the experiences gained during the elaboration of the integrated workflow-oriented services in the former EU FP7 agINFRA project. The new Agrodat search tool, on the one hand, takes the advantage of the HPE IDOL (Intelligent Data Operating Layer) enterprise search engine that offers advanced functionalities for conceptual and contextual analysis of such structured, semi-structured and unstructured data. On the other hand, the new tool integrates the AGROVOC multilingual agricultural thesaurus that consists of over 34,000 concepts available in 29 languages and published by the Food and Agriculture Organisation. In this way, the tool enables data managers, researchers, and the broader user community to access dozens of international repositories (registered in the CIARD RING, http://ring.ciard.net) and also to use them for data exploration and analytical purposes. As one of the latest developments, the Agrodat platform is able to calculate and visualise NDVI (Normalised Difference Vegetation Index) from Sentinel 2 multispectral images.

The targeted end-users include farmers, experts, service companies in the agriculture sector, insurance companies, universities and research communities. The most welcome functionalities of the decision support system by the farmers are the visualisation of the collected data and the automatic issuing of various pest control warnings in case of special weather conditions that make it possible for the farmers to intervene (e.g. preventive steps) in a timely and targeted fashion, which increases the expected crop yield.

**Link:**
[L1] www.agrodat.hu

**References:**
[1] G. Paller, P. Szármes, G. Élő: "Power Consumption Considerations of GSM-connected Sensors in the AgroDat.hu Sensor Network", Sensors and Transducers, 189:6 (2015)
[2] J. Kovács, P. Kacsuk: "Occopus: a Multi-Cloud Orchestrator to Deploy and Manage Complex Scientific Infrastructures", Journal of Grid Computing, 16:1 (2018)
[3] A. Csaba Marosi, A. Farkas, R. Lovas: "An adaptive cloud-based IoT back-end architecture and its applications", PDP 2018 conference (2018)
[4] P. Kacsuk, J. Kovács, Z. Farkas: "The Flowbster Cloud-Oriented Workflow System to Process Large Scientific Data Sets", Journal of Grid Computing, 16:1 (2018)

**Please contact:**
Zsolt Büki, eNET Ltd., Hungary
+36 30 902 8846 info@agrodat.hu

# Information Systems for Precision Agriculture: Monitoring Computation of Prescription Maps

by Simone Lanucara, Cristiano Fugazza, Paolo Tagliolato and Alessandro Oggioni (CNR-IREA)

***Implementing a semantics-aware dashboard allowing for automated flagging of missing data.***

Precision agriculture (PA) is increasingly becoming a data-intensive discipline characterised by the collection and processing of digital data whose level of granularity is augmenting day by day. On the one hand, individual in-the-field activities are becoming increasingly accurately traced (e.g., via telemetry); on the other, digitally collected information can be exploited by in-silico models in order to suggest actions to farmers. In this scenario, data management must be capable of considering heterogeneous sources, either site-specific information directly collected by farmer's devices or derived from these and external data sources, such as meteorological data and satellite data. Major issues in the advancement of PA include articulating and understanding the huge amount of data that is collected and implementing sound strategies for field management. However, most data and information sources are fragmented, encoded in different formats, and awkward to find and use. It is then necessary to structure and harmonise information in order to check which data are available, what are they for, and their sources. Once this information is made available, it is possible to process data in order to obtain more refined data products.

The integration of heterogeneous data from multiple sources is major issue in the domain of spatial information technology. Different approaches (see [1] for a review of these) propose Service Oriented Architectures (SOA) to ease integration of heterogeneous and distributed sources for PA; as implemented in other domains (geologic, oceanographic, ecological), SOA enable web sharing of data and related metadata. In the context of the SATFARMING project, in order to foster interoperability, we are implementing a semantically enabled SOA based on Open Geospatial Consortium (OGC) standards. The SATFARMING project, funded by a private company, IBF Servizi S.p.A., is addressing the aforementioned issues in a comprehensive way. The SOA takes into account PA models, workflows, and processes through diverse software components, services, and tools accelerating integration of data-sources, and their processing, creating PA data products and sharing them.

In SATFARMING, in addition to allowing for harmonisation, management, identification, and sharing of multi-temporal, multi-scale, and external data sources, we formalise the flows leading to PA products in terms of the necessary data and processes, enabling automated flagging of missing inputs in a user-friendly web-based dashboard. The first application scenario is constituted by prescription maps, PA products specifying the actions (in terms of what, where, and when) the farmer shall execute on a given field. Figure 1 depicts the transitions leading from abstract flows and heterogeneous data sources to the synoptic dashboard pinpointing the missing components for processing the data product constituting prescription maps. On the one hand, the abstract flow is translated into a digital representation that is amenable to processing (upper-left side); on the other, the heterogeneous data sources are collected and

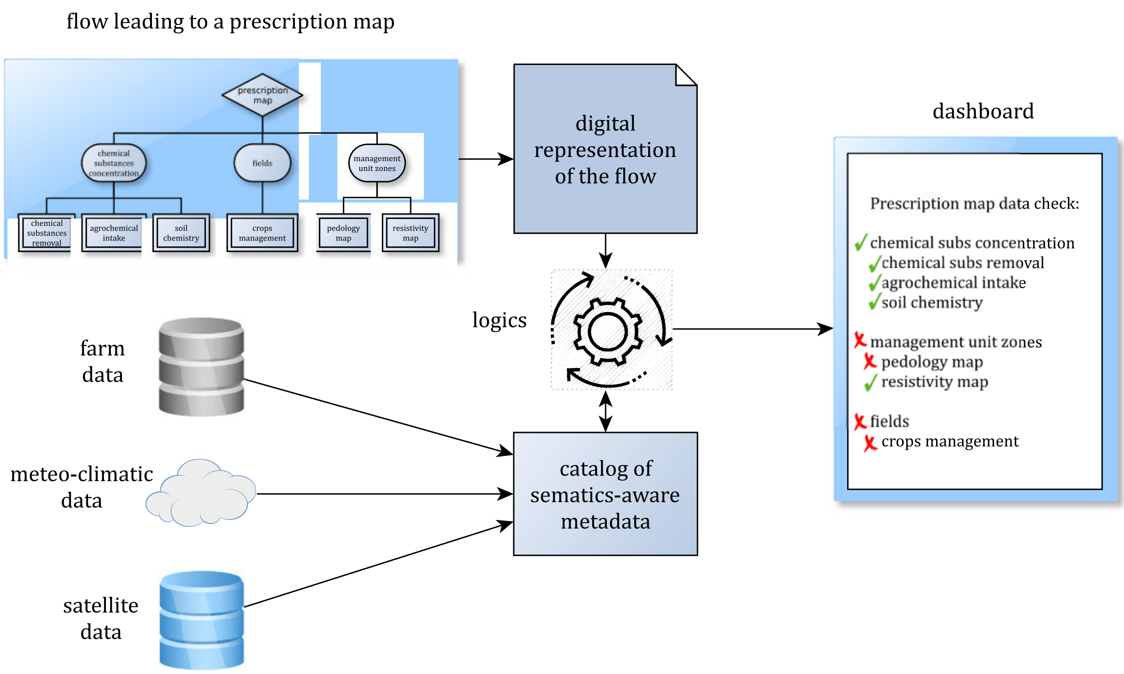



*Figure 1: Semantics-aware dashboard in the SATFARMING SOA.*

annotated (bottom-left side). Both these activities are supported by the controlled vocabulary that allows their representations to be semantically enriched. The added semantics allows logics to process these data structures in order to cross-reference them and compute the dashboard on the right-hand side of the Figure.

The dashboard presented here is implemented as a component of the SOA being developed in the SATFARMING project; hence it is readily available via the mechanisms defined by SOA. As an example, the component could be exploited for dissemination of PA knowledge (models, workflows and products) from PA-savvy agronomists to agronomists in developing countries who do not have first-hand experience of these techniques. Agronomists with expertise in this area may define the flow leading to a given PA product so that an inexperienced agronomist can understand, by using the dashboard, which data and processing are required to achieve the final PA product. This is in line with the United Nations 2030 Agenda on Zero Hunger Goal which aims to: 1) ensure sustainable food production systems, 2) implement resilient agricultural practices that increase production, and 3) improve agricultural productivity of small-scale food producers.


**References:**

[1] Janssen et al.:"Towards a New Generation of Agricultural System Data, Models and Knowledge Products: Information and Communication Technology." Agricultural Systems 155. 2017



**Please contact:**

Simone Lanucara, CNR-IREA, Italy
+39 3496082982
lanucara.s@irea.cnr.it


# Integrating Agronomic Knowledge into Precision Agriculture


by Marco Napoli, Anna Dalla Marta, Marco Mancini and Simone Orlandini (University of Florence)



***Technology is increasingly being used to make farming processes more efficient. We are integrating agronomic knowledge and precision farming techniques in order to achieve an adequate agronomic planning at the farm or consortium level and then to transfer this knowledge to farmers, facilitators and farm advisors.***


Modern agriculture is based on the control of in-field variability, which is determined by the interactions of numerous factors such as soil, climate and crop. However, growing areas in Italy are often characterised by a high heterogeneity in pedological, orographic and climatic conditions. This variability can be seen at a small scale, within fields of a few hectares, which results in highly variable grain yields and quality, posing a challenge for agronomic planning at the farm or consortium level. Furthermore, some limitations (e.g., too shallow soil or clayey soils) cannot be amended for technical or economic reasons and in these cases, it would be appropriate to limit inputs in function of the expected potential yield. Nevertheless, sometimes in an effort to improve the quantity and quality of yield in low productivity areas, farmers increase the chemical and energetic input rate without a real agronomic or economic justification. Therefore, understanding the causes and the spatial extent of poor performing areas, of agronomic options and of potential production can provide farmers with tools to evaluate and implement appropriate management techniques.

Our study focusses on wheat and the area is located between the Val d'Orcia and the Val d'Arbia, two rural areas of Tuscany (Central Italy), characterised by a typical Mediterranean climate (13.6 °C and 715 mm of annual average temperature and cumulated rainfall, respectively) and fine-textured soil. The "Department of Agrifood Production and Environmental Sciences of the University of Florence (DISPAA)" and the "Siena Provincial Agricultural Consortium (CAPSI)" have been carrying out these activities since 2009, as part of the projects "Precision farming for the Tuscany pasta producers (APPCoT)" and "Old wheat varieties – New cultivation techniques (GRANT)" which were partially funded by the Rural Development Plan of Tuscan Region 2007–2013 and 2014-2020, respectively.

In Central Italy, wheat is usually fertilised by two applications of nitrogen (N) as a function of vegetation vigour and expected yield potential of the field; one following tilling and the second just before the end of stem elongation. Owing to the in-field variability, the potential patterns of yield, as observed over several seasons, cannot reliably drive N fertilisation. On the contrary, crop vigour monitoring allows the crop yield potential to be predicted for each sub-area, thus enabling a site-specific distribution of the fertiliser with positive implications for both economic and environmental sustainability.

For this reason, some Tuscan farmers, with a pioneering spirit, wired and equipped their combine harvesters and tractors with Trimble GPS sensors and computers for parallel driving. Moreover, a Real Time Kinematic (RTK) positioning was installed to enhance the precision of position data derived from satellite-based positioning systems. Combines were further equipped with yield monitoring sensors to measure and record information such as grain flow and moisture, area covered and location (Figure 1). Kuhn fertiliser spreaders, fitted with an electronic control box and actuators controlling the outlets, were adopted, so the fertiliser flow is automatically adjusted depending on the forward speed. DISPAA calculated vegetation indices, based on the RapidEye satellite images (resolution of 5×5 m), which

*Figure 1: Combines equipped with yield monitoring sensors for measuring and recording information such as grain flow and moisture, area covered and location.*

carried a multi-spectral sensor recording radiance in five broad bands of the electromagnetic spectrum. Two satellite images were acquired annually at the end of wheat tilling and during stem elongation, for timing the first and the second top-dressed fertilisation, respectively [1]. Moreover, to produce prescription maps for variable-rate fertiliser application, the CERES-wheat model simulations were performed to identify the appropriate fertiliser rate to optimise grain yield [2] and the grain protein concentration [3] as function of both the field zonal characteristics and the seasonal meteorological forecasts. These prescription maps are then transferred to farmers through the CAPSI's farm advisors to distribute the fertiliser.

The tests carried out in recent years have shown that it is possible to optimise fertilisation to achieve good production results and at the same time reduce the economic and environmental costs associated with inappropriate distribution of fertilisers.

**References:**

[1] A. Dalla Marta, et al.: "Durum wheat in-field monitoring and early-yield prediction: Assessment of potential use of high resolution satellite imagery in a hilly area of Tuscany, Central Italy", Journal of Agricultural Science, 2015, 153(1), 68-77.

[2] A. Dalla Marta, et al.: "A simplified index for an early estimation of durum wheat yield in Tuscany (Central Italy)", Field Crops Research, 2015,170, 1-6.

[3] F. Orlando, et al.: "Modelling durum wheat (Triticum turgidum L. var. durum) grain protein concentration", Journal of Agricultural Science, 2017, 155, 930–938.

**Please contact:**

Marco Napoli

Department of agrifood production and environmental sciences of the University of Florence

Piazale delle Cascine 18, 50144, Florence, Italy

+39 0552755746

marco.napoli@unifi.it

# Farming with Mathematical Precision

by Simon van Mourik, Peter Groot Koerkamp and Eldert van Henten (Wageningen Universiteit)

***A key challenge in precision farming is complex decision making under variable and uncertain circumstances. A possible solution is offered by mathematical models and algorithms.***

The enormous quantities of fertiliser, pesticides, fresh water, and fossil fuels that have been used in modern farming over the last century have had an enormous impact on the local and global environment. Furthermore, resources such as fresh water, phosphate, oil, and affordable labour are getting scarce. This situation is not sustainable, especially given the rapid growth of the world population.

In our group we research precision farming, a farm management approach that uses information and precision technology to address these sustainability issues; firstly by automation to save on labour input, and secondly by precision management to save on resources, and reduce environmental impacts. The term precision in this context reflects accurate dosage of resources, the timing with which they are applied, and sometimes even how these dosages are allocated; for example feeding per individual animal or fertilising at specific field locations.

We apply the precision farming approach in a range of agricultural settings, for example in arable farming, greenhouse horticulture, and livestock farming. Well known applications are milking robots, automated irrigation, air conditioning, and individual cow feeding. However, most aspects of management are not yet automated or precise. Identifying what is the optimal dosage, timing, and allocation is complicated for a number of reasons. The response of crops, animals, indoor climate, or soil, towards application of food, fertiliser, water, and so on, is generally governed by a complex nonlinear response. Some processes respond quickly, such as indoor climate toward air conditioning, and some processes respond slowly, such as the development of crops or livestock. Hence, we deal with multiple time scales. On top of that come several sources of variation – differences between crops, animals, soil properties, weather changes, and disease occurrence to name a few. Our challenge is to design models and algorithms that come up with optimal management strategies under all of these circumstances.

An automated system consists of three key elements; sensing (by sensors or cameras), processing (decision making software based on sensor information and decision rules), and actuation (by operating valves or switches, by machines, or robots). An example of such as system is the climate inside a greenhouse. With sensors the temperature, humidity, light intensity, and carbon dioxide levels are measured, and based on these measurements a processor decides which windows or screens should be opened or closed, whether lights should be switched on, and how much heat should be pumped into the greenhouse. The interaction between the actuator, farming system, sensor, and processor, forms a feedback loop. This interaction resembles one of the core principles within systems and control theory – feedback control.

Control algorithms allow us to design decision rules that are optimal with respect to a chosen performance criterion. This criterion typically weighs costs of input against revenues of output, for example heating and lighting costs against the yield of greenhouse tomatoes. A control algorithm that works well in practice requires a mathematical model that describes all relevant inputs, dynamic states, and measured output of the system, as well as all relevant disturbances coming from outside such as weather changes or sensor noise. This is quite an exercise, but in the end it pays off. For example, we showed that optimal control of greenhouse air conditioning can potentially save up to 47% on heating energy [1].

A big challenge is introduced by sources of variation that are not understood or are hard to measure in practical situations. To unravel the roles of different sources of variation, we carry out uncertainty analyses. For example, during transport about 5-10% of consumption eggs get fractured, at considerable economic loss. We used a Bayesian uncertainty analysis to predict that variation in collisions has far more impact on fracture rates than variation in egg quality [2]. To deal with variation that cannot be explained or controlled, control algorithms are adapted by online parameter tuning based on sensor data. A promising application for this method is pest control that is adapted to specific fields, greenhouses, or as our research showed, barns [3]. Another promising application we are investigating is a controller that steers the timing of harvest per specific greenhouse and per crop type.

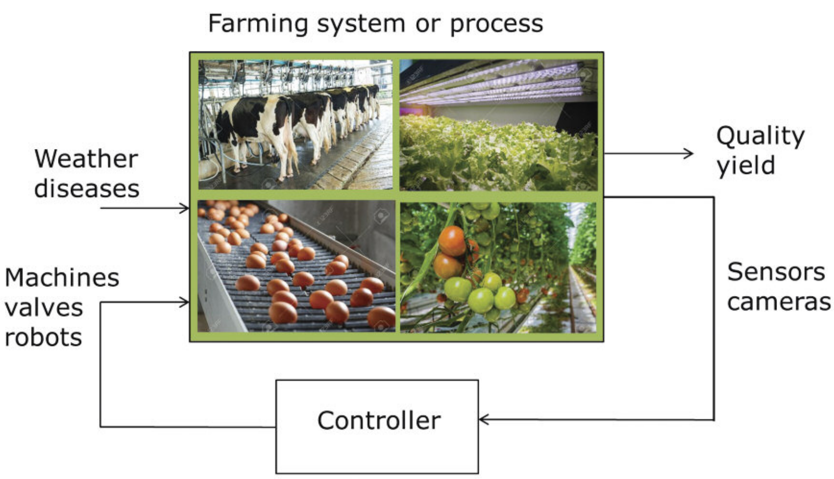



*Figure 1: The feedback control loop projected onto some farming processes (source photos: www.123rf.com)*

Even if full automatic control is not yet feasible, automated decision support in the form of a prognosis, or diagnosis, can already help a farmer considerably. For example, by generating advice for variable input rates of water and fertilizer depending on local field conditions, or by predicting what the consequence of additional carbon dioxide input in the greenhouse will be regarding yield and input costs. Some methods we apply to this end are model integration, state filtering (for dynamic systems), and machine learning algorithms (mostly for static systems).

For a fully automated, and precisely controlled farm, a lot of mathematical and engineering challenges still have to be addressed. For example, we need more accurate input-state-output models for crop and animal systems, and control algorithms that perform reliably under multiple disturbances (such as weather, model errors, variation, and sensor noise). Nevertheless, since the methods of systems and control theory apply so beautifully to the concepts of precision farming, we believe there is a world to gain here.

**Link**: www.fte.wur.nl

**References:**


[1] P. van Beveren, et al.: "Optimal control of greenhouse climate using minimal energy and grower defined bounds", Applied Energy, 2015. 159: p. 509-519.

[2] S. van Mourik, et al.: "Predicting hairline fractures in eggs of mature hens", Poultry Science, 2016.

[3] M.F. Mul, et al.: "Development of a model forecasting Dermanyssus gallinae's population dynamics for advancing Integrated Pest Management in laying hen facilities", Veterinary parasitology, 2017. 245: p. 128-140.


**Please contact:**


Simon van Mourik
Wageningen Universiteit, The Netherlands
+31 317 481275
simon.vanmourik@wur.nl

# Exploiting Low-cost Depth Cameras for Body Measurement in the Livestock Sector

by Andrea Pezzuolo, Luigi Sartori and Francesco Marinello (University of Padua, Italy)

***Animal dimensions play a key role in providing data in support of management decisions regarding livestock production. In the last fifty years, manual measurements have been the most common way to get an indication of animal growth. An alternative approach, which overcomes the limitations of manual direct measurements, is to use techniques based on optical detection instruments.***

Monitoring animal growth parameters in a frequent and quantitative way is useful to help the maintenance of animal health and to maximise production efficiency. In particular, the live weight is of great interest for breeding and management, as it serves as an index for animal growth, health and readiness for market. Over the last fifty years, the best way to measure individual animal's body live weight, as well as their physical development, has been through the use of traditional or electronic weighing systems. However, this approach is laborious and difficult, and it may be stressful for both the animals and the stockman. As a consequence, it can be very time consuming, expensive and may cause injury to the stockman [1].

An alternative approach, which overcomes the limitations of manual direct measurements, is with the introduction of techniques based on optical detection instruments. Optical systems are often used for agricultural and livestock study and applications. In different studies, different kinds of optical sensors (for example, 2D cameras) have been used, but there is a greater interest in 3D sensors, like TOF (Time-Of-Flight) camera or CTS (Consumer Triangulation Sensor) system [2].
Another important sensor that could be implemented is Microsoft Kinect™. Three-dimensional reconstruction by the Kinect camera is achieved by combining an infrared laser emission source with an infrared-sensitive camera. Since its launch in 2010, tens of millions of units of this sensor have been sold, and there is a growing interest in how its high quality three-dimensional imaging can be used in many applications including the agriculture and livestock sectors [3].

Some studies have already discussed the use of the Kinect sensor in livestock applications. However, there is a lack of information about metrological performance, and this is severely limiting the actual application of the method, since it is not clear to what extent Kinect 3D data can support or replace manual measurements, particularly when the animal's body is in different positions.
The present study proposes the implementation of the same low-cost Kinect v1 depth camera for fast, non-contact measurements of cow (Figure 1) and pig (Figure 2) body dimensions such as heart girth, length and height. Such body values can be directly related to the weight which can be thus estimated without moving or contacting the animal.

Research experiments compared estimated data with the results from manual measurements during the fattening period, highlighting relatively high coefficients of determination ($R^2$>0.9), in particular for parameters arising from three-dimensional measurements collected on low-depth surfaces, as is the case of hip distance or withers to tail

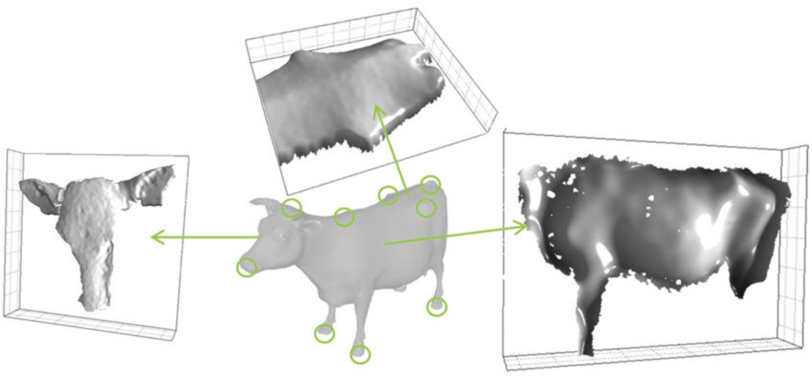


*Figure 1: The system was implemented in order to allow a fast extraction and monitoring of different cow body parameters and in particular heart girth, length and height*

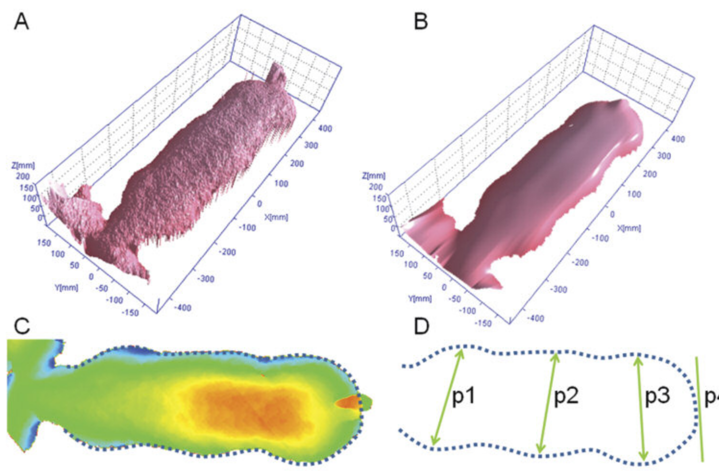



*Figure 2: Example of a portion of the animal back, as revealed by the top sensor before (A) and after (B) Gaussian filtering. Definition of the pig external profile (C) and localisation of the reference cross sections (D).*

head length. Slightly lower accuracy is achieved when analyses are carried out on high aspect ratio body shapes or on purely three-dimensional shapes, as found in the cases of average height or chest girth data comparison (with an $R^2$=0.7-0.9). The lowest performances are observed when parameter values are estimated as a combination of multiple parameters, due to uncertainty propagation: this has been shown in the cases of back slope or depth experiments ($R^2$<0.5).

Furthermore, compared to manual measurements, the adoption of the Kinect sensor for the quantification of body parameters allows a significant reduction in the absolute error means. Indeed, compared to the use of manual measurements, the non-contact Kinect approach enables the collection of more data in a shorter time and is not affected by difficulties associated with ruler reading or uncertainties related to animal movements that are characteristic of slow manual measurement approaches.

The method clearly needs pre-industrialisation engineering activity to allow automatic data collection and extraction. Additionally, studies are needed in livestock environments to understand the effects of prolonged dust exposure on the Kinect infrared projector or sensor, and the presence of high vibration levels due to the frequent passage of tractors or other vehicles in the proximity of the device installation.

However, considering that livestock farms are generally increasing their stock densities, the possibility of replacing manual with non-contact measurements is of great interest in order to optimise herd management based on individual animal's health, welfare and growth data monitoring, with no induced stress on animals.

**References:**

[1] A. Pezzuolo, et al.: "On-barn pig weight estimation based on body measurement by means of a Kinect v1 depth-camera", Computers and Electronics in Agriculture, 148: 29-36, 2018..

[2] M. Dubbini, et al.: "Last generation instrument for agriculture multispectral data collection", CIGR Journal, 19: 87-93, 2017.

[3] A. Pezzuolo, M. Guarino, L. Sartori, F. Marinello: "A feasibility study on the use of a structured light depth-camera for three-dimensional body measurements of dairy cows in free-stall barns", Sensors, 18(2): 673, 2018.

**Please contact:**

Andrea Pezzuolo
Department of Agroforesty and Landscape, University of Padua - Italy.
andrea.pezzuolo@unipd.it

# Data Science Techniques for Sustainable Dairy Management

by Kevin Fauvel (Inria), Véronique Masson (Univ. Rennes), Philippe Faverdin (INRA) and Alexandre Termier (Univ. Rennes)

***A multi-disciplinary team of experts from Inria and INRA are working towards improving farmers' income and working conditions in an environmentally friendly manner.***

As underlined in the last EU Agricultural Outlook 2017-2030 of the European Commission, growing global and EU demand is expected to support world dairy markets in the long term. However, farm economic viability encourages an increasing herd size. As a response, precision livestock farming (PLF) is a way to improve farm economic performance [1]. PLF is the use of continuous information to optimise individualised animal management. In addition, studies reveal that PLF can help enhance animal welfare [2], environmental sustainability [2] and working conditions [3].

## Dairy Management

In dairy farming, data is collected through different types of sensors based on animals (e.g., temperature, accelerometer, feed intake, weight), buildings (e.g., temperature, humidity) or milking robots (e.g., volume, milk composition). These sensors are common in today's farms and facilitate activity reporting.

However, the level of data analysis provided does not allow the farmer to make a decision such as inseminate or give medical treatment. Indeed, alerts from current devices often contain too many false positives to be reliable and the level of detail is too high to reduce the workload. For example, false positives for disease detection may result in extra medical treatment and monitoring costs, the inverse of the initial intent. Another key management aspect is heat detection for insemination purposes. It is fundamental because a cow must have a calf to begin lactating and is expected to give birth once a year thereafter to maintain a certain level of production. Thus, false positives for ovulation phase detection can lead to suboptimal production, additional semen costs and decisions to cull cows due to reproductive problems.

## Research Project

Current techniques used by animal scientists focus mostly on mono-sensor approaches (e.g., accelerometer), which are often insufficient to reduce false positives and ease decision-making. For specific diagnostics such as ovulation detection, there exist efficient mono-sensor approaches such as progesterone dosage in milk. However, such technologies remain too expensive for a massive implementation in dairy farming.

Our goal is to use widespread sensors to provide reliable recommendations to facilitate farmers' decision-making about insemination, disease detection and animal selection. Our aim is to combine common sensors in dairy management (e.g., accelerometer and temperature) in order to diagnose events

*Figure 1: A cow equipped with an activity meter.*

and elaborate upon these recommendations. This problem structure requires the use of machine learning methods. Thanks to an experimental farm, labeled data is available. The challenge is to design new algorithms which take into account data heterogeneity, both owing to their nature (e.g., temperature, weight, feed intake) and time scales (e.g., each five minutes, twice a day, daily).

Indeed, our approach will rely on multivariate time series classification and no existing method is designed to efficiently handle data having different time scales per dimension. First, most methods are window based: they extract features with the same temporal granularity per variable. Then, among methods specific to each variable properties, no exhaustive and efficient way to characterise different events has emerged.

The goal is to provide solutions to be transferred to the agricultural sector. This work is led by Inria LACODAM and INRA PEGASE teams in France under the #DigitAg initiative. #DigitAg is the French Digital Agriculture Convergence Lab, which brings together 360 researchers and higher education teachers from leading French organisations in this field. It is supported by the French state through the National Agency for Research funding with the reference ANR-16-CONV-0004.

**References:**


[1] W. Steeneveld et al. : "Characterization of Dutch dairy farms using sensor systems for cow management", Journal of Dairy Science, 2015

[2] D. Berckmans: "Precision livestock farming technologies for welfare management in intensive livestock systems", Revue Scientifique et technique, 2014

[3] RL. Schewe, D. Stuart: "Diversity in agricultural technology adoption: How are automatic milking systems used and to what end?", Agriculture and Human Values, 2015


**Please contact:**


Kevin Fauvel, Inria, France
kevin.fauvel@inria.fr

Alexandre Termier, Univ. Rennes, France
alexandre.termier@inria.fr


# Measuring Micro-climate to Create Disease Prediction Models

by Christian Hirsch (TU Wien)

***In crop production there are still many diseases where the time of infection and outbreak in connection with the weather is still not known. Some farmers have weather stations in their fields to monitor weather conditions in order to prevent a possible spread of a disease. However, the measurements of the weather stations represent the climate within an area of several 100 m$^2$. To overcome this problem, TU Wien and BOKU Wien are preparing an IoT infrastructure to measure micro-climate in vineyards. Many sensors are placed all over the field and the collected data will be used to create disease prediction models using machine learning algorithms.***

In addition to natural disasters, crop diseases are still a major risk to crop yields in agriculture. Some diseases are already well known and studied, and there are models for these diseases that can tell a farmer the risk of its occurrence based on weather information and weather forecast. With this information the farmer is able to take steps to protect the crops right on time. However, the main causes of other diseases are still unclear, making it difficult for farmers to take steps to protect their crops. One of these diseases is the powdery mildew in vineyards.

Many fields and vineyards, used for educational and scientific purposes, are equipped with weather stations. These stations typically measure wind, precipitation, humidity, temperature and leaf wetness. The problem, however, is that the measurements of a weather station represent a whole field of several 100 m$^2$, making it impossible to detect micro-climate changes that might trigger a certain disease. To overcome this problem, TU Wien [L1] and the University of Natural Resources and Life Sciences, Vienna (BOKU Wien) [L2] will set up an Internet of Things (IoT) infrastructure consisting of small battery powered sensors [L3]. Like the weather station, the sensors will measure humidity and temperature. Additionally sensors for atmospheric pressure, soil moisture, soil temperature and $CO_2$ equivalents will be deployed. Many sensors of the same kind will be used and placed all over the field, enabling the scientists to record climate differences within a field.

The IoT infrastructure will consist of three basic parts: the swarm (Figure 1a), the fog (Figure 1b) and the cloud (Figure 1c). The swarm represents the set of sensors used to measure micro-climate. Typically those swarm nodes use wireless means of communication with a base station or GSM. The nodes are normally also equipped with a

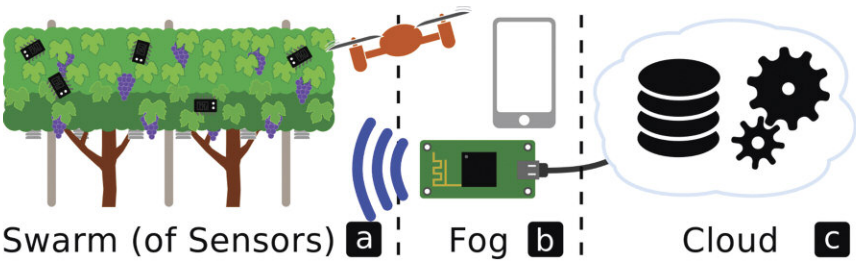



*Figure 1: Simple IoT infrastructure consisting of a) the swarm (of sensors and actuators), b) the fog and c) the cloud.*

microprocessor powerful enough to pre-process measurements in order to reduce the amount of data that has to be transmitted. The second part of the IoT infrastructure is the fog. The fog's main job is to communicate with the swarm and the internet, i.e., it acts as a router. The fog receives the measurements from the swarm via a wireless communication stack and transmits the data to the cloud storage. But the fog is more than just a simple router; it can also do pre-processing and simple analysis of the data before forwarding it to the cloud, known as fog- or edge computing [2]. This means that the fog can already trigger tasks, like notifications to a user, a robot or other actuators (e.g., irrigation systems) [1, 2]. Last but not least, the data from the fog will be transmitted to the cloud, which is mainly used as a database that collects all the data. In addition to acting as a simple database, the cloud also provides an interface to the data, which can, for example, visualise information for a user (user interface) and provide tools to analyse the data. Tasks done by the cloud are typically not time critical, e.g., machine-learning [1, 3].

For the specific case of analysing powdery mildew in vineyards, special swarm nodes, based on the Simblee BLE System on a Chip (SoC) have been developed. This SoC consists of an ARM-Cortex M0 processor with a BLE communication stack and an integrated antenna. This chip is used on two different swarm nodes: one that measures soil moisture, and soil temperature; the other one measures air temperature, humidity, atmospheric pressure and $CO_2$ equivalents in the air. The fog nodes are based on the Raspberry Pi mini computer. This computer already has the ability to connect to the swarm nodes via BLE. It receives data from the swarm nodes and forwards it the cloud. The cloud is based on the open-source IoT platform Thingsboard. It already has the capability to receive data and store it in a database, and it offers a web-based user interface to present the data, as well as other interfaces to query data.

This spring, the prototypes of the swam nodes will be evaluated in the green house. Based on the results of the evaluation phase, the sensors will be upgraded in order to meet the requirements necessary for outdoor use. This includes the housings, and further deficiencies occurred during the evaluation phase. Then, the actual data recording begins and sensors are placed outdoors in the vineyards. After collecting the data from at least one season, the data analysis can start, and scientists can hunt the triggers of diseases.

**Links:**
[L1] https://ti.tuwien.ac.at/cps
[L2]https://www.dnw.boku.ac.at/en/ps/
[L3] http://cpsiot.at/

**References:**

[1] A. Al-Fuqaha, et al.: "Internet of things: A survey on enabling technologies, protocols, and applications," in IEEE Communications Surveys & Tutorials, vol. 17, no. 4, pp. 2347–2376, 2015.
[2] W. Shi, et al.: "Edge computing: Vision and challenges," IEEE Internet of Things Journal, vol. 3, no. 5, pp. 637–646, Oct 2016.
[3] M. Armbrust, et al.: "A view of cloud computing," Commun. ACM, vol. 53, no. 4, pp. 50–58, Apr. 2010.


**Please contact:**

Christian Hirsch
TU Wien, Computer Engineering / Cyber-Physical Systems
christian.hirsch@tuwien.ac.at

Markus Redl
BOKU Wien, Department of Crop Sciences / Divison of Plant Protection
markus.redl@boku.ac.at


# Drone Swarms in the Field

by Dario Albani and Vito Trianni (ISTC-CNR)

***Drawing inspiration from the behaviour of honeybees, a buzzing swarm of small unmanned aerial vehicles (UAVs) will fly over hectares of cultivated fields seeking weeds.***

Want the grass to be greener on your side of the fence? Robots are the new frontier for agriculture and will soon be making this a reality. The Swarm Robotics for Agricultural Applications (SAGA) project is one of the frontrunners in deploying robot swarms in the field: drawing inspiration from the behaviour of ants and bees [1], it uses a swarm of small unmanned aerial vehicles (UAVs) to process information onboard and to constantly monitor the presence of weeds and the status of the crops. More efficient monitoring can help with eliminating weeds, whilst reducing costs and improving crops.

SAGA tackles the problem of sugar beet fields infested by volunteer potatoes—a common benchmark in agricultural robotics. Volunteer potatoes originate from tubers that remain in the soil after harvesting. During the next season, when sugar beets are grown in the same field, volunteer potatoes are a major threat, as they spread diseases and facil-

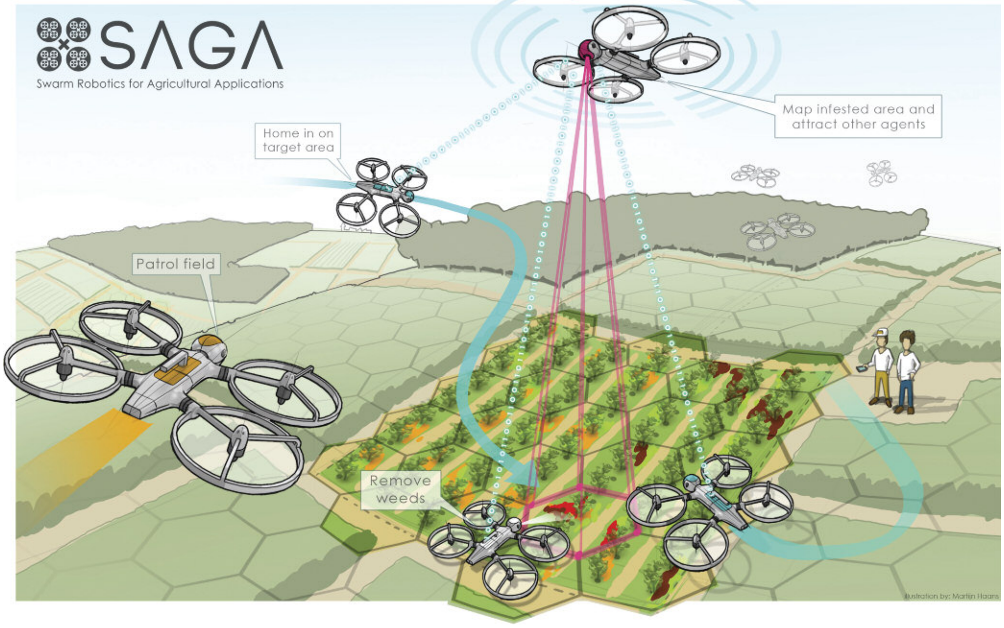



*Figure 1:The Swarm Robotics for Agricultural Applications (SAGA) project is deploying robot swarms for weed monitoring and mapping.*

itate harmful soil nematodes. Regulations require farmers to control this weed — a costly operation that in the Netherlands has an estimated cost that goes up to 300€ per hectare per growing season. This cost could be drastically reduced by an efficient and automated weeding system.

Drones represent a valid solution and are expected to have a big impact on precision agriculture. A single UAV traversing and inspecting the field could provide the farmer with huge streams of data. A swarm not only guarantees the same output in a fraction of the time but, thanks to distributed algorithms, drastically increases accuracy and precision.

Controlling a bevy of robots is not trivial and requires specific techniques to ensure robustness to faults and scalability to large groups. To this end, UAVs employ an algorithm relying on random walks, an exploration technique based on stochastic movements [2]. First, the field is discretised into a 2D grid with fixed size cells. For every decision step, probabilities are assigned to the cells in a way to optimally deal with individual effort and social influence. To avoid costly manoeuvres, both in terms of energy and time, bigger probabilities are assigned to cells that are in the same direction of motion. To obviate long relocations, only a subset of cells in the vicinity is taken into consideration. Moreover, drones collaborate and send information about their position and already visited locations. This information is translated into attracting and repulsive forces: attraction toward points of interest and repulsion from each other. These mechanisms can be controlled with a few parameters that allow the system to be tuned and adapted to specific working conditions, achieving an efficient, reliable and fast mapping of the field [2].

Furthermore, the swarm is designed to exploit all three dimensions. Many applications require high-resolution data only in certain areas while other parts can receive lower attention. Hence, UAVs can adjust the altitude to increase the resolution only where needed, making non-uniform strategies efficient both in time and energy expenditure [3]. A close inspection guarantees higher image resolution and greater accuracy, but at the cost of higher inspection time due to the limited footprint of the camera. To obtain the best trade-off between accuracy and speed, the SAGA swarm implements a decentralised deployment strategy inspired by honeybee behaviour [1], that dynamically assigns UAVs to different areas to be monitored, and suitably re-assigns them to other areas when needed. In this way, UAVs only inspect some areas at low altitude, while mildly monitoring other cells. Thanks to this self-organising strategy, the swarm can partition the monitoring task in an optimal way and autonomously allocate the required resources only where and when needed.

The successful demonstration of the SAGA project raises the bar for swarm robotics research. The natural step forward is the extension of the agricultural swarms concept toward heterogeneous systems made of ground and flying robots, tackling all sort of tasks within the farm, from weed removal to pest control, from optimal usage of fertilisers to harvesting.

**References:**

[1] A. Reina, J. A. R. Marshall, V.Trianni, T.Bose, "Model of the best-of-N nest-site selection process in honeybees" Phys. Rev. E, vol.95, no.5, 2017, pp. 52411–52415.

[2] D. Albani, D. Nardi, V. Trianni, "Field Coverage and Weed Mapping by UAV Swarms," in IEEE International Conference on Intelligent Robots and Systems, 2017, pp. 4319–4325.

[3] S. A. Sadat, J. Wawerla, R. T. Vaughan, "Recursive non-uniform coverage of unknown terrains for UAVs," in IEEE International Conference on Intelligent Robots and Systems, 2014, pp. 1742–1747.

**Links:**

http://laral.istc.cnr.it/saga/

**Please contact:**

Dario Albani, (+39) 06 44595 341, dario.albani@istc.cnr.it

Vito Trianni, (+39) 06 44595 277, vito.trianni@istc.cnr.it

# Satellite and Aerial Image Processing for Smart Farming and Biodiversity Conservation

by Andrea Manno-Kovács, András L. Majdik and Tamás Szirányi (MTA SZTAKI)

***Over the last three years, the Institute for Computer Science and Control (SZTAKI) of the Hungarian Academy of Sciences (MTA) has been conducting research in the area of smart farming. The Machine Perception Research Laboratory [L1] has been integrating drone-captured aerial images with freely available European Space Agency (ESA) Sentinel satellite imagery to develop an operational system for smart farming applications.***

Our first goal is to enhance the limited resolution of satellite images of the test sites by capturing high resolution video footage of the same test areas. To achieve this, we co-register geo-referenced images from different sources and resolutions. Secondly, our goal is to detect changes between images collected periodically with different technologies, and more importantly, to detect and estimate damage in crop production and fruit plantation areas as well as to monitor the natural habitats representing a major role in sustainable agriculture and ecological farming.

In our first project, during the summer of 2017 we periodically collected data with a quadcopter equipped with an RGB and a four channel (Green-550 nm, Red-660 nm, Red edge -735 nm, Near infrared -790 nm) multispectral camera system (i.e., Parrot Sequoia) in three different farming sites in the vicinity of Biatorbágy town, Budapest metropolitan area, Hungary. In the course of the multi-session mapping we captured more than 25,000 narrow-band 1.2 mega pixel images covering 41 hectare (ha) of wheat field, 18 ha of cornfield, and 8 ha of fruit and vineyard. Next, by computing the geo-referenced orthomosaic photo corresponding to every multispectral band we analysed the normalised difference vegetation index (NDVI) at different time instances. Similarly, the NDVI index was estimated using satellite imagery whenever data was provided by the ESA Sentinel satellite in order to track the changes in terms of live green vegetation quantity at the smart farming sites (Figure 1).

Beside, the conservation of natural habitats is also connected to smart farming and sustainable agriculture. Climate change and agriculture are related processes; therefore special emphasis is placed on such methods that preserve and preferably regenerate the ecosystem. EU directives require the preservations of natural habitats (e.g. the Natura2000 habitats) and also the Hungarian government made steps for biodiversity conservation concentrating on three types of habitats: forests, grasslands and wetlands.

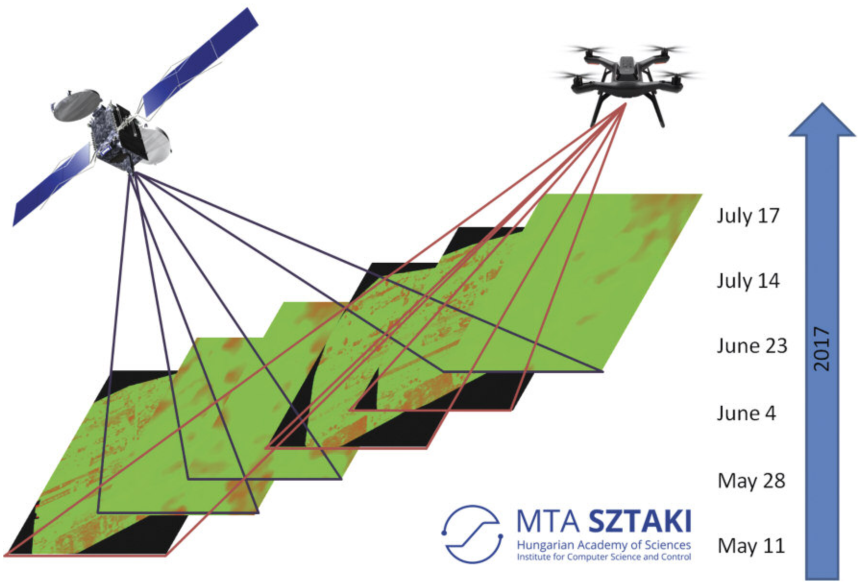



*Figure 1: Summary of results: NDVI indexes computed from satellite imagery and low altitude (70 meter above ground) aerial mapping captured at different time instances over the summer of 2017.*

While small wetland patches are closely connected to smart farming, this vegetation in the mosaic of the landscape receives far less attention. However, these small wetlands are especially important and should not be modified by agriculture and urbanization, as they provide sanctuaries for wildlife sensitive to human presence. Moreover, they also play an important role in regulating water quality and quantity through pollution absorption, groundwater recharge and flood retention. The small patches of wetland areas give an important contribution to the co-habitation inside agricultural, urban and traffic areas. At the same time, agricultural land acts as a buffer zone around wetlands, protecting them from developing industrial zones and urban areas. Due to this connected relation between agricultural areas and wetlands, smart farming should also concentrate on the tracing and tracking of these small habitats. To monitor such areas, our ESA OWETIS (Observation of local WETland areas from Satellite Imaging) project concentrates on the automatic localization and observation of potential small backcountry wetlands.

In our second project (OWETIS), MTA SZTAKI, together with the Balaton Limnological Institute, Centre for Ecological Research, Hungarian Academy of Sciences (Andras Zlinszky) and Airbus DS GEO Hungary Ltd. (Gyorgy Domokos) performs a nation-wide observation of small local wetland areas from satellite imaging [L3]. As a first step, a GIS dataset has been built, focusing on small, backcountry wetlands which are not connected to large, open water surfaces. The wetland candidates were first selected using Google Maps, then verified in field campaigns with UAVs, the identification is based on the collection

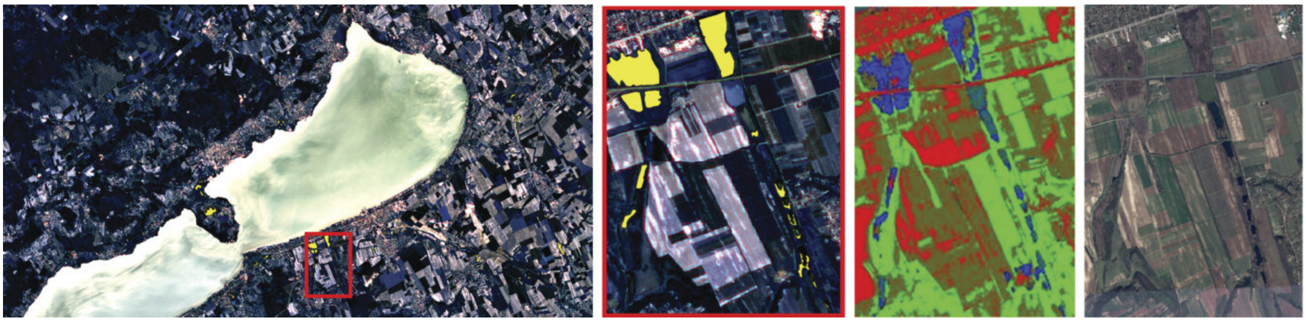


*Figure 2: Wetland mapping on Lake Balaton - The test patches of the wetland database are shown in yellow, followed by the MRF-based multitemporal segmentation [2] and the test site in Google Maps.*

of GPS points and polygon data together with field photographs. Beside the location and extent, local water regime, habitat quality (biodiversity) and observation of flora and fauna is also included in the dataset. The database includes the surrounds of the three largest Hungarian lakes: Lake Balaton, Lake Tisza and Lake Neusiedl (Hungarian side). The database contains 150 test patches, out of which 113 are wetlands, providing a good representation of the variety of Hungarian wetlands. Besides wetlands, there are 23 grasslands (also an important habitat type) and 14 croplands.

The main focus of the project is to use this database to investigate temporal, spectral and spatial characteristics of wetlands using Sentinel multispectral satellite imagery, and from this to develop an automatic detection process, which can provide further information about natural habitats. The detection process applies a fusion based Markov Random Field (MRF) technique [1] on multitemporal and multispectral image series. In our fusion MRF model an unsupervised method tracks temporal changes of wetland areas by comparing the class labelling of different time layers [2]. A classification map based on existing airborne laser scanning data [3] has been used for wetland classification improving the discrimination of land-cover classes with similar spectral characteristics. The proposed method can be extended by machine learning and feature based classification for more accurate monitoring of wetlands.

**Links:**
[L1] http://mplab.sztaki.hu
[L2] https://kwz.me/hbv
[L3] https://kwz.me/hbw

**References:**
[1] T. Szirányi and M. Shadaydeh: "Segmentation Of Remote Sensing Images Using Similarity Measure Based Fusion-MRF Model", IEEE Geosci. Remote Sens. Lett., vol. 11, no. 9, pp. 1544-1548, 2014
[2] M. Shadaydeh, et al.: "Wetland Mapping by Fusion of Airborne Laser Scanning and Multitemporal Multispectral Satellite Imagery", Int. J. Remote Sens., vol. 38, no. 23, pp. 7422-7440, 2017
[3] A. Zlinszky, et al.: "Categorizing wetland vegetation by airborne laser scanning on Lake Balaton and Kis-Balaton, Hungary", Remote Sens., vol. 4, no. 6, pp. 1617-1650, 2012

**Please contact:**
Andrea Manno-Kovacs
MTA SZTAKI, MPLab
andrea.manno-kovacs@sztaki.mta.hu

# Case-based Reasoning for Forecasting the Allocation of Perennial Biomass Crops

by Florence Le Ber (Université de Strasbourg, ENGEES), Jean Lieber (Université de Lorraine, Inria) and Marc Benoît (INRA)

***A farmer's decision to plant new perennial biomass crops is a complex process that involves numerous parameters and can change the food / non-food balance. The paradigm of case-based reasoning is able to deal with the kind of complex and sparse information that is available in this situation, to model and forecast the allocation of these crops.***

Perennial biomass crops are new renewable energy resources that could play an important role in replacing fossil fuels. The expansion of these crops seems unavoidable, and will raise global issues such as competition for space between food and non-food crops. To forecast their introduction in farming systems and their spatial allocation is thus a main challenge. Whilst several land-use change models deal with biomass crop allocation, most of them simulate large-scale allocation processes, taking into account numerous biophysical variables but only a few true-to-life human variables. In contrast, we aimed to model farmers' allocation decisions regarding new perennial biomass crops as a complex agricultural management system, coupling social, technical and environmental variables.

As this objective raises knowledge acquisition and knowledge integration methodological issues, we proposed to model biomass crop allocation relying on case-based reasoning (CBR), a problem-solving paradigm. CBR consists of solving new problems by reusing the solution of similar problems that have already been solved [1]. A case corresponds to a problem-solving episode usually represented by a problem-solution pair. Cases are recorded in a case base. A source case is an element of the case base. The CBR process consists of solving a new problem, the target problem, using the case base. A common way of doing this is to select a source case similar to the target problem (case retrieval) and to modify the retrieved case so that it provides a solution that hypothetically solves the target problem (case adaptation). After these inference steps, some learning steps are sometimes implemented.

This approach was tested on a case study located in Burgundy (East of France), where a new perennial biomass crop *Miscanthus sp*. has been established since 2010.

In our model [2], a case is defined as a specific experience of *Miscanthus* allocation (or non-allocation) in a farm field. The case base includes 82 farm fields that have been described by farmers in past interviews (2011-2012) as having *Miscanthus* allocation potential. The problem-solution pair is a farm field and its allocation potential for *Miscanthus*. Each case is represented with a vector of qualitative values, divided into two parts: the problem part is a set of attributes, giving the farm field characteristics, as described by the farmer; the solution part describes the *Miscanthus* allocation potential of the farm field (0-no allocation, 1-allocation, or 2-possible allocation). The model also relies on a set of rules that formalise the elements given by farmers when explaining their decision to plant (or not plant) *Miscanthus* in a field.

To define a similarity measure between cases, we assumed that similar farm management and biophysical constraints of farmland enable analogue farmers' decisions about crop allocation. The comparison of the source cases and the target problem is thus based on a comparison of their attribute values in terms of cropping plan, farm biophysical features and spatial farmland features.

In the adaptation step, we use the rules of the farmer associated to the source case to build the solution. Adaptation knowledge allows the appropriate rule to be chosen and applied, according to a given adaptation context: the system

***Figure 1: A Miscanthus field in Burgundy.***

promotes the rules with conclusion 0 (when the context is not favourable for *Miscanthus*, e.g., because its price is low compared with traditional crops) or those with conclusion 1 or 2, if the economic context is favourable for *Miscanthus*.

Experiments showed the central role of the user, who has to choose parameters and adaptation contexts, and to examine results step by step: retrieved cases, available rules, proposed solutions. Rules in particular can be analysed to highlight the field characteristics that are important with respect to the farmer's decision. Future work will deepen the use of rules for adaptation and explanation. This model could also be tested on other innovative crops. Finally, a formal model for the adaptation of crop allocation has been developed that is based on the application of belief revision in the qualitative algebra RCC8 setting (Dufour et al., 2012 ): this model still needs to be applied and validated on real farm data.

Farm surveys and knowledge modelling were funded by FUTUROL project [L1] and the French government.

**Link:**
[L1] https://www.projetfuturol.com

**References:**

[1] C.K. Riesbeck, R.C. Schank: 3Inside Case-Based reasoning", Lawrence Erlbaum Associates Publishers, 1989.
[2] F. Le Ber, et al.: "A Reasoning Model Based on Perennial Crop Allocation Cases and Rules", ICCBR. Springer, LNAI 10339, 61-75, 2017.
[3] V. Dufour-Lussier, et al.: "Adapting Spatial and Temporal Cases", ICCBR. Springer, LNAI 7466, 77-91, 2012.

**Please contact:**

Florence Le Ber, ICube, Université de Strasbourg, CNRS, ENGEES, France
+33 3 88 24 82 30
florence.leber@icube.unistra.fr

Marc Benoît, ASTER, INRA, France
+33 3 29 38 55 01
marc.benoit@inra.fr

# Distributed Management of Scientific Workflows for High-Throughput Plant Phenotyping

by Christophe Pradal (CIRAD), Sarah Cohen-Boulakia (Univ. Paris-Saclay), Gaetan Heidsieck (Inria), Esther Pacitti (Univ. Montpellier), François Tardieu (INRA) and Patrick Valduriez (Inria)

***High-throughput phenotyping platforms allow acquisition of quantitative data on thousands of plants required for genetic analyses in well-controlled environmental conditions. However, analysing these massive datasets and reproducing computational experiments require the use of new computational infrastructure and algorithms to scale.***

Plant species will need particular characteristics to survive in the world's changing climate. To this end, plant scientists are analysing traits of interest to identify both the natural genetic variations in plants and the genetic control of their responses to environmental cues.

*Figure 1 The Phenoarch phenotyping platform is one of the Phenome node in Montpellier. It has a capacity of 1,500 plants per experiment with a controlled environment (e.g., temperature, humidity, irrigation) and automatic imaging through time (13 images per plant per day).*

In the last decade, high-throughput phenotyping platforms have allowed acquisition of quantitative data on thousands of plants required for genetic analyses in well-controlled environmental conditions. These platforms in controlled conditions are essential to understand the variability of yield in the field depending on climatic scenarios [3]. Recently, national infrastructures such as Phenome have used high-throughput platforms to observe the dynamic growth of a large number of plants under in field and platform conditions. These instrumented platforms produce huge datasets (thousands of images, environmental conditions and sensor outputs) with complex in-silico data analyses that result in the definition of new, complex variables [2]. The seven facilities of Phenome produce 200 terabytes of data annually, which are heterogeneous (images, time courses), multiscale (from the organ to the field) and originate from different sites. This infrastructure paves the way for the collection of data at an even larger scale. Indeed, farmers and breeders can increasingly capture huge amounts of diverse data, which at this stage is mostly environmental, but which can involve detailed maps of yield in fields (precision agriculture), and images originating from remote sensing and drone imaging. Hence, the major problem becomes the automatic analysis of these massive datasets and the ability to reproduce large and complex in-silico experiments.

Such experiments can actually be represented and executed in an efficient and reproducible way by means of scientific workflows in which computational tasks can be chained (e.g., upload input files, preprocess the data, run various analyses and simulations, aggregate the results). OpenAlea is a scientific workflow system that provides methods and software for plant modelling at different scales [L1]. We have used it in the context of Phenome, by developing the OpenAlea Phenomenal software package dedicated to the analysis of 3-D plant architecture, whose outputs can be used for ecophysiological modelling. Phenomenal provides fully automatic workflows dedicated to the 3D reconstruction, segmentation and tracking of plant organs. It has been tested on maize, cotton, sorghum and apple trees. OpenAlea radiative models are used to estimate the light use efficiency and the in-silico crop performance in a large range of contexts.

Executing workflows on large datasets is time-consuming, and thus often incompatible with the scientist's way of working, where trial and error is an essential component. We have designed the infrastructure InfraPhenoGrid [1] to distribute the computation of workflows using the EGI/France Grilles computing facilities. EGI provides access to a grid with multiple sites, each with one or more clusters. This environment is now well suited for data-intensive computing, with different research groups collaborating at different sites. In this context, the goal is to address two critical issues in the management of plant phenotyping experiments: 1) scheduling distributed computation and

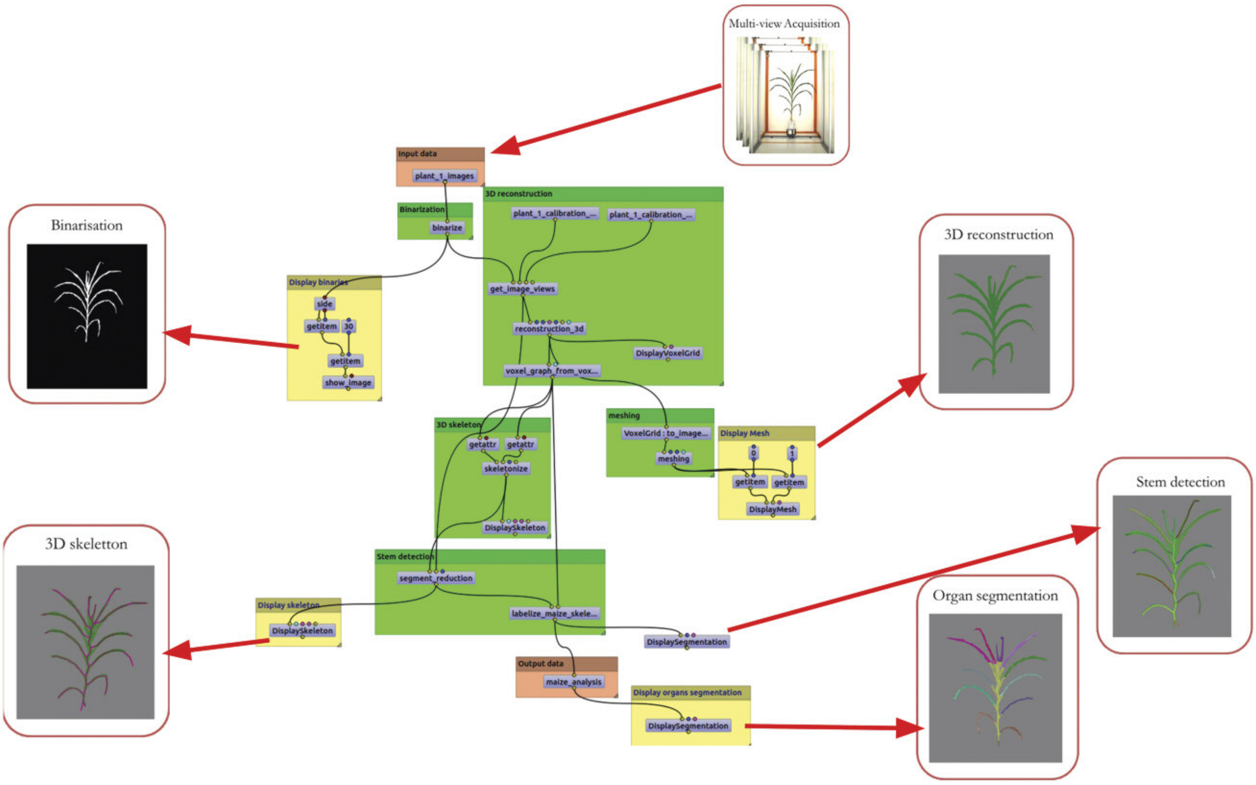



*Figure 2: The 3D reconstruction workflow Phenomenal in the OpenAlea visual programming environment is applied to one plant at a given time. The same workflow is run on thousands of plants through time, consuming several terabytes of data, and distributed transparently on the European Grid infrastructure (EGI).*

2) allowing reuse and reproducibility of experiments:

1. Scheduling distributed computation
We have adopted an algebraic approach suited to the optimisation and parallelisation of data-intensive scientific workflows. The scheduling problem resembles scientific workflow execution in a multisite cloud [2]. In the context of the #DigitAg project [L2], our objective is to propose new scalable and elastic heterogeneous scheduling algorithms that will transparently distribute the computation of these very large computational experiments on local servers, where the data are stored but with limited resources, on multisite clouds, and on the european grid, which provide computing power but with lower availability and longer delays to access to the data.

2. Allowing reuse and reproducibility of experiments
The second challenge we addressed was to help scientists to foster discovery of novel traits and mechanisms based on the processed datasets, while providing a robust methodology with support for reproducibility and reuse. Modern scientific workflow systems are now equipped with features that offer this support. The provenance information (parameter settings and data sets consumed and produced) can be systematically recorded. Moreover, InfraPhenoGrid and OpenAlea workflows enhance reproducibility and reuse by providing users with the means to interact with provenance information through Jupyter electronic Notebooks [L3] [1].

The authors acknowledge the support of France-Grilles for providing computing resources on the French National Grid Infrastructure.

**Links:**
[L1] https://openalea.gforge.inria.fr
[L2] http://www.hdigitag.fr
[L3] https://jupyter.org

**References:**
[1] C. Pradal, S. Artzet, J. Chopard, et al.: "InfraPhenoGrid: A scientific workflow infrastructure for plant phenomics on the Grid", Future Generation Computer Systems 67: 341-353, 2017.
[2] J. Liu, E. Pacitti, P. Valduriez, et al.: "Multi-objective scheduling of Scientific Workflows in multisite clouds", Future Generation Computer Systems 63: 76-95, 2016.
[3] F. Tardieu, et al.: "Plant Phenomics, From Sensors to Knowledge", Current Biology 27(15):R770-R783, 2017.

**Please contact:**
Christophe Pradal, Inria, France
+33 4 67 61 97 94
christophe.pradal@cirad.fr

Patrick Valduriez, Inria, France
+33 4 67 14 97 26
patrick.valduriez@inria.fr

# Virtual Research Environments to Support Agriculture and Food Research Communities: The AGINFRA+ Project

by Panagiotis Zervas, Leonardo Candela and Pythagoras Karampiperis (Agroknow)

***Virtual research environments are proposed as a prominent cloud-based solution for agricultural and food scientists willing to collaborate and seamlessly access, use and reuse research resources such as datasets, mathematical models, software components results and publications.***

Current developments in ICT potentially provide innovative and more effective ways to support agricultural and food researchers to work with extremely large data sets and handle use cases involving big data. This new paradigm raises new research questions, new methods and approaches to perform research in agriculture and food. To support this paradigm, research facilities and e-infrastructures need to be revisited and new partnerships among academic institutions and private companies should be developed. This might lead to the involvement of scientists from areas of science and technology, who are not involved in the agricultural and food sector, to boost the innovation process in this sector. This also creates new challenges and opportunities for informed decision making from the micro scale (i.e., precision farming) to the macro scale (i.e., policy making) [1]. To this end, it is evident that large-scale, cloud-based infrastructure assets need to be utilised to support the agriculture and food research communities in the process of data storage and indexing, algorithm execution, as well as results visualisation and deployment.

A prominent existing cloud-based solution is the Virtual Research Environments (VREs) provided by the D4Science Initiative. VREs are web-based, community-oriented, collaborative, user-friendly, open-science-compliant working environments for scientists and practitioners working together on a research task [2].

In this context, the AGINFRA+ project [L1] is exploiting the VREs paradigm for three (3) prominent research communities, namely:

- The agro-climatic and economic modelling research community: this community focuses use cases related to crop modelling and crop phenology estimation. Current cases aim to support the workflow of researchers, intermediaries and business analysts working on crop modelling and yield forecasting and related activities in the area of policy and decision support in food security, farm management advice and related activities. In particular, the aim of these use cases is to bring researchers from their current usually local, single computer and mostly peer network based work space to a cluster compute and cloud based collaborative work environment. Providing the community with such advanced options for virtual research will increase the buy-in to use such frameworks and support the community in making the transition to effective collaborative, cloud-based research.
- Food safety risk assessment research community: this community focuses on use cases to support scientists in the multidisciplinary field of risk assessment and emerging risk identification as there is currently a strong need to create new software-based solutions that support the knowledge integration processes relevant for these tasks. Domain-specific, cloud-based research environments are a promising solution to overcome current limitations and frustrations in these domains. Specifically, features that facilitate scientific collaboration as well as features to store and share data and knowledge are important. Furthermore, the usage of broader infrastructural and technical solutions can facilitate the creation and adoption of standards which will

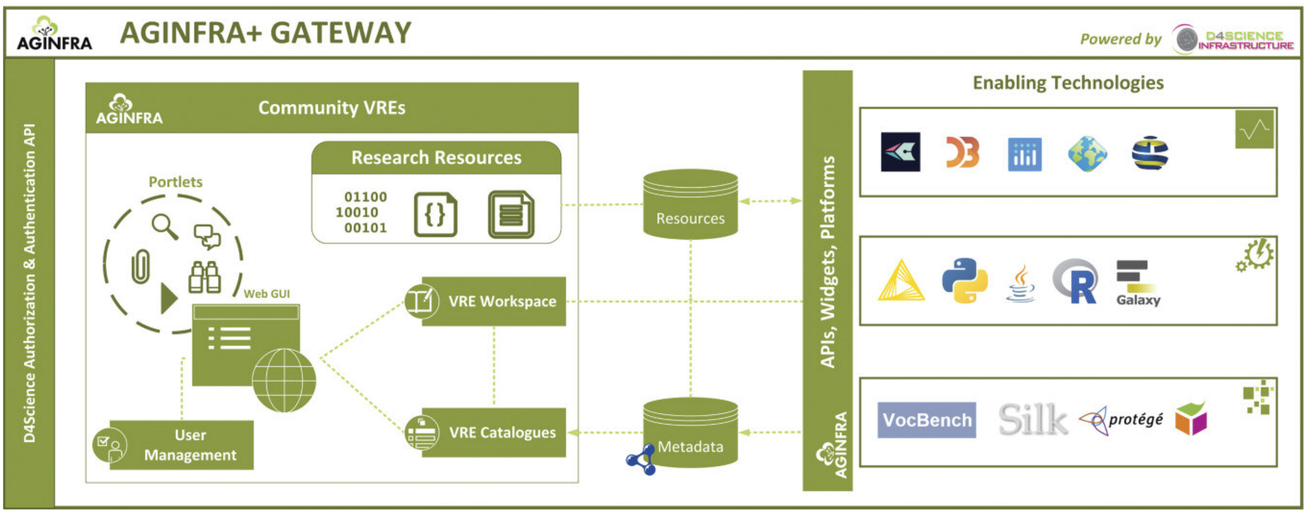



*Figure 1: AGINFRA+ Technical Infrastructure based on research community VREs.*

increase efficiency along all scientific processes.
- Food security research community: this community focuses on use cases related to high-throughput phenotyping to support phenomics researchers to select plant species and varieties which are the most adapted to specific environments and to global changes. High-throughput phenotyping produces a large amount of data which need to be analysed immediately for decision making, as a result cloud-based research environments are a promising solution to support this process.

In order to support the aforementioned research communities, AGINFRA+ project has collected requirements from these communities and appropriate VREs have been set-up [L2]. These VREs encapsulate the technical solutions serving the community requirements within a collaborative environment that allows the setup, execution, monitoring and sharing of research activities and their results. More specifically, as presented in Figure 1, the VREs provide researchers access to research resources such as data, models and publications, in order to design and execute their research. The resources are findable via the VREs' Catalogues, which are informed by semantically rich metadata organised and generated using the AGINFRA+ Data & Semantics Layer technologies. Experiments are carried out via the execution of the available models over the services provided by the Analytics & Processing Layer. Finally, results are visualised, organised and shared using the technologies incorporated in the AGINFRA+ Visualisation & Publishing Layer.

In summary, the aforementioned approach aims to serve the needs of three adjacent but not fully connected user communities, which perform research on multi-disciplinary and multi-domain problems related to agriculture and food. Finally, it aims to support open science to research resources for agriculture and food.

**Links:**
[L1] http://www.plus.aginfra.eu/
[L2] https://aginfra.d4science.org/explore

**References:**
[1]. S. Wolfert, L. Ge, C. Verdouw, M-J. Bogaardt: "Big Data in Smart Farming – A review". Agricultural Systems, Vol. 153, pp. 69–80, 2017 10.1016/j.agsy.2017.01.023
[2] L. Candela, D. Castelli, P. Pagano: "Virtual Research Environments: An Overview and a Research Agenda" Data Science Journal. vol. 12, pp. GRDI75–GRDI81, 2013 10.2481/dsj.GRDI-013

**Please contact:**
Panagiotis Zervas
Agroknow - Athens, Greece
+302106897905
pzervas@agroknow.com

# PERSEPHONE: Using Biogas Residues to Mitigate Agricultural Nitrogen Pollution

by Bella Tsachidou, Philippe Delfosse and Christophe Hissler (LIST)

***At the Luxembourg Institute of Science and Technology (LIST), a multidisciplinary scientific group is looking to prove how recycling of biogas residues back to agricultural soils has the potential to mitigate nitrate leaching, enhance long term storage of nutrients in the agro-ecosystems and improve agronomic performance.***

As the world population grows, crop yields need to be increased to provide sufficient quantities of food and stock feed. To meet this challenge, the farming system has to overcome limiting factors such as nutrient depletion of arable soils. Nitrogen being the main mineral element required for plant growth, has led to its industrial fixation and extensive use in the form of chemical fertilisers. As a result, the nitrogen cycle has been greatly modified, resulting in one of the most challenging environmental problems: nutrient pollution of aquatic and terrestrial ecosystems, as well as atmospheric contamination. Reactive nitrogen (Nr), released by intensive agricultural and industrial activities, causes problems owing to its great transformation capacity and enduring cascade between environmental media. Inefficient use and losses of Nr to the environment, have over time contributed to major issues such as the greenhouse effect and climate change, deterioration of soil quality, aquatic eutrophication and loss of biodiversity, elevated levels of nitrates in waters, and consequently, impacts on human health and society.

Different aspects of Nr pollution have led to the establishment of treaties, commissions and legal instruments in Europe. Contamination of drinking water was, inter alia, the main driver for the enactment of the European Nitrates Directive 91/676/EEC which aims to protect water bodies from agricultural nitrate leaching. The EU Member States are legally bound to adopt and implement the directive, and failing to comply with its Action Programmes may inflict financial penalties. To date, particular measures within the Nitrates Action Programmes are considered impractical and hard to implement, leading to additional costs for farmers. Limitation of livestock manure application to 170kg N/ha, in accordance with the ANNEX I, has resulted in excessive mineral fertilisation and consequent increase of nitrate loads in the soil and groundwater. Therefore, in order to solve the nitrogen pollution problem, whilst contributing to the circular economy, it seems reasonable to work toward N recycling.

To prevent further acceleration of environmental pollution caused by Nr from agricultural sources, and at the same time integrate the biogas sector into the new bio-economy, we propose the recycling of anaerobically digested organic wastes from various sources such as agriculture, food/feed industries, households and more, back to arable lands. Anaerobic digestion (AD) is an environmentally sound and cost competitive multipurpose process that generates both biogas and digestate (biogas residues, BRs). The digestate resulting is not a waste but a valuable co-product which contains all the nutrients initially present in the substrate, and therefore can be used as biofertiliser, recycling nutrients such as N and P back to agricultural soils. Moreover, the on-farm anaerobic digestion of readily available organic resources can cover the demands of the farm in heat and power (Figure 1). Even though the residues from the biomethanation process are a new type of organic fertiliser, there is a considerable number of research publications that highlight their beneficial effects on soil [1] and plants [2].

However, owing to the wide range of organic materials used in the process of anaerobic digestion, the residues generated create concerns about their potential content of heavy metals, organic pollutants, pesticides and pathogens. The interest in AD by the EU is driven by critical issues such as climate change, high fossil fuel prices and need for renewable energy, excessive use of mineral fertilisers, and the treatment and disposal of organic waste. The number of biogas plant reactors is drastically increasing as a response to worldwide energy demands and the growing dependency of the European Union on energy imports from non-member countries. This growth produces large amounts of biogas residue that requires immediate handling.

Within the framework of the “Persephone” Project [L1] carried out by LIST and partners, in situ trials of 17 different fertilisation schemes are in place across the grasslands of the Greater Region. Strategic fertiliser application is performed with respect to the guidelines of Nitrates Directive, while soil sampling and forage harvest are dictated by environmental factors affecting microbial activity and plant growth. The overarching purpose of the study is to investigate the contribution of biogas residues to the mitigation of nitrate leaching in the soil and the improvement of agronomic performance. Another aim, stemming from the ANNEX I of Nitrates Directive, is to demonstrate, on one hand, the rampant polluting nature of chemical fertilisers, and on the other hand, the environmental benefits resulting from their substitution by biogas residues. This approach could spur societies to move down a new path towards a sustainable nitrogen future. Furthermore, since biological nitrogen fixation is one of the largest suppliers of Nr in nature, we focus on the impact of biogas residue on soil microbial community composition, function and nitrogen cycling capacity - with emphasis on the transcripts involved in nitrogen transformation - by exploiting metagenomic and metatranscriptomic approaches.

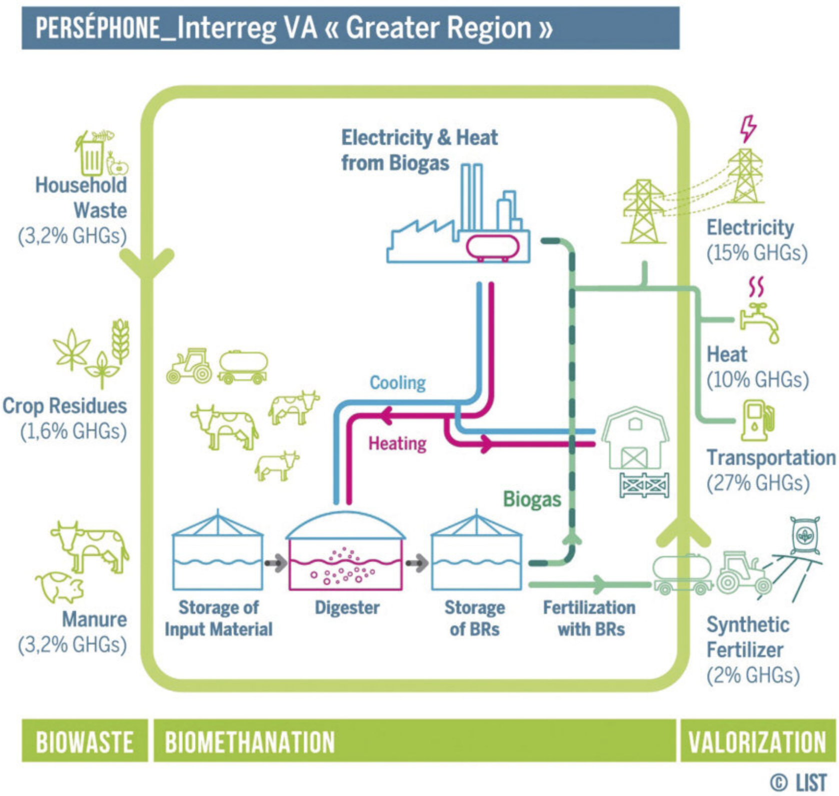



*Figure 1: On-farm production of electricity, heat and biofertiliser from biomethanation of organic waste. Implementation on a global scale has the potential to reduce greenhouse gases by 10%*

The project is a Transfrontier Program for European Territorial Cooperation and is co-financed by the European Union - European Regional Development Fund 2014-2020 INTERREG VA «Greater Region» and the Ministère du Développement Durable et des Infrastructures (MDDI) of Luxembourg.

Acknowledgments:
The authors acknowledge the contribution of the project partners: Au pays de l’Attert ASBL, Agria Grand Est, Ama Mundu Technologies, Sustain Water, Biogas Vereenenegung a.s.b.l., Biogas Beckerich, Université Libre de Bruxelles, Ecole Nationale Supérieure d’Agronomie et des Industries Alimentaires, Université de Liège, Agra-Ost, IZES gGmbH, Naturgas Kielen, La Ferme du Faascht, Biogas Rohlingerhof, Bio Recycle sarl, Lycée technique Agricole.

**Link:** [L1] https://kwz.me/hbx

**References:**
[1] M. Odlare, M. Pell, K. Svensson: “Changes in soil chemical and microbiological properties during 4 years of application of various organic residues,” in Waste Management, vol. 28:7, s. 1246–1253, Waste Management Elsevier, 2008.
[2] M. Koszel, E. Lorencowicz: “Agricultural Use of Biogas Digestate as a Replacement Fertilizers,” Agric. Agric. Sci. Proc., vol. 7, pp. 119–124, Jan. 2015.

**Please contact:**
Bella Tsachidou, Luxembourg Institute of Science & Technology
+352 470 261 5033,
bella.tsachidou@list.lu

# Computational 3D X-ray Imaging for the Food Industry

by Joost Batenburg and Robert van Liere (CWI)

***The Computational Imaging group at CWI develops mathematical techniques and algorithms for 3D image reconstruction. One of the main application areas is industrial quality control and rapid inspection, which is becoming increasingly important in the food processing industry. Characteristics that need to be tested and possibly also quantified include appearance of the product (size, shape, colour), texture, flavour, as well as internal factors (chemical, physical, microbial).***

The food processing industry has a rich history of applying 2D X-ray imaging for the analysis of their products. The advantage of X-ray imaging is that it is a non-destructive technique, i.e. the internal morphology of the products can be inspected, without having to cut open each product. However, traditional 2D X-ray imaging cannot be used to determine the exact 3D location of an anomaly and relevant features can sometimes not be seen in the individual X-ray images, requiring computational 3D imaging techniques that combine the information from multiple X-ray images.

The key challenge of introducing new X-ray imaging techniques in an industrial environment is that on one hand the image acquisition and algorithmic analysis must take place in a very short time, as throughput requirements are extremely high. On the other hand, the fact that only a few X-ray images can be taken in this short time results in a highly underdetermined image reconstruction problem, as the data is insufficient to compute a full 3D volumetric image.

## Tomography

The problem of reconstructing a 3D image of an object from its projections leads to a mathematical inverse problem known as tomography. The Filtered Backprojection algorithm (FBP) is by far the most common algorithm for computing a reconstruction, as it is capable of computing accurate reconstructions with high computational efficiency, provided that high quality projections have been acquired for a sufficiently large number of angles, limiting its usefulness for real-time reconstruction in an industrial setting.

## Algebraic algorithms

Algebraic algorithms model the reconstruction task as an optimisation problem that is solved iteratively. Compared to FBP, algebraic methods have the advantages that limited data problems can be modelled accurately by adjusting the optimisation problem and that noise in the projections can be effectively averaged. Moreover, certain types of prior knowledge, such as non-negativity of the attenuation coefficients and image sparsity, can be effectively incorporated in the reconstruction algorithm using the regularisation function. However, at present, the computational requirements of algebraic algorithms are far too high for the real-time requirements of an industrial production setting.

## Algebraic filter methods

In recent work, we have demonstrated that in some cases the computational bottlenecks of algebraic methods can potentially be overcome by introducing a new class of hybrid reconstruction algorithms that combine the favourable computational properties of FBP with the power and versatility of algebraic methods [1,2]. Our new approach, which we refer to as algebraic filter methods, translates the analytical prop-

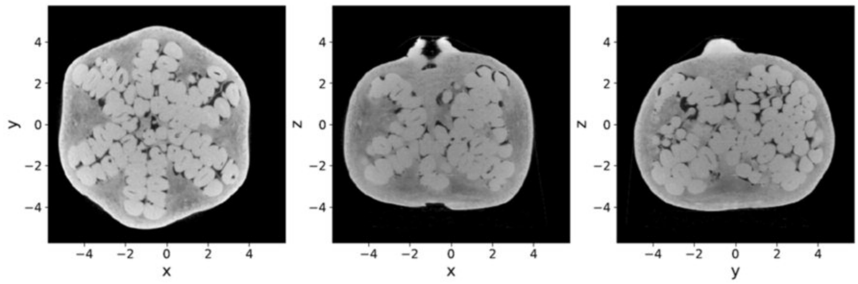



*Figure 1: 3D reconstruction of a pomegranate from high-resolution X-ray CT data, clearly showing the complex internal morphology and textures. This dataset is publicly available [3].*

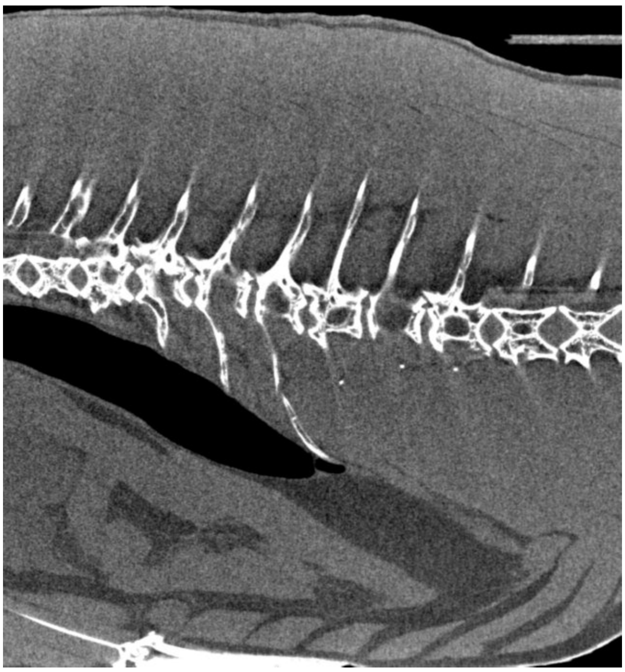


*Figure 2: Reconstruction of a small part of a fish, showing the bones, muscle and some organs. The 3D location of bones can be computed in near real-time (approx. 200ms). Picture: CWI.*

erties of the tomography problem into the domain of algebraic methods, replacing the generic matrix representation with convolution-based filter operators. The idea is to compute one or more filters h for the FBP algorithm in such a way that the resulting FBP reconstruction resembles the image computed by a more expensive iterative method. Surprisingly, it is possible to obtain image quality almost identical to that of iterative methods with the computational performance of FBP, making the approach suitable for industrial use in high-throughput, limited-data scenarios.

In May 2017, CWI started the FleX-ray lab, where a highly flexible X-ray CT system is combined with high-performance computing capabilities, making it possible to rapidly prototype novel computational methods for 3D X-ray imaging, and experiment with various geometrical scanning configurations. Due to their combination of computational efficiency and accuracy in limited data scenarios, algebraic filter methods can potentially enable the effective use of 3D imaging for quality control in a broad range of food products, which we are currently exploring further in ongoing research projects with high-tech partners from the food industry.

3D reconstruction of a pomegranate from high-resolution X-ray CT data, clearly showing the complex internal morphology and textures. This dataset is publicly available [3].

**References:**

[1] D.M. Pelt and K.J. Batenburg: "Improving Filtered Backprojection Reconstruction by Data-Dependent Filtering", IEEE Transactions on Image Processing, 23(11), 4750-4762, 2014

[2] L. Plantagie and K.J. Batenburg: "Algebraic filter approach for fast approximation of nonlinear tomographic reconstruction methods", J. Digital Imaging, 24(1), 013026, 2015

[3] DOI: 10.5281/zenodo.1144086

**Please contact:**
Joost Batenburg, Robert van Liere
CWI, The Netherlands
joost.batenburg@cwi.nl,
robert.van.liere@cwi.nl

# Security Threats and Risk Analysis of an IoT Web Service for a Smart Vineyard

by Massimo Borrelli, Vanes Coric, Clemens Gnauer, Jennifer Wolfgeher and Markus Tauber (FH Burgenland)

***The grape-growing industry is changing as growers increasingly combine technology with traditional growing methods in smart vineyards. If wine makers want to maximise the potential of their plants, it is no longer enough to rely on gut-feeling, but rather on locally gathered environmental data. These data help to accurately plan individual tasks, such as fertilisation, plant protection, and harvesting. This is where automated and IT-supported farming, or smart farming, comes into play.***

While smart farming is a significant trend with substantial literature in the areas of IoT and the Cloud, very little attention has been paid to security in this area. Carrara et al. [1] have established an IoT based management program collecting data on temperature and relative humidity. Zachariadis and Kaskalis [2] measured critical environmental measurements with the data being sent via SMS, and Patil and Thorat [3] monitored water, climate, pests and diseases of the grape and provided this information via a Webinterface. But none of these studies have addressed related security concerns.

To this end, we have developed a representative service oriented architecture (SOA) based application and conducted a structure security analysis to identify typical security issues in smart farming applications – in our use case, a smart vineyard. For such a use case it is important to get accurate and actual climate data from the vineyard. To achieve this, a device was established over a period of three months that collects local data in the vineyard. Information like temperature, wind strength, direction and soil humidity is then stored in a database and transmitted via a web service so that it can be accessed easily through a web interface. For this a prototype has been built with a Raspberry Pi v3 and a DHT22 temperature sensor and a humidity sensor. The Raspberry Pi uses the operating system Raspbian without GUI. An Apache2 web server is installed for the delivery of the weather information on a web interface. The database server that we use is MySQL where the data from the sensors is stored, as is the hardcoded data for wind speed and wind direction. A python script is used to collect the data from the sensor and to store this data into the database. PHP7 is also installed in the device which is used for the programming of the web service. Figure 1 shows a representation of the prototype, its assets and data flows.

The development of this prototype has helped us acquire an understanding of the impact of threats to the system components. The assets that have been considered for the security threats and risk analysis include: the Raspberry Pi, the database and the web service. Out of the CIA triangle (availability, integrity and confidentiality), availability is the most important asset, while confidentiality and integrity apply only to the data included in the database.

To critically assess the prototype, the OWASP Top Ten [L1] and OWASP IoT Top Ten [L2] catalogues have been consolidated to identify the threats that would affect such a solution. These catalogues have been chosen because of the IoT nature of our service, as well as being relevant to rate the web application security.

Each of the three assets have been assessed against all attack vectors included in these catalogues, to determine the major threats and the assets most likely affected, as showed in the following charts. The number of points given to each vector is based on the DREAD methodology for risk assessment. Ranging from 0 to 15, the figures 2 and 3 show the threats with the greatest impact on our assets (where not applicable a value of 0 was used).

The results of the security threats revealed that the following risks might impact the identified assets:
- Data loss and corruption is the main consequence of the various attack vectors affecting the database, undermining the integrity of the data
- A denial of service attack would bring the service down and thus poses a significant issue to the service availability
- The Raspberry Pi, being the device where the software is running and the sensors are connected to, needs to be protected against physical attacks that might remove storage media or access the software via USB ports. This would also affect the service availability and compromise its functionality.

Deploying a smart farming web service and relying on IoT technology is a fundamental step ahead to improve productivity within the vineyard industry, but the security aspects of the web service need to be considered as well. The topic of security needs to be further discussed by assessing different and much broader IoT setups. Understanding threats and the risks arising from the service infrastructure and IoT is the most relevant starting point to design safe solutions that reduce risks or alternatively help to minimise the impact of security breaches.


*Figure 1: Prototype, assets and data flows.*

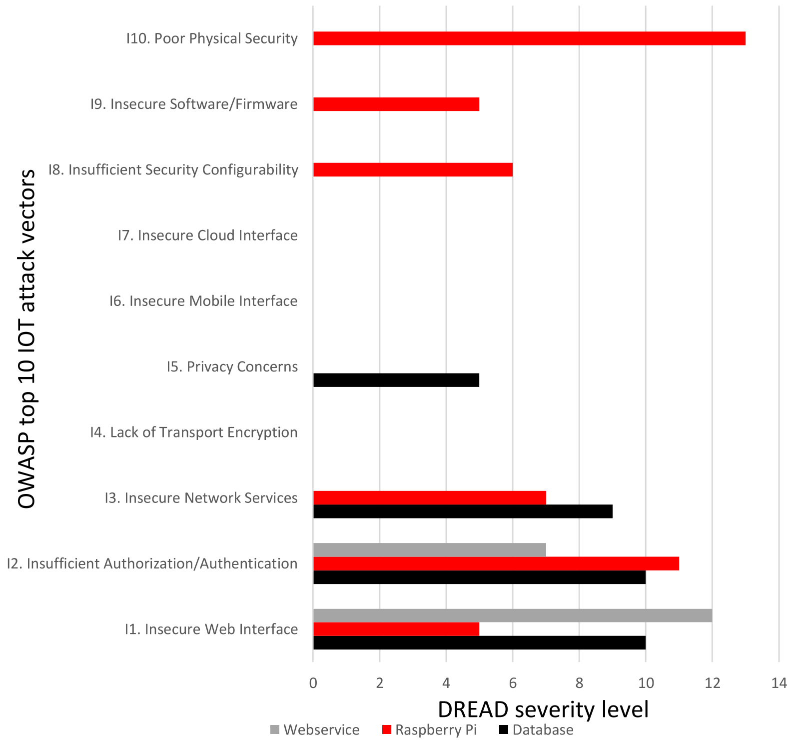



*Figure 2: Weighted impact of OWASP top 10 IOT attack vectors on prototype assets.*

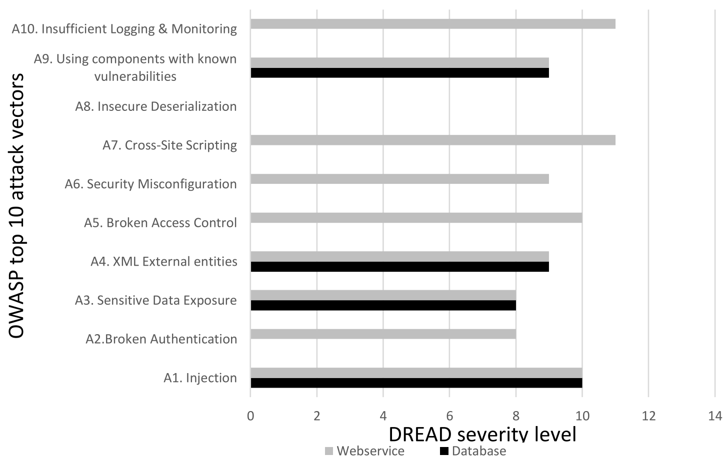



*Figure 3: Weighted impact of OWASP top 10 attack vectors on prototype assets.*

**Links:**
[L1] https://kwz.me/hbz
[L2] https://kwz.me/hbA


**References:**

[1] M. Carrara, et al.: "An innovative system for vineyard management in Sicily", Journal of Agricultural Engineering, 41(1), pp. 13-18, 2010.

[2] S. Zachariadis, T. H. Kaskalis: "An Embedded System for Smart Vineyard Agriculture", 2nd Pan-Hellenic Conference on Electronics and Telecommunications - PACET´12, Thessaloniki, Greece, 16-18 March 2012.

[3] S.S. Patil, S.A. Thorat: "Vinayard monitoring and recommendations using wireless sensor network: a study", International conference on computing, communication and energy systems (ICCCES-16), January 2016.



**Please contact:**
Massimo Borrelli
FH Burgenland, Austria
+43 664 3823105
1710781023@fh-burgenland.at

Markus Tauber
FH Burgenland, Austria
+43 5 7705 4321
markus.tauber@fh-burgenland.at


# A Research Agenda on Quantum Algoritmics

by Luis S. Barbosa, Alexandre Madeira (QuantaLab, and HASLab INESC TEC, Universidade do Minho)

***Quantum algorithmics is emerging as a hot area of research, with the potential to have groundbreaking impacts on many different fields. The benefits come at a high price, however: quantum programming is hard and finding new quantum algorithms is far from straightforward. This area of research may greatly benefit from mathematical foundations and calculi, capable of supporting algorithm development and analysis. The Quantum Algorithmics Agenda at QuantaLab is contributing by seeking a suitable semantics-calculus-logic trilogy for quantum computation.***

Software is a critical factor in the reliability of computer systems. While the development of hardware is assisted by mature science and engineering disciplines, software science is still in its infancy. This situation is likely to worsen in the future with quantum computer systems, as these will be more difficult to program and test. Moreover, quantum resources are critical. Thus, the need for correct-by-construction formal techniques [1] in quantum software development is even bigger than in classical computation. A lack of reliable approaches to quantum computer programming will put at risk the expected quantum supremacy of the new hardware.

Interest in quantum algorithmics has recently been boosted, with the identification of quantum technologies and computation as a strategic R&D domain for the EU. Over the last decade, the number of algorithmic techniques has grown significantly, exploring quantum effects, namely entanglement, in a number of ways. However, finding new and effective quantum algorithms has proved a challenging task. This entails the need for more mature mathematical techniques to provide rigorous semantics, effective program development calculi and suitable logics for specification and verification.

The class of quantum algorithms that encompasses classical control (deterministic, non-deterministic and probabilistic control structures) offers a wide scope of applications. This includes most common algorithms, as well as new applications to machine learning, computer graphics and quantum simulation.

The Quantum Algorithmics Agenda is a recently launched initiative at QuantaLab—the new collaborative research centre established by the International Iberian Nanotechnology Laboratory and the Universidade do Minho, located in Braga, Portugal. It brings together researchers in the broad areas of Quantum Materials and Technologies, covering experimental, theoretical, and computational research.

Our starting point is that, just like classical computation, quantum algorithmics will greatly benefit from a mathematically based approach, able to conceptualise, and predict behaviour, and to provide a rich, formal framework for specifying, developing and verifying quantum algorithms. Such foundations are produced in the confluence of several mathematical disciplines [2]: category theory (categorical logic, categorical semantics), logic (dynamic logic, lambda-calculus, linear logic, proof theory, type theory), theory of computation (complexity theory, information theory).

The goal of our research agenda is to develop such a framework and the corresponding mathematical techniques. We intend to fill a gap that we have identified in the literature at the following interrelated levels:
- Appropriate semantic structures, able to comply with the different types of classical control structures and quantum data, as well as to capture the notion of (quantum) program approximation;
- An algorithmic calculus consistent with the semantics above for the systematic derivation of quantum programs in a compositional way;
- A new family of dynamic logics, obtained in a parametric way [3] to support the formulation of contracts for quantum algorithms and their compositional verification;
- Proof systems for quantum logics, emphasising their role as type systems to ensure intentional properties of quantum algorithms, such as complexity and termination.

We are fully aware that such a broad goal cannot be achieved unless a particularly challenging application domain is selected as a testbed for theoretical results. Therefore, the construction of quantum convolutional codes was chosen as a case-study, not only because of its intrinsic difficulty, but also for its societal relevance. Modern societies depend heavily on the security level with which data is exchanged among complex, heterogeneous systems.

Quantum communication channels open the door for new, fast and secure communication schemes. For instance, in cryptography, the quantum key distribution protocol, that produces and distributes keys with an absolute guarantee of secrecy (relying on entanglement of quantum states) is already used and commercialised. Quantum error correction codes constitute simultaneously a rich source of semantic challenges and a vital tool in quantum technology, with a wide potential range of applications, from the protection of quantum information in noisy quantum channels, to the avoidance of decoherence in quantum computers.

**References:**

[1] R. Bird, O. de Moor: "Algebra of Programming. Series in Computer Science",.Prentice-Hall, 1997.
[2] B. Coecke (Ed.): "New Structures for Physics", Springer (Lecture Notes in Physics), 2011.
[3] A. Madeira, R. Neves, M.A. Martins: "An exercise on the generation of many-valued dynamic logics", Jour. Logical and Algebraic Methods in Programming 85(5), 2016.


**Please contact:**

Luis Soares Barbosa, QuantaLab, and HASLab INESC TEC, Universidade do Minho, Portugal
+351 253604463, lsb@di.uminho.pt


# An Innovative Technology for Vision Aids


by Carmelo Lodato and Patrizia Ribino (ICAR-CNR)



***Vitreoretinal diseases are widespread and on the rise. These pathologies can cause the degeneration of parts of the retina, and in particular of the macula. The visual perception, compromised by these ocular pathologies, can be improved with our non-therapeutic corrective technology, incorporated into a display device, which is viewed by the user in binocular vision.***


We are developing a technology [1] that improves binocular vision in people suffering from eye dysfunctions that cause the perception of defects in various portions of the visual field, in central or peripheral vision, when the user looks at a video played on a display device, as shown in Figure 1.

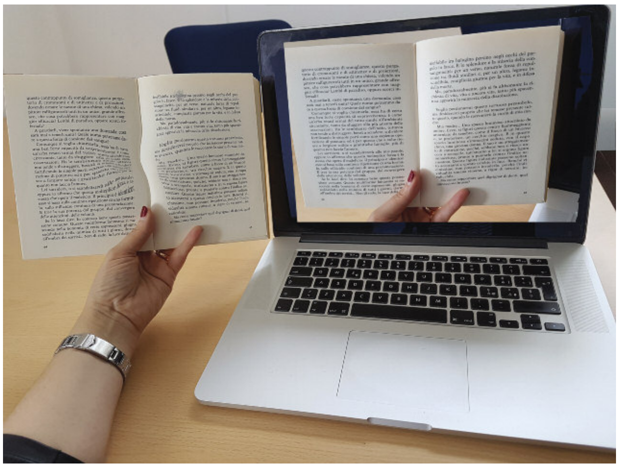


*Figure 1: A possible implementation of the technology: A computer helps a user with distorted vision to read a book by improving their visual perception.*

This type of disorder can derive, for example, from vitreoretinal diseases of diverse origins or traumas. Some ocular pathologies such as macular senile degenerations (AMD) are very widespread, especially in older adults. These pathologies can cause visual defects in various areas of the visual field and often compromise the central vision. For example, a person suffering from AMD might see a portion of an image as distorted or blurred, and, in especially severe cases, the perceived image may be partially deprived of visual information.

Therapeutic or non-therapeutic remedies are available. Therapeutic approaches are often invasive, and have a reduced success rate if not promptly performed. Conversely, non-therapeutic approaches may include vision aids that can improve vision without intervening in their physiological causes. Magnifiers or projectors, which improve vision by enlarging an image, are common vision aids for some daily activities, such as reading. Other more sophisticated solutions try to approach visual distortion issues [2, 3].

Our method uses techniques of 3D visualisation, eye tracking, and image processing (see Figure 2). It uses two corrective schemes, one for each eye, containing digital filters configured according to the visual defects of the user. Each corrective scheme is customised defining shape, size, and position of the image portions affected by defects. Moreover, sets of digital filters are associated with the portions of the image mentioned above to obtain various local effects. The binocular vision is improved by appropriately applying the corrective schemes to the video frames to be conveyed to the user's eyes.

A corrector device transforms an input video stream, coming from a video source, in two distinct flows applying the respective corrective scheme to each frame of the input stream. Each corrective scheme is applied in a dynamic mode congruently to the instantaneous variations of the visual fixation point and the position of the head revealed by the eye tracking system. A stereoscopic or autostereoscopic system, therefore, allows each video flow to be conveyed to the corresponding eye.

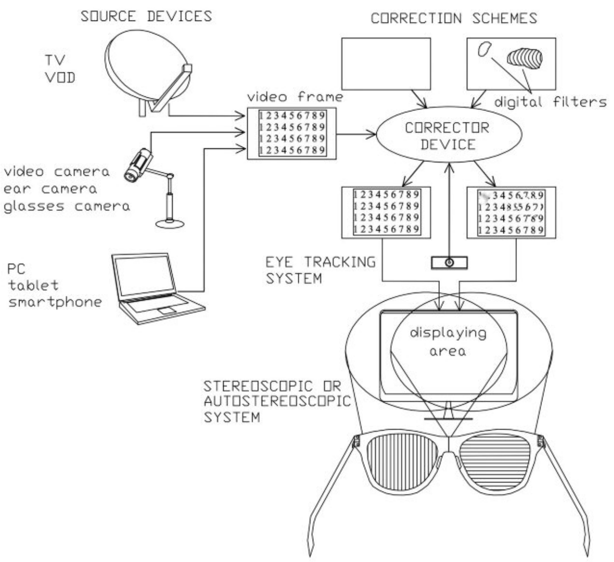



*Figure 2: An informal representation of the main system components.*

This solution can significantly improve quality of life, reducing the impact of the social costs and loss of productivity typically associated with this type of disorder. This is especially true for older people for whom even partial recovery of some daily activities in binocular vision could be of utmost importance. It is a non-invasive solution, and it allows the correct binocular visual perception even when more defects are localised in corresponding retinal points, providing that the visual impairment does not involve a loss of information. The technology can also significantly improve vision in the case of a loss of information, provided that the loss of information relates to non-corresponding retinal points.

The technology could be incorporated into vision aids consisting of an apparatus connected to a CCTV such as a magnifier for close up viewing, distance viewing and self-viewing, or in other devices that are connected to a television signal receiver or a computer's video output and provide vision aids for watching TV or using a computer. Other embodiments can give support for diagnosis and for monitoring the progression of the disease.

**References:**

[1] C. Lodato, P. Ribino. "Method for modifying at least one frame video of a digital video stream to be displayed on a displaying area and related system", International patent Application No. PCT/IT2017/000089.

[2] Kohn et al. "Characterization and correction of macular distortion" US Patent No. 8,708,495. Apr. 29, 2014

[3] Dalal et al. "Central vision impairment compensation" US Patent No. 2014/0210970. Jul. 31, 2014.

**Please contact:**

Carmelo Lodato, ICAR-CNR, Italy
+39 091 8031065
carmelo.lodato@icar.cnr.it

# Beyond Summaries for Ontology Exploration

by Haridimos Kondylakis, Georgia Troullinou, Kostas Stefanidis and Dimitris Plexousakis (ICS-FORTH)

***The recent explosion of the web of data and the associated Linked Open Data (LOD) initiative have led to an enormous number of widely available RDF datasets. These datasets often have extremely complex schemas, which are difficult to comprehend, limiting the exploitation potential of the information they contain. As a result, there is now, more than ever, an increasing need to develop methods and tools that facilitate the quick understanding and exploration of these data sources.***

To this direction, many approaches focus on generating ontology summaries. Ontology summarisation is defined as the process of distilling knowledge from an ontology in order to produce an abridged version. Although generating summaries is an active field of research, the generated summaries in most of cases are static, limiting the exploration and the exploitation potential of the information they contain. As a result, there is now, more than ever, an increasing need to develop methods and tools to help facilitate the understanding and exploration of various data sources.

RDFDigest [4] is a tool that was initially developed in the context of eHealthMonitor FP7 project, for summarising multifaceted linked health data, exploiting measures such as relevance and centrality to identify the most important nodes and novel algorithms for linking them together [3]. The tool was extended [1] within the MyHealthAvatar FP7 and the iManageCancer H2020 projects, including seven more measures for identifying importance according to various approaches in the literature. In addition, new algorithms were introduced, trying to minimise the additional non-important nodes that should be added for linking the selected

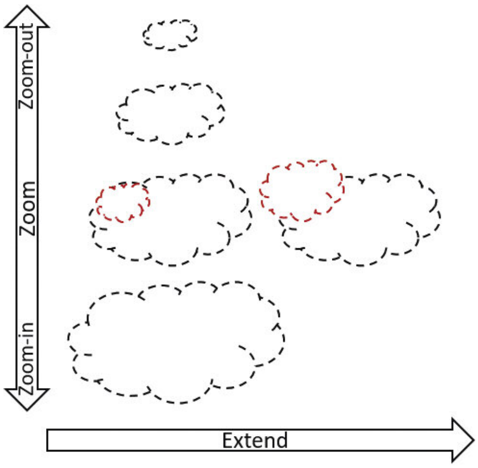



*Figure 1: Zoom and extend operators.*

most important nodes in the final summary produced. Although end-users can select among various importance aspects and linking algorithms, the user might still find the presented information overwhelming and would ideally see less information, focusing only on a specific subset of the presented nodes.

To this direction, the second version of the tool, RDFDigest+ will soon be released, enabling users to actively further explore data sources, starting from the summaries presented to them. Zoom-in and zoom-out operators, allow users to explore the contents of a data source at a higher or lower granularity level, respectively. Furthermore, users might want to have more detailed information not only on the whole schema graph but on a selected subset of it. This can be achieved by selecting some nodes and requesting more detail on those, extending a specific graph part. The extend operation, in essence provides additional nodes dependent on the selected ones, giving more details on the selected schema part.

From a different perspective, we also focus on how to explore changes in evolving ontologies [2]. Specifically, given the big number of ontologies that constantly evolve, there is a clear need to monitor and analyse the changes that occur on them. Traditional approaches for studying the evolution of data focus on providing humans with deltas that include loads of information. In our approach, we propose a processing model that recommends evolution measures taking into account particular challenges, such as relatedness, transparency, diversity, fairness and anonymity. Our goal is to support humans with complementary measures that offer high-level overviews of the changes to help them understand how data of interest evolves.

**References:**

[1] A. Pappas, G. Troullinou, G. Roussakis, H. Kondylakis, D. Plexousakis, “Exploring Importance Measures for Summarizing RDF/S KBs”. ESWC (1) 2017: 387-403.

[2] K. Stefanidis, H. Kondylakis, D. Plexousakis, “On Recommending Evolution Measures: A Human-aware Approach”, Proc. of DESWeb @ ICDE 2017.

[3] G. Troullinou, H. Kondylakis, E. Daskalaki, D. Plexousakis, “Ontology understanding without tears: The summarization approach”. Semantic Web 8(6): 797-815 (2017).

[4] G. Troullinou, H. Kondylakis, E. Daskalaki, D. Plexousakis, “RDF Digest: Efficient Summarization of RDF/S KBs”. ESWC 2015: 119-134.

**Link:**

[L1]: http://rdfdigest.ics.forth.gr

**Please contact:**

Haridimos Kondylakis, FORTH-ICS, Greece

kondylak@ics.forth.gr

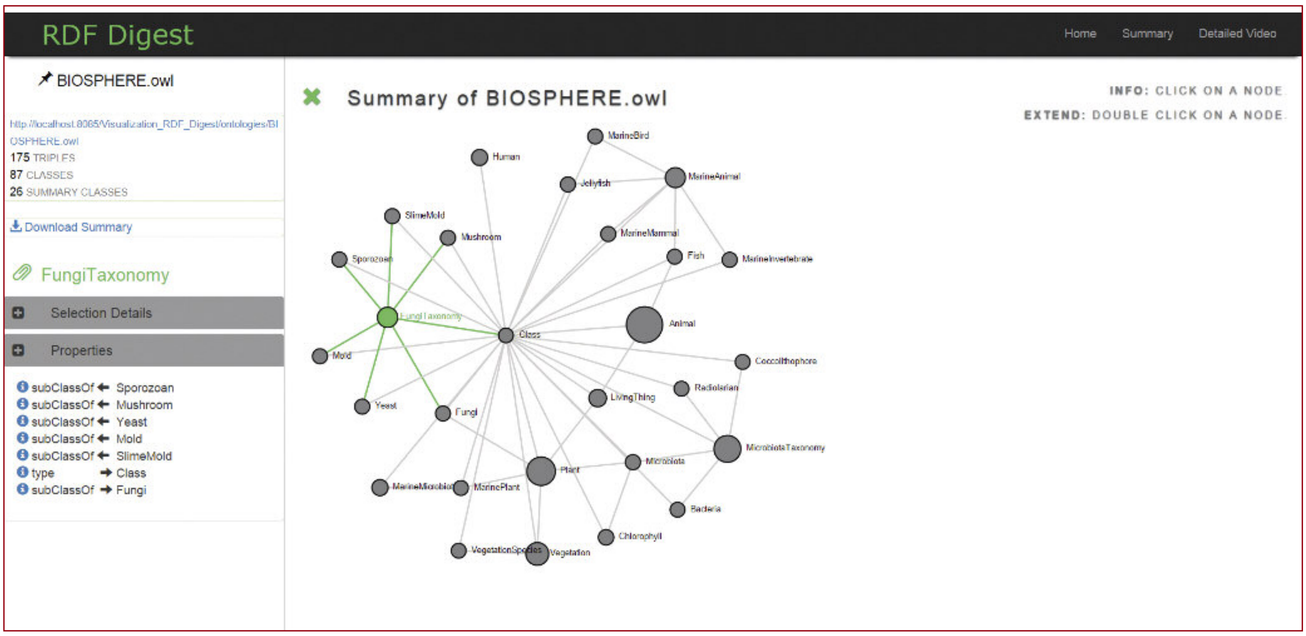



*Figure 2: The RDFDigest+ system.*

# Enabling 3D-TLC NAND Flash in Enterprise Storage Systems

by Roman Pletka, Nikolas Ioannou, Nikolaos Papandreou, Thomas Parnell and Saša Tomić (IBM Research Zurich)

***The sustained increase in storage density and simultaneous cost reductions of NAND flash memories have ensured their continued popularity. But these positive attributes are linked to less desirable trends such as a reduction in the reliability of memory cells, requiring an increasingly overwhelming amount of sophisticated mitigation techniques, which become most apparent when applied to enterprise-level storage systems.***

The continuous success of NAND flash is a result of the enormous progress achieved over the past few decades. Initially driven by the scaling of the process technology node, then assisted by an increase in the number of bits stored per cell, today, the aggressive stacking of layers in the third dimension continues to preserve this trend. Read and write operations are performed on a page granularity – where the page size is currently in the order of 16KiB – with several hundred pages being organised into a block. But an entire block must be erased prior to being programmed. Therefore, Flash controllers run a Flash translation layer (FTL), that transforms "writes" into a sequential write stream and maintains the mapping between logical and physical pages.

An increasing number of physical variables affect the reliability of NAND Flash memory cells [1]: Charge being trapped in the insulating tunnel oxide because of repeated program-erase cycling is a permanent effect that wears out cells and reduces their health, resulting in endurance limitations. Temporary effects influence the raw bit error rate (RBER) only up to the next erase operation but tend to be more pronounced as the program-erase cycle (PEC) count increases. Typical temporary effects are: (1) Cell-to-cell interference upon programming causing charge to be falsely deposited on adjacent cells due to parasitic capacitance coupling. (2) Retention limitations, where a cell gradually leaks charge after being programmed. This effect is accelerated by the amount of charge trapped, but also depends on the number of electrons stored in the cell. (3) Read disturb happens upon a read operation, where a high voltage is applied to all unselected wordlines in a block. As this voltage must be higher than the highest programmed threshold voltage, it has essentially the same effect as a tiny program impulse and therefore causes a slight shift in the threshold voltage. (4) The accumulation of static charge over time evokes an electric field that temporarily shifts the threshold voltage distributions. Only after reading a block, is the built-up charge removed. The influence on the threshold voltage distributions for the most important effects are graphically illustrated in Figure 1.

All these effects coexist in NAND Flash-based storage devices and cause widening as well as shifts of threshold voltage distributions. Today, up to four bits are stored in a single flash cell resulting in 16 charge levels that must be programmed accurately and distinguished properly upon reading. Thus, the threshold voltage distributions are very tight and tend to overlap even in fresh devices. Strong error correction codes (ECC), capable of correcting an RBER of up to 10-2 are essential to address the widening of the threshold voltage distributions but they are not sufficient.

At IBM Research - Zurich, we are exploring novel ECC algorithms and flash-management techniques that address these challenges holistically. We have developed key technologies including block calibration, a technique to determine optimal read voltage levels in the background, novel garbage-collection, data-placement, and wear-leveling algorithms that balance block health variability and – together with our hardware compressor – reduce overall write amplification.

Our technologies tightly integrate data path processing directly implemented in hardware with control path activities in firmware inside a flash card or SSD to achieve continuous mitigation in the background (Figure 2).

Because of the effects mentioned above, the threshold voltage distributions shift in different directions. Therefore, most flash memory devices support read-retry operations that apply a sequence of different threshold voltages in an attempt to lower the RBER. But consecutive reads result in a proportionally higher read latency, an unwanted effect in enterprise storage systems causing significant application-level tail latency increases.

Prospective read testing facilitates unbiased reads by ensuring that built-up charge is removed and determines

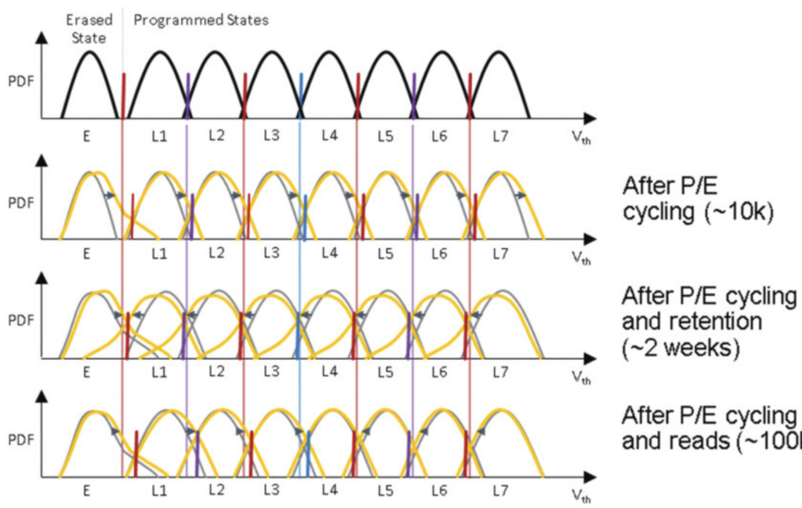



*Figure 1: Effects on the threshold voltage distribution Vth of a 3-bit per cell (TLC) NAND Flash memory cell from program-erase cycling, retention, and read disturbs.*

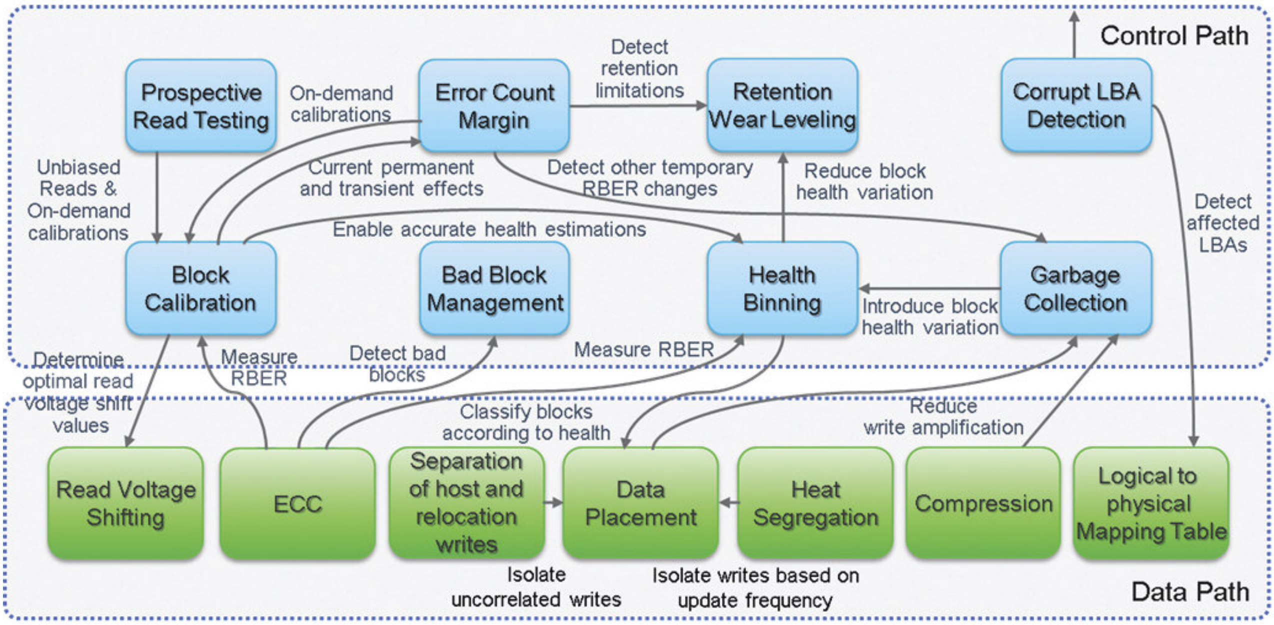



*Figure 2: Overview and interactions of 3D TLC NAND flash management functions.*

blocks that need their threshold voltage shift values adapted by calibration. The block calibration engine uses the RBER measured by the ECC during background read scrubbing to determine the current set of optimal read voltage shift values [2]. The measured RBER is also used by health binning to classify blocks according to their health as well as to detect bad blocks [3]. The current permanent and transient effects on a block together with large-scale characterisation data are used to determine an upper bound of the RBER that is applied to trigger block re-calibrations and to detect early retention limitations or other excessive temporary effects on the RBER.

It has been shown that real-world workloads tend to be skewed in their access-pattern distributions. We leverage the separation of host and relocation writes as well as the segregation of writes based on their update frequency in the data placement to minimise the amount of relocated data in skewed write workloads. Both data placement and compression help to reduce internal write amplification, i.e., additional writes needed to relocate still valid data from blocks that are going to be erased such that they can be used to accommodate future writes.

At the same time, we place write hot data on the healthiest available blocks and write cold data on less healthy blocks, resulting in an efficient balancing of block wear depending on the current block state.

The compression engine, built as a streaming engine in hardware, is based on Lempel-Ziv coding combined with pseudo-dynamic Huffman codes and a 4KiB history buffer matching the logical block address (LBA) size and it can deliver inline compression at more than 1M IOPS at the system level. Finally, due to compression, a physical flash page can hold a large amount of compressed logical pages. In the rare event that the ECC cannot correct the data read, all logical pages stored in that physical page will be affected even though only a single page may have been read. Due to the log-structured organisation of data, affected logical pages are typically not adjacent and need to be detected to prevent silent data corruption. This is also applicable when data is RAID-protected by the SSD controller and there are more physical page failures than can be corrected by the given RAID level. Our algorithms allow such error events to be recovered using the array-level data protection layer and hence operate as a cooperative two-dimensional RAID scheme.

We have been able to demonstrate that using our technologies we can increase the device endurance by up to 15x without affecting performance and using acceptable additional resources only. Furthermore, we have reduced the host read retry probability by more than four orders of magnitude and lessened the tail latencies of the 99 percentiles by up to 2.7x compared to traditional wear leveling algorithms. We implemented and verified our techniques in a simulator as well as the IBM FlashSystem® all-flash array. The simulator enables us to evaluate scenarios that exceed the capabilities of the hardware environment.



**Link:**
https://www.zurich.ibm.com/cloudstorage/

**References:**

[1] Y. Cai et al., Error Characterization, Mitigation, and Recovery in Flash-Memory-Based Solid-State Drives, Proc IEEE 2017
[2] N. Papandreou et al., Using Adaptive Read Voltage Thresholds to Enhance the Reliability of MLC NAND Flash Memory Systems, GLSVLSI, 2014
[3] R. Pletka and S. Tomić, Health-Binning: Maximizing the Performance and the Endurance of Consumer-Level NAND Flash, SYSTOR, 2016


**Please contact:**
Roman Pletka, IBM Research Zurich, Switzerland
rap@zurich.ibm.com

# CasAware: A Semantic-based Context-aware System Enabling Ambient Assisted Living Solutions

by Daniele Spoladore, Enrico G. Caldarola, Atieh Mahroo and Gianfranco E. Modoni (ITIA-CNR)

***The CasAware research project combines context awareness with ambient assisted living by using a semantic-based framework that enables the smart home paradigm for everybody.***

In traditional home design, collaboration among appliances is often neglected: each appliance performs its tasks and provides its services in isolation, with very limited exchange of information – or none at all. This approach hinders that development of home automation and customisation of domestic services embodied by the smart home paradigm. In fact, in the smart home, domestic appliances should cooperate in a distributed and interconnected system, by acquiring context data and processing them in order to provide customised services. In this context, the CasAware project tries to implement the smart home paradigm exploiting a semantic-based system grounded on a set of ontologies named “CasAwareOnt”.

By exploiting first-order logic based languages (RDF and OWL), the CasAware’s system leverages the formal representation of concepts and relationships belonging to different knowledge domains, such as domestic sensors (e.g., CO2 concentration, indoor temperature, indoor illuminance, etc.), the expected behaviour of appliances (programs, actuator responses, etc.) and residents’ preferences (when and how to activate specific programs, preferred comfort metrics, etc.). This knowledge is formally represented in the CasAwareOnt and is expected to grow over time, according to the modifications of residents’ preferences – which are identified by monitoring inhabitants’ interactions with the appliances and the house. The ability to detect new preferences so that they can be suggested to the residents and added to the CasAwareOnt (thus dynamically updating the knowledge-base), is achieved by exploiting machine-learning techniques in addition to semantic reasoning capabilities. The inhabitants can interact with the system via an app, which acts as the main interface and can be operated via tablet or smartphone, allowing remote management of the house as well.

Residents can take advantage of CasAware to help manage several domestic tasks, such as supervising the energy consumption of the appliances and preventing unexpected power-cuts: in fact, CasAwareOnt provides a description of the energy required by each appliance and programs to run and matches the current consumption with the set power threshold. If the activation of a new device generates a power-cut by surpassing the power threshold, the application warns the dweller and prevents the device from turning on. Notifications via app can also be used to warn the inhabitants when potential energy wastage is detected – for instance, if the A/C system is on and a window is accidentally left open. The system can also access several open-data sources to provide tailored services and help the inhabitants with their daily activities. For instance, the CasAware system is aware of the inhabitant’s habits, such as wake up time and time required to get to the workplace and the route usually taken. The CasAware system is able to acquire information regarding the traffic and weather conditions, then analyse traffic delays and set the alarm to go off early to prevent the dweller from getting to work late.

CasAware is able to monitor the dwellers’ behaviour and to provide customised services relating to their safety in the house, appliance management and energy consumption management. CasAware can also manage the residents’ health-related data: this can be particularly interesting for ambient assisted living (AAL) solutions. In fact, AAL aims at providing elderly people or those with disabilities with a set of services and appliances capable of anticipating and responding to the specific needs of the inhabitants, taking advantage of tailored services provided by diverse devices. CasAwareOnt can be enriched with the knowledge regarding the physiological status of the dwellers – formalised using

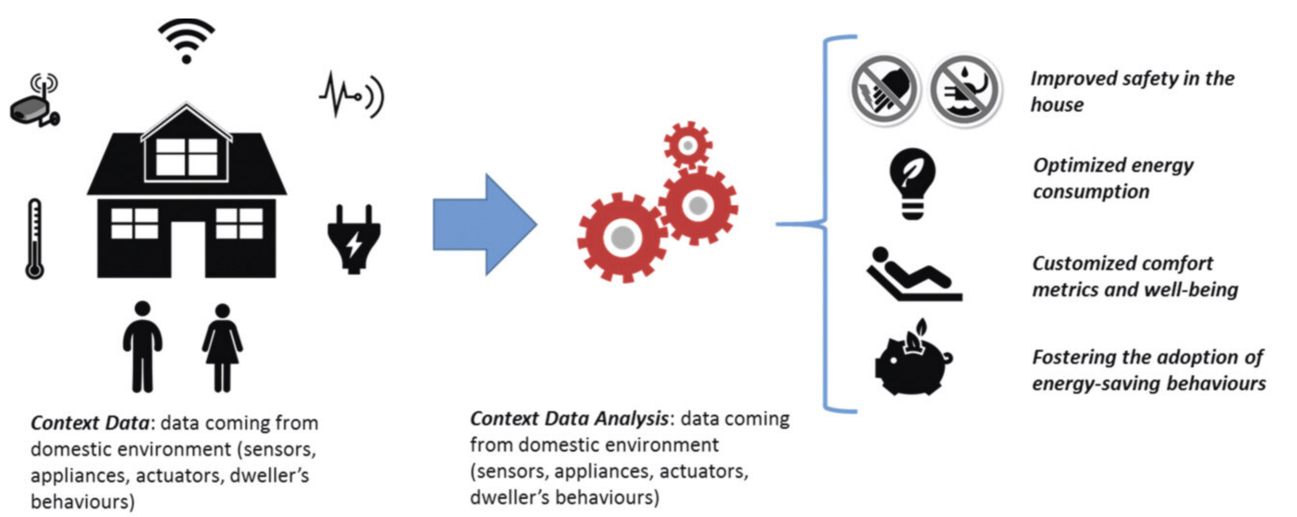



*Figure 1: Overall approach of CasAware.*

the World Health Organisation's International Classification of Functioning, Disability and Health – while semantic reasoning can be leveraged to provide the most suitable solution for each inhabitant, according to the individual's needs [1].

CasAware continues the work of the "Design For All" project [2] by enhancing the semantic-based framework and extending its use to real-time monitoring of the domestic environment, while preserving and specialising the AAL vocation of these solutions. The whole system benefits from the seamless integration of the devices leveraged by the semantic-based architecture [3], while the monotonicity of semantic reasoning is overcome by adopting machine-learning techniques. A pilot CasAware system is being developed in the CNR's Living Lab, which will allow us to acquire invaluable insights into potential issues and bottlenecks and to test solutions in new scenarios.

CasAware is a research project approved by Lombardy Region (ID 147152) involving the Institute of Industrial Technologies of Automation (ITIA) of the Italian National Research Council with two ICT companies – Teorema Engineering and S@it.

**References:**


[1] D. Spoladore, "Ontology-based decision support systems for health data management to support collaboration in ambient assisted living and work reintegration". In Working Conf. on Virtual Enterprises, Springer, Cham.

[2] D. Spoladore, S. Arlati, and M. Sacco, "Semantic and Virtual Reality-Enhanced Configuration of Domestic Environments: The Smart Home Simulator", Mobile Information Systems.

[3] M. Sacco et al., "Supporting the design of AAL through a SW integration framework: the D4All project", in International Conf. on Universal Access in Human-Computer Interaction.


**Please contact:**


Daniele Spoladore, ITIA-CNR, Italy
daniele.spoladore@itia.cnr.it


# Durchblick - Combining Sensor Outputs and Supporting Forensic Investigations in Robot-Assisted Analysis of Suspicious Objects


by Michael Sonntag (Johannes Kepler Universität Linz) and Michael Hofstätter (AIT)



***Terrorism represents a serious threat to human life and infrastructure, and it is vital that police officers are equipped to investigate suspicious luggage, rubbish bins that might have been tampered with, etc., for improvised or unconventional explosive devices. Common investigation techniques currently include video cameras and transmission x-rays on mobile robots. Modern and advanced sensors are being incorporated into a prototype to allow more detailed investigations, while simultaneously ensuring the data is usable as evidence.***


When a multitude of sensors are contained within a robot that investigates suspicious objects, it is important to be able to combine their output: E.g. if chemicals or radioactive material are detected, it is necessary to know their location in relation to the visual video. Similarly, infrared pictures or x-rays must "overlay" the normal video. As the robot will often be needed to operate in areas where no detailed 3D-plans of the building or outdoor area exist, acquiring an environment model and the exact localisation of the robot are also necessary for efficient handling. All this requires the fusion of a multitude of sensors outputs to rapidly produce an integrated display.

The Durchblick projects aims to create a prototype of such a system, based on robots already used by the police and the military, by extending them with new sensors and a significantly enhanced base station (see Figure 1). The first part of the project is therefore to investigate physically mounting multiple sensors (with the option for quick exchange of the sensor loadout) and integrating their communication. Based on this a prototype with selected sensors will be built and tested. Because of size and weight, some sensors will not be carried by the robot for the tests (mass spectroscopy), but with future miniaturisation or more powerful platforms, these might enter practical use too. Another part of the project is the communication with the base station, where the fusion of the sensor data will take place. A large amount of data needs to be transferred quickly (e.g. a laser scan of the surrounding environment (e.g. room or the local area) might require hundreds of MB), but at other times a continuous stream is necessary (e.g. visible/IR video). Because of the high bandwidth that is sometimes required the prototype will use a glass fibre cable for communication. To be independent of the transmission medium and fulfil security properties, a separate encryption layer is necessary. To simplify its use, secure USB modules (physically tamper-proof) are under

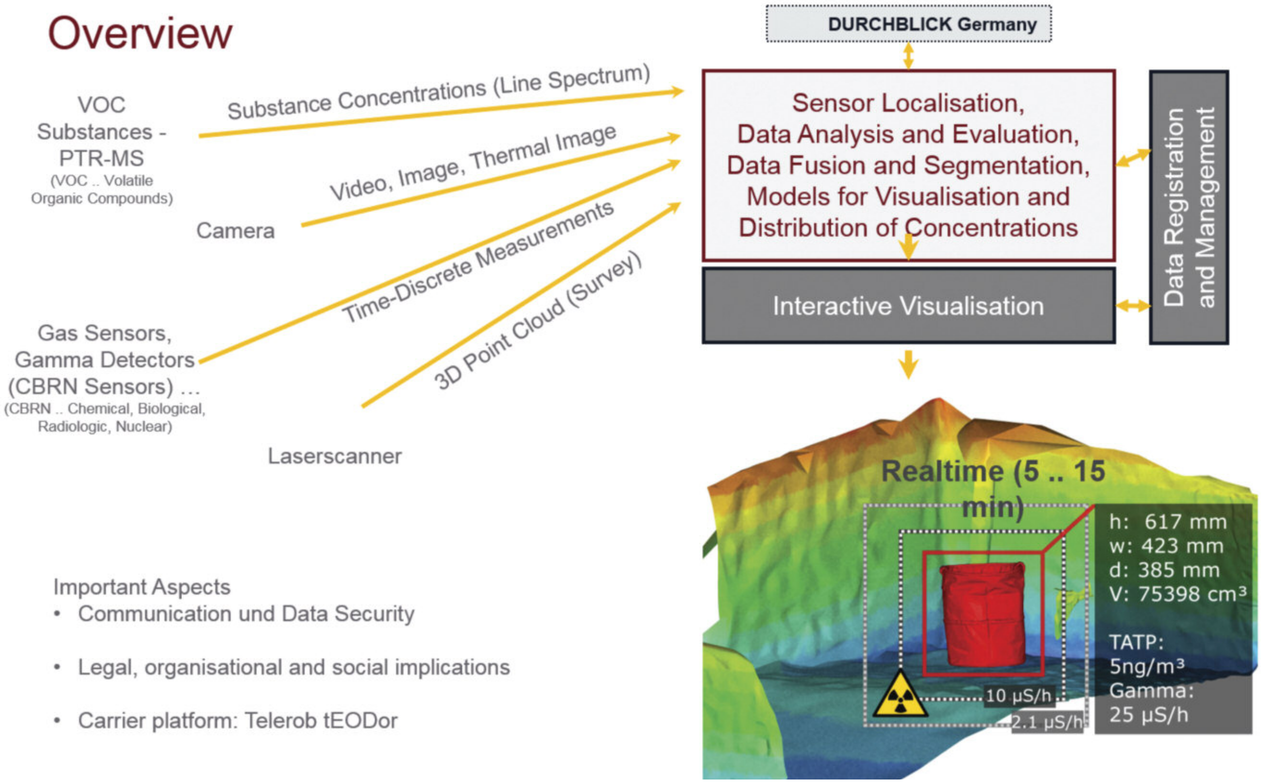



*Figure 1: Overview of the project.*

investigation to distribute/install the keys needed at the various elements.

The individual sensor data is combined at the base station and not the robot, as this requires significant computing power (weight and power on the robot are limited). This includes not only the visualisation of individual sensors, but also the combination of several sensor's outputs. An additional goal is the improvement of the human-machine interface. Investigating/defusing is a stressful situation and important data must be displayed prominently, while not producing additional stress or sensory overload. Any elements detected as dangerous must be clearly highlighted in all kinds of display. An additional complication is that operators might wear heavy protective gear, limiting manual dexterity as well as perception.

A final task of the project is to investigate gathering and preserving evidence, i.e. forensics support. This is particularly important because little may be left of a bomb following an explosion, and equally, it is important to be able to provide valid justification in situations in which harmless luggage is investigated or even damaged/destroyed. For this a separate forensics server is planned (see Figure 2). All data transmission to and from the robot will be stored as a network dump. This is not directly usable but enables detailed investiga-

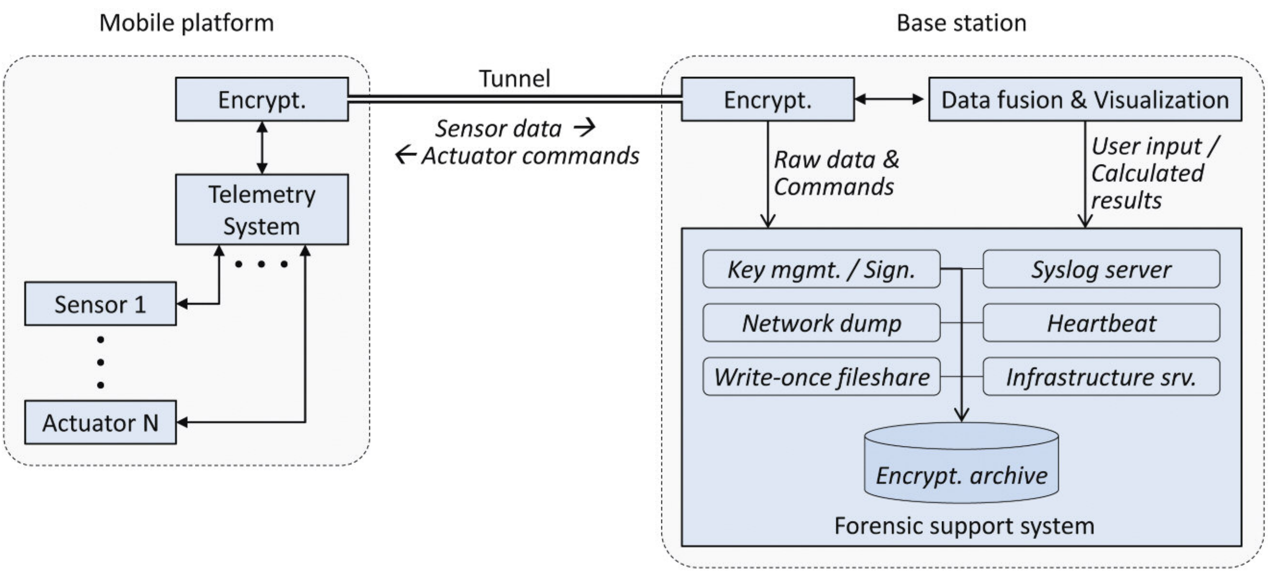



*Figure 2: System with detailed view on the forensic support.*

tions. As the sensor protocols and commands employed are unencrypted (security is provided by an additional layer; see above), this is feasible. All sensors are accessed via this network, so their presence is also automatically identified (logging of DHCP, regular scans for devices etc). Because this only provides access to raw data, but no sensor data fusion output, for example, these results will have to be added and stored as files. For security and privacy reasons encrypted and signed data storage to prevent modifications and ensure confidentiality are investigated. Again, for simplified use of the security functions, small hardware security modules might be employed.

In conclusion, existing technology needs to be integrated to become usable in this field of work, whereby numerous outside restrictions (size, weight, power, security, evidence etc) need to be fulfilled. The Durchblick project also cooperates with a similar German project (of the same name), which focuses on a larger set of different sensors, but leaves out data fusion and forensics as a second focus.

This project has been funded by the Austrian security research program KIRAS of the Federal Ministry for Transport, Innovation and Technology (bmvit) (Grant Nr. 854756).

**Please contact:**
Michael Sonntag, Johannes Kepler University Linz, Austria
+43 732 2468-4137, michael.sonntag@ins.jku.at

Michael Hofstätter, AIT Austrian Institute of Technology, Austria, +43 50550-4202, michael.hofstaetter@ait.ac.at

# Engineering Cryptography for Security and Privacy in the Cloud

by Stephan Krenn, Thomas Lorünser and Christoph Striecks (AIT)

***A large body of research in cryptography was established over recent decades but little of it has been commercialised. This is to the detriment of cloud computing, where the new methods would help to establish end-to-end security, which is not yet common practice.***

Cloud computing has made its way into all domains and remains one of the major growth areas in information and communication technology. It will also form the main backend platform of the internet of things and therefore play a crucial role in information society. Nevertheless, it still suffers from its outsourcing paradigm and intrinsic security and privacy problems.

Cryptographic research has made substantial progress over recent years and now provides a portfolio of mature cryptographic primitives and protocols suitable for addressing several of these problems in an effective and efficient way. Nevertheless, there still exists a substantial gap between what is possible and what is actually available in the cloud.

In PRISMACLOUD [L1] and CREDENTIAL [L2] we analysed the situation, extracted main inhibitors and developed promotors for the use and integration of cryptography in engineering work flows of interdisciplinary projects.

## Inhibitors

Going from specification to implementation requires substantial cryptographic knowledge to be able to translate the very compact and abstract specifications provided in the literature and very often they focus on formal specifications and proofs, but leave real-world aspects unspecified. Furthermore, implementing cryptographic schemes also requires non-standard skills from software engineers to exclude side-channel attacks (e.g., timing attacks). Overall, this makes the secure, sound, and efficient realisation of cryptographic protocols published in the literature a time-consuming, expensive and potentially error-prone task.

While even theoretical cryptographic research is often motivated with applications, typically no market-driven requirements engineering is performed. Specifically, the motivation aims at other cryptographers rather than practitioners, and therefore does not address real-world problems adequately. In fact, the gap between cryptographic theory and practice used to be huge.

While security can be used as market differentiator, we see little commercial interest in comprehensive practical usage of cryptography for this purpose. This is because the used technologies are mostly hidden from the user for best usability and there is normally no gain in functionality that is visible to a user, hence, the added value is hard to grasp.

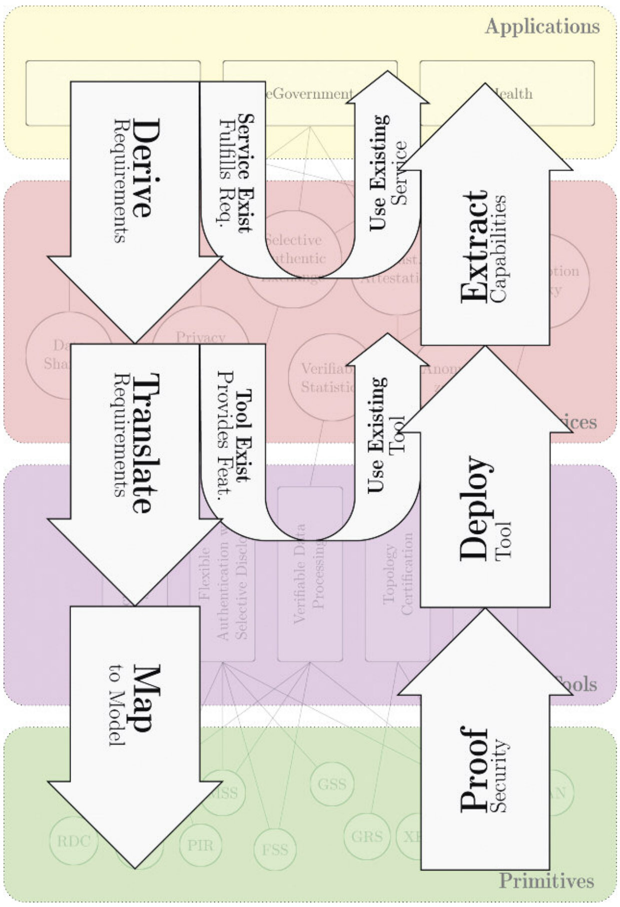
Establishing trust into novel algorithms before they get adopted by industry often takes significant amounts of time, e.g., because of security engineers waiting for them to be standardised. Although a conservative approach makes sense, many results are already mature enough to be picked up for commercial use, especially in the field of public key cryptography, where rigorous security proofs can at least partially replace the test of time.

Finally, due to the missing knowledge about novel technologies, it is not easy for developers and architects to understand what could be possible and which solutions really have practical relevance from the large amount of published work in the literature.

## Promotors

Systematically addressing the multidisciplinarity in our projects required the development of a system architecture as a first step. It was important to encapsulate expert knowledge in different domains ranging from theoretical cryptography, up to service and application development.

The architecture was then used as a basis for a modern tool and service based engineering approach, which was essential for integration into modern software development life cycles. A tool in our context is a well specified combination of cryptographic methods which serves a specific purpose




*The engineering methodology overview*

and comes with readymade cryptographic software and hardware components to system developers. The tools must be easily configurable and usable without cryptographic knowledge, only support a limited set of options and, ideally, already come preconfigured for most typical application scenarios. Additionally, by integrating the tools into ready-made security enhanced cloud (micro-)services it became even easier for developers to integrate them into applications. By doing so, we could further reduce the security knowledge required for system architects and developers to apply the cryptography and therefore significantly lower the entry bar for use in applications.

Although standardisation can be a tedious process, it provides multiple opportunities for an accelerated uptake of cryptographic cloud solutions. However, for best proliferation of the technologies the full architectural stack must be addressed in the standardisation efforts. In our case, it was not enough to initiate standards for new cryptographic schemes, we also had to make sure that the new cryptographic capabilities became visible in cloud service level agreements (SLAs), as the ISO/IEC 19086 series.

Another issue of importance is usability for all stakeholders, which are end users, application designers, cloud service designers as well as cryptographic tool designers who need to share a common understanding of the basic technology. The tool designers and the cloud service designers need to understand the requirements of the end users, while the cloud service designers need to understand the capabilities of the cryptographic tools so that they employ them in the intended way. These complex requirements can efficiently be addressed by using design patterns, and specifically cloud security and privacy design patterns, as well as human computer interaction (HCI) patterns.

This short list of promotors turned out to be the most effective approaches and procedures during our project work in recent years and helped to mitigate many of the main roadblocks mentioned above.

**Links:**
[L1]: https://prismacloud.eu,
https://at.linkedin.com/in/prismacloud
twitter:@prismacloud, http://twitter.com/prismacloud
http://cordis.europa.eu/project/rcn/194266_en.html
[L2] https://credential.eu/

**References:**
[1] T. Lorünser, S. Krenn, C. Striecks, T. Länger: "Agile cryptographic solutions for the cloud", Elektrotechnik und Informationstechnik 134(7), pp. 364-369 (2017)
[2] F. Karegar, et al.: "Opportunities and Challenges of CREDENTIAL - Towards a Metadata-Privacy Respecting Identity Provider", Privacy and Identity Management 2016, pp. 76-91
[3] T. Lorünser, D. Slamanig, T. Länger, and H. C. Pöhls: "PRISMACLOUD Tools: A Cryptographic Toolbox for Increasing Security in Cloud Services", ARES 2016, pp. 733-741

**Please contact:**
Thomas Lorünser
AIT Austrian Institute of Technology GmbH
+43 664 8157857, Thomas.Loruenser@ait.ac.at

# A 3D Modelling and Photorealistic Simulation Pipeline for 3D Heritage Edifices

by Maria Pateraki and George Papagiannakis (ICS-FORTH)

***Methods and technologies for creating photorealistic 3D models have undergone an impressive evolution with a positive impact in the cultural heritage sector. Using widely available standard tools we describe a modelling and simulation pipeline, which generally provides for most visualisation projects with an acceptable trade-off between visual consistency, accuracy and automatism.***

The creation of 3D models and the assumed simulation process for cultural heritage applications depends on the individual application requirements and mixed reality presentation method of choice [1, 2]. Primarily these are related to the geometric combined with photometric accuracy of the reconstructed 3D models as well as the level and number of details of the final product based on existing or, as in our case, not fully existing and partially reconstructed monument. Furthermore, the final 3D model portability in different platforms, the production cost and model size efficiency all need to be considered [3]. In this project we focused on the 3D reconstruction of the Macedonian tomb of Phillip II, one of the royal tombs excavated in the ancient city of Aigai in Vergina, Greece. The 3D model was created to give the interested user a sense of the geometry of the volume and spaces of the edifice and not necessarily a faithful and detailed representation, emphasising the aesthetic part of the 3D model. This figurative volumetric model resembles a digital mock-up, and can be experienced in an offline rendering or interactively rotated in 360 degrees, exhibiting a high-degree of realism and presenting the special characteristics of the materials.

Drawings and plans, provided by the IZ Ephorate of Prehistoric and Classical Antiquities, were used as the main source of information and the polygonal modelling technique was the preferred choice for the 3D model creation. Both the interior and exterior parts of the Tomb were modelled using SketchUp as well as MeshLab and Autodesk Maya tools. Special emphasis was given to modelling details of the façade and columns. The selection of materials and images for texturing was based on the analysis of textual descriptions and a collection of images retrieved from the web portal of the Ministry of Culture and accredited archaeological sources. Texturing of the individual surfaces was also performed in the above tools. For texturing the façade, in addition to the material library, a single image was used that was taken from the real scene. In order to achieve optimal results the original image was cropped into smaller images corresponding to individual planar or cylindrical surfaces of the façade, and draping was performed for each surface/image part.

For the photorealistic rendering of the reconstructed scene, we employed the open-source Mitsuba renderer [L1]. The reason behind this choice was that this modular, portable C++ renderer, implements latest unbiased as well as biased rendering techniques, and contains heavy optimisations targeted towards current CPU architectures. After several experimentations, we opted for the Monte Carlo path tracing technique, solving the complete rendering equation, using the generic parallelisation layer of the renderer. Thus, we were able to achieve fast frame renderings, suitable for a complete movie production that otherwise would demand a large render-farm to complete on-time with other commercial, production-ready renderers.

In order to allow efficient interchange of the various 3D scene elements between the different modelling and reconstruction software packages involved in the complete pipeline, we employed the Khronos Group open, COLLADA scene file format.

The figures below display frontal and side views of the different versions of the 3D model using Monte Carlo path-tracing global illumination. These include the untextured geometric version (first row) and the textured versions based on the material library (second row) and the image draping (third row) as part of a production movie clip with a sample camera fly-through around the 3D scene of the reconstructed edifice.

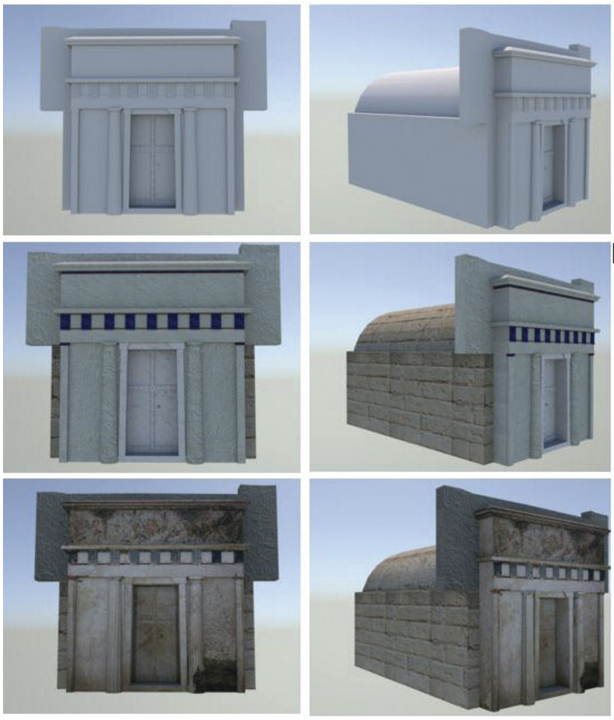


*Figure 1: Different versions of the 3D model of the Macedonian tomb using Monte Carlo path-tracing global illumination.*

A modelling and simulation pipeline was presented related to the creation process of photorealistic 3D reconstruction of heritage edifices, using widely available standard tools. However, it is important to stress that the optimal choice of the modelling strategy to be followed depends on the specific objectives and the complexity of the heritage item being reconstructed.

**Link:**
[L1] www.mitsuba-renderer.org

**References:**

[1] M. Ioannides, N. Magnenat-Thalmann, G. Papagiannakis, Eds., Mixed Reality and Gamification for Cultural Heritage, Springer, DOI: 10.1007/978-3-319-49607-8, 2017.

[2] E.P. Baltsavias, A. Gruen, L. Van Gool, M. Pateraki, Eds., Recording, Modeling and Visualization of Cultural Heritage. ISBN 0 415 39208 X, Taylor & Francis, Balkema: London, 2006.

[3] F. Remondino, S. El-Hakim, Image-based 3D modelling: A review, The Photogrammetric Record 21 (115), 269-291, 2006.


**Please contact:**

Maria Pateraki , George Papagiannaki
ICS – FORTH, Greece
pateraki@ics.forth.gr, papagian@ics.forth.gr

# Having Trouble Deciding Which Wine to Order? Try Le Sommelier Recommendation System

by Nikoleta Tsabanaki, Theodore Patkos and Dimitris Plexousakis (ICS-FORTH)

***Le Sommelier is a recommendation system that helps users choose one or more different bottles of wine that pair nicely with their joined orders. This is a frequent issue met by restaurant customers who have no specialised knowledge on the subject, and requires the satisfaction of a combination of complex rules. Our system represents, to a large extent, the expert's knowledge of the field in the form of both strict and weak prioritised rules. In addition to the formalisation of general empirical knowledge, an important feature of the system is the utilisation of users' preferences. Our system's research contribution is the modelling of a complicated combinatorial problem in an accessible way and the execution of reasoning that can be fine-tuned with user-friendly defined adjustments based on the customers' or restaurateur's preferences and needs.***

Le Sommelier is simple for users, since the interaction is mainly based on clicks over the virtual restaurant menu and wine list (Figure 1).

The system supports multiple users at a time and its purpose is to find an optimal match of wines to all the ordered dishes and user preferences. It displays the wine suggestions in the central part of the screen, while the left part is occupied by the order information. In addition to wine information, the suggestion part includes a justification generated by the system, which provides an explanation "why" the specific wine is suggested, based on the activated rules (Figure 2).

Initially, all available wines are considered, but gradually they are filtered with strict and weak rules so that they can be properly ranked. Strict rules impose exclusion of wines that do not match with the basic ingredient of an ordered dish or exceed the customer's budget. They also take into consideration other user preferences (e.g., preference for local wines). Weak rules are based on three rule categories, which represent general rules of thumb often used by experts. By using priorities among the rule categories, the system aims to satisfy as many rule categories as possible, while suggesting wines as close to the customer's budget as possible.

Le Sommelier relies on Answer Set Programming (ASP) and the tool Clingo for reasoning with the various constraints. ASP is a declarative language oriented towards difficult search problems. It is based on the stable model (answer set) semantics of logic programming. The Answer Set Programming tool Clingo comprises the Gringo grounder and the Clasp solver. The GUI is a web-based user interface that relies on Http Servlet requests on a Java API that offers the functionality to call the reasoner and get the wine recommendations.

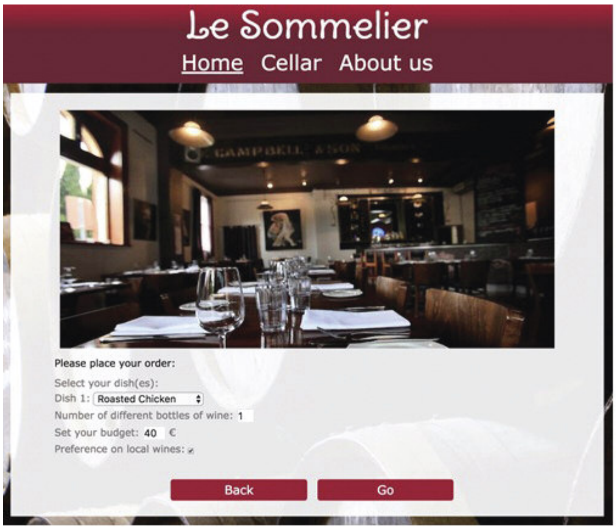



*Figure 1: A part of Le Sommelier user interface depicting the order form.*


*Figure 2: A part of Le Sommelier user interface depicting the recommended wine along with a justification.*

Our future activity includes extending the system with extra features such as the exploitation of the user's history and profile to provide even more personalised recommendations. Furthermore, thorough evaluations with real users and various restaurant menus/wine lists will be carried out, including expert walk-through evaluations of the system and adjustments based on user expert feedback.

**Links:**
www.ics.forth.gr/isl/sommelier/
www.ics.forth.gr/isl/sommelier-video

**Please contact:**
Nikoleta Tsabanaki, ICS-FORTH, Greece
tsaban@ics.forth.gr

# International Workshop on Computational Intelligence for Multimedia Understanding

by Davide Moroni (ISTI-CNR), Behçet Uğur Töreyin (İTÜ) and A. Enis Cetin (Bilkent University)

***Around twenty researchers attended the International Workshop on Computational Intelligence for Multimedia Understanding (IWCIM) organized annually by the working group Multimedia Understanding through Semantics, Computation and Learning (MUSCLE) of the European Research Consortium for Informatics and Mathematics (ERCIM), which took place as a satellite workshop to EUSIPCO-2017 held in Kos, Greece, September 2, 2017, with eleven original research papers. The IWCIM 2017 website is hosted by Istanbul Technical University and full-text papers may be accessed via MDPI Proceedings.***

Multimedia understanding is an important part of many intelligent applications in our social life, be it in our households, or in commercial, industrial, service, and scientific environments. Analyzing raw data to provide them with semantics is essential to exploit their full potential and help us managing our everyday tasks. Nowadays, raw data normally come from a host of different sensors and other sources, and are different in nature, format, reliability and information content. Multimodal and cross-modal analysis are the only ways to use them at their best. Besides data analysis, this problem is also relevant to data description intended to help storage and mining. Interoperability and exchangeability of heterogeneous and distributed data is a need for any practical application. Semantics is information at the highest level, and inferring it from raw data (that is, from information at the lowest level) entails exploiting both data and prior information to extract structure and meaning. Computation, machine learning, statistical and Bayesian methods are tools to achieve this goal at various levels.

Besides the regular track, a special theme for 2017 was selected: "Signal Processing for Surveillance and Security Applications.

Signal processing is an indispensable technology and research field for surveillance and security applications. With the advance of sensor technology and increased computational power, multimodal security and surveillance systems exploiting several modalities become more prevalent than conventional surveillance systems depending solely on single-channel visible-range video. In that respect, methods and techniques for multimodal sensor and signal analysis play instrumental role for surveillance and security applications.

Researchers had the opportunity to attend the invited talk delivered by Prof. Alptekin Temizel (METU) on "First Person Activity Recognition with Multimodal Features". Prof. Temizel presented the state-of-the-art on the subject followed by a discussion on a comparative performance assessment of several features for first person activity recognition.


*Prof. Alptekin Temizel, invited speaker from Middle East Technical University, Ankara, Turkey, presenting the state-of-the-art techniques on first-person activity recognition with multimodal features.*

**Links:**
http://wiki.ercim.eu/wg/MUSCLE/
http://iwcim.itu.edu.tr
http://www.mdpi.com/2504-3900/2/2
https://kwz.me/hbQ

**Please contact:**
Davide Moroni, ISTI-CNR, Italy
davide.moroni@isti.cnr.it

Behçet Uğur Töreyin, ITU, Turkey
toreyin@itu.edu.tr

A. Enis Cetin
Bilkent University, Turkey
cetin@bilkent.edu.tr

Call for Papers

# IWCIM 2018: 6th International Workshop on Computational Intelligence for Multimedia Understanding

at the 11th International Conference on Multimedia & Network Information Systems (MISSI 2018), Wrocław, Poland, September 12-14, 2018

The International Workshop on Computational Intelligence for Multimedia Understanding (IWCIM) is the annual workshop organized by the ERCIM Working group Multimedia Understanding through Semantics, Computation and Learning (MUSCLE). This year, IWCIM will take place as a satellite workshop to MISSI-2018, to be held in Wroclaw, Poland, on September 12-14, 2018.

Accepted papers will be published in the conference proceedings, in a bound volume by Springer-Verlag in its Advances in Intelligent Systems and Computing Series.

The scope of IWCIM 2018 includes, but is not limited to the following topics: Multisensor systems; Multimodal analysis; Crossmodal data analysis and clustering; Mixed-reality applications; Activity and object detection and recognition; Text and speech recognition; Multimedia labelling; semantic annotation, and metadata; Multimodal indexing and searching in very large data-bases; Big and linked data; Case studies and applications.

Important dates:
- Submission of papers: 10 April 2018
- Notification of acceptance: 20 May 2018
- Camera-ready papers: 31 May 2018
- Conference date: 12-14 September 2018

**More information:**
http://missi.pwr.edu.pl
http://iwcim.itu.edu.tr/

# ERCIM Membership

After having successfully grown to become one of the most recognized ICT Societies in Europe, ERCIM has opened membership to multiple member institutes per country. By joining ERCIM, your research institution or university can directly participate in ERCIM's activities and contribute to the ERCIM members' common objectives playing a leading role in Information and Communication Technology in Europe:

- Building a Europe-wide, open network of centres of excellence in ICT and Applied Mathematics;
- Excelling in research and acting as a bridge for ICT applications;
- Being internationally recognised both as a major representative organisation in its field and as a portal giving access to all relevant ICT research groups in Europe;
- Liaising with other international organisations in its field;
- Promoting cooperation in research, technology transfer, innovation and training.

## About ERCIM

ERCIM – the European Research Consortium for Informatics and Mathematics – aims to foster collaborative work within the European research community and to increase cooperation with European industry. Founded in 1989, ERCIM currently includes 15 leading research establishments from 14 European countries. ERCIM is able to undertake consultancy, development and educational projects on any subject related to its field of activity.

ERCIM members are centres of excellence across Europe. ERCIM is internationally recognized as a major representative organization in its field. ERCIM provides access to all major Information Communication Technology research groups in Europe and has established an extensive program in the fields of science, strategy, human capital and outreach. ERCIM publishes ERCIM News, a quarterly high quality magazine and delivers annually the Cor Baayen Award to outstanding young researchers in computer science or applied mathematics. ERCIM also hosts the European branch of the World Wide Web Consortium (W3C).

**"Through a long history of successful research collaborations in projects and working groups and a highly-selective mobility programme, ERCIM has managed to become the premier network of ICT research institutions in Europe. ERCIM has a consistent presence in EU funded research programmes conducting and promoting high-end research with European and global impact. It has a strong position in advising at the research policy level and contributes significantly to the shaping of EC framework programmes. ERCIM provides a unique pool of research resources within Europe fostering both the career development of young researchers and the synergies among established groups. Membership is a privilege."**

## Benefits of Membership

As members of ERCIM AISBL, institutions benefit from:

- International recognition as a leading centre for ICT R&D, as member of the ERCIM European-wide network of centres of excellence;
- More influence on European and national government R&D strategy in ICT. ERCIM members team up to speak with a common voice and produce strategic reports to shape the European research agenda;
- Privileged access to standardisation bodies, such as the W3C which is hosted by ERCIM, and to other bodies with which ERCIM has also established strategic cooperation. These include ETSI, the European Mathematical Society and Informatics Europe;
- Invitations to join projects of strategic importance;
- Establishing personal contacts with executives of leading European research institutes during the bi-annual ERCIM meetings;
- Invitations to join committees and boards developing ICT strategy nationally and internationally;
- Excellent networking possibilities with more than 10,000 research colleagues across Europe. ERCIM's mobility activities, such as the fellowship programme, leverage scientific cooperation and excellence;
- Professional development of staff including international recognition;
- Publicity through the ERCIM website and ERCIM News, the widely read quarterly magazine.

## How to Become a Member

- Prospective members must be outstanding research institutions (including universities) within their country;
- Applicants should address a request to the ERCIM Office. The application should inlcude:
    - Name and address of the institution;
    - Short description of the institution's activities;
    - Staff (full time equivalent) relevant to ERCIM's fields of activity;
    - Number of European projects in which the institution is currently involved;
    - Name of the representative and a deputy.
- Membership applications will be reviewed by an internal board and may include an on-site visit;
- The decision on admission of new members is made by the General Assembly of the Association, in accordance with the procedure defined in the Bylaws (http://kwz.me/U7), and notified in writing by the Secretary to the applicant;
- Admission becomes effective upon payment of the appropriate membership fee in each year of membership;
- Membership is renewable as long as the criteria for excellence in research and an active participation in the ERCIM community, cooperating for excellence, are met.

**Please contact the ERCIM Office:** contact@ercim.eu

Call for Participation

## IFIP NETWORKING 2018

Zurich, 14 – 16 May 2018

The IFIP Networking 2018 Conference (NETWORKING 2018), to be held in Zurich, Switzerland, from May 14 to May 16, 2018 is the 17th event of the series, sponsored by the IFIP Technical Committee on Communication Systems (TC6). Accepted papers will be published in the IFIP Digital Library and submitted to the IEEE Xplore Digital Library.

The main objective of Networking 2018 is to bring together members of the networking community, from both academia and industry, to discuss recent advances in the broad and quickly-evolving fields of computer and communication networks, to highlight key issues, identify trends, and develop a vision for future Internet technology, operation, and use.

The conference features three keynotes:
- “Blockchain for Cyber Physical Systems” by Salil Kanhere, University of New South Wales, Sydney, Australia;
- “Flexibility Matters: On the Design and Evaluation of Softwarized Networks” by Wolfgang Kellerer, Technical University of Munich (TUM), Germany;
- “Artificial Intelligence in Network Operations and Management” by Jürgen Quitteck, NEC Laboratories Europe, Germany;

**More information:**
http://networking.ifip.org/2018/

## In Memoriam Gerrit Blaauw

In March 2018 the Dutch computer pioneer Gerrit Blaauw passed away, 93 years old. He did pioneering research at the Mathematisch Centrum (MC) in Amsterdam, currently called CWI. Blaauw worked from 1952 until 1955 at the MC, where he contributed to the design of the ARRA II, one of the first computers built in the Netherlands, under supervision of Aad van Wijngaarden. Blaauw was also involved in the design of the FERTA computer for Fokker Aircraft. Later in his career he co-designed the architecture of the IBM System/360 mainframe computer in the USA.


*Gerrit Blaauw in 2015. Picture (©): E. Verhagen.*

He became famous for his methodological manner of building computer machines, making a difference between architecture, implementation and realization of a machine. Later in his career Blaauw became a professor at the University of Twente and a member of the Royal Netherlands Academy of Arts and Sciences (KNAW).

## Evaluation of Virtual Research Environment Architectures

Edinburgh, Scotland, 15 June 2018

The EU VRE4EIC project will hold a workshop on Evaluation of Virtual Research Environment Architectures co-located with the 10th International Workshop on Science Gateways (IWSG18) on 15 June in Edinburgh, Scotland.

Researchers can access and use more and more research data in the digital age. They can use this data to obtain new insights, especially by combining datasets with other data. Discoverability of data is very important as a precondition for combining datasets and data. Data could be described by metadata in many ways and searching those metadata improves discoverability. Various projects are already producing e-Research Infrastructures to give researchers access to publicly funded research and open research data, and are developing towards Virtual Research Environments (VREs). VREs provide access to data, tools, and resources from different research infrastructures, and facilitate co-operation or collaboration between researchers at the same or different institutions, at the intra- and inter-institutional levels, and preserve data and other outputs.

The objective of this half-day workshop, which will take place in the afternoon following the close of the IWSG 2018 international workshop, is to review the architectural requirements for virtual research environments and their relationship to science gateways, virtual laboratories and collaboratories. The workshop targets people interested in using, developing and extending virtual research environments. The expected outcome will be a written record of group discussions, a SWOT analysis and data collected from a questionnaire during the workshop.

Participation is free but registration is mandatory

VRE4EIC is a Horizon 2020 research project coordinated by ERCIM.

https://kwz.me/hbZ
https://kwz.me/hbX

ERCIM European Research Consortium for Informatics and Mathematics

**ERCIM – the European Research Consortium for Informatics and Mathematics is an organisation dedicated to the advancement of European research and development in information technology and applied mathematics. Its member institutions aim to foster collaborative work within the European research community and to increase co-operation with European industry.**

**ERCIM is the European Host of the World Wide Web Consortium.**


Consiglio Nazionale delle Ricerche
Area della Ricerca CNR di Pisa
Via G. Moruzzi 1, 56124 Pisa, Italy
http://www.iit.cnr.it/

Norwegian University of Science and Technology
Faculty of Information Technology, Mathematics and Electrical Engineering, N 7491 Trondheim, Norway
http://www.ntnu.no/

Centrum Wiskunde & Informatica
Science Park 123,
NL-1098 XG Amsterdam, The Netherlands
http://www.cwi.nl/

RISE SICS
Box 1263,
SE-164 29 Kista, Sweden
http://www.sics.se/

Fonds National de la Recherche
6, rue Antoine de Saint-Exupéry, B.P. 1777
L-1017 Luxembourg-Kirchberg
http://www.fnr.lu/

SBA Research gGmbH
Favoritenstraße 16, 1040 Wien
http://www.sba-research.org/

Foundation for Research and Technology – Hellas
Institute of Computer Science
P.O. Box 1385, GR-71110 Heraklion, Crete, Greece
http://www.ics.forth.gr/

Magyar Tudományos Akadémia
Számítástechnikai és Automatizálási Kutató Intézet
P.O. Box 63, H-1518 Budapest, Hungary
http://www.sztaki.hu/

Fraunhofer ICT Group
Anna-Louisa-Karsch-Str. 2
10178 Berlin, Germany
http://www.iuk.fraunhofer.de/

University of Cyprus
P.O. Box 20537
1678 Nicosia, Cyprus
http://www.cs.ucy.ac.cy/

INESC
c/o INESC Porto, Campus da FEUP,
Rua Dr. Roberto Frias, nº 378,
4200-465 Porto, Portugal

Universty of Warsaw
Faculty of Mathematics, Informatics and Mechanics
Banacha 2, 02-097 Warsaw, Poland
http://www.mimuw.edu.pl/

Institut National de Recherche en Informatique
et en Automatique
B.P. 105, F-78153 Le Chesnay, France
http://www.inria.fr/

I.S.I. – Industrial Systems Institute
Patras Science Park building
Platani, Patras, Greece, GR-26504
http://www.isi.gr/

VTT Technical Research Centre of Finland Ltd
PO Box 1000
FIN-02044 VTT, Finland
http://www.vttresearch.com