

FP7 Project ASAP
Adaptable Scalable Analytics Platform



D10.1 Dissemination Strategy Document

This report describes the strategy, channels, and methods for dissemination of project results and maximizing its impact

WP 10 - Exploitation and Dissemination

Nature: Report

Dissemination: Public

Version History

Version	Date	Author	Comments
0.1	16 Jul 2014	A. Scharl	Initial Version
0.2	10 Aug 2014	A. Scharl	First Major Revision
0.3	15 Aug 2014	N.N.	Various Contributions by Consortium Partners
1.0	31 Aug 2014	A. Scharl	Final Revision and Layout

Acknowledgement

This project has received funding from the European Union's 7th Framework Programme for research, technological development and demonstration under grant agreement number 619706.

1. Executive Summary

This report outlines the dissemination and impact activities to be performed in the ASAP project, what future activities are planned, and how the project results will be made available to the European industry. This document will be updated during the project.

All partners will continue to play active roles in dissemination, placing special emphasis upon scientific channels of publications including journal and conference papers, but also actively engaging in workshops, summer schools and other courses, invited talks, and various online dissemination activities. External networks and advisors will form a key part of the dissemination strategy towards non-scientific users, as they are providing outreach towards citizens and organizations in diverse sectors - e.g., press releases, hosting of events, referencing the project in online publications, and presentations aimed at non-scientific users.

The main public interface of ASAP is the project web site (www.asap-fp7.eu), which contains information on the project objectives, partners, R&D activities, and summaries of results as they become available. Furthermore, the scientific results of the project are being submitted to major international conferences and journals. The project partners will take part in relevant national and European concertation events.

In order to promote maximum use and dissemination, major technology components from ASAP are being made available as open source and are thus easily exploitable both commercially and for research purposes. In order to ensure support beyond the project's lifetime, the source code of these components will be made available on well-known public platforms such as GitHub.com or SourceForge.org.

2. Knowledge Creation and Dissemination

For maximum impact, ASAP has adopted a **multi-channel dissemination approach**, which will be outlined in the following - this includes setting up online spaces for supporting communities of interest, participating in scientific conferences and exhibitions, as well as publishing at high impact journals and conferences.

These dissemination activities are aimed at creating a widespread **awareness** and **understanding** of the benefits of ASAP, interacting with stakeholders and

demonstrating the work to businesses and **potential customers**, sharing **technological achievements** with researchers and practitioners, establishing and maintaining a favourable reputation of the project, and fostering further collaborations and exploitation activities.

The consortium recognises that a favourable reputation created through dissemination can lead to **competitive advantage** and generate market demand for the products or services being created and exploited. This will support the participating companies who are actively engaged in implementing and delivering commercial software offerings based on ASAP.

The following sections will first describe **general dissemination activities** (e.g., designing a logo, setting up the project Website), and then describe the dissemination plans of individual **project partners**, who will contribute to project dissemination in ways that fit their research domain and technology development.

2.1 General Dissemination Activities

2.1.1 ASAP Logo

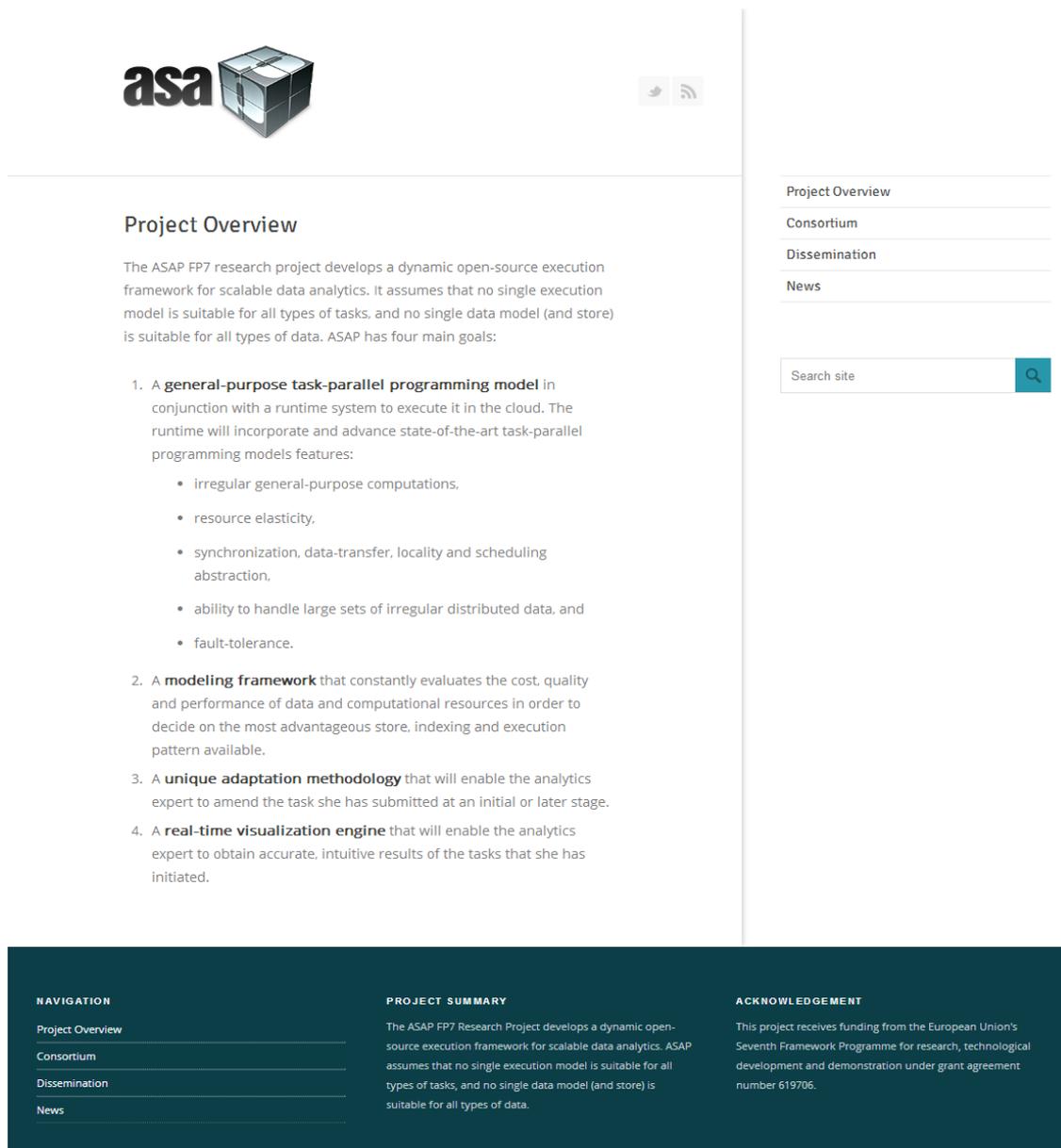
The following logo below has been designed for ASAP and is used in all dissemination materials; it is provided as a shaded bitmap version (left) as well as a black-and-white vector version (right) for high-quality print reproduction.



2.1.2 Project Web Site

The project has set up and maintains a public **Website** at www.asap-fp7.eu, allowing distributed editing among consortium members. Following an initial static HTML version, a dynamic site based on the WordPress content management system with a professional theme has been activated in August 2014. WordPress supports a more interactive development process and the independent editing of document by consortium partners.

This Website functions as a project dissemination and documentation tool and single access point that references other ASAP applications and services. In addition to documenting the scientific output of the project, examples and descriptions will outline how the information services can be used by third parties, independently or in conjunction with other Web applications.



The screenshot shows the ASAP FP7 Project website. At the top left is the ASAP logo, which consists of the letters 'asa' in a bold, lowercase font next to a 3D cube icon. To the right of the logo are social media icons for Twitter and RSS. Below the logo is a horizontal line. The main content area is titled 'Project Overview'. It contains a paragraph describing the project's goal: 'The ASAP FP7 research project develops a dynamic open-source execution framework for scalable data analytics. It assumes that no single execution model is suitable for all types of tasks, and no single data model (and store) is suitable for all types of data. ASAP has four main goals:'. This is followed by a numbered list of four goals, each with a brief description and a bulleted list of features or characteristics. On the right side of the page, there is a vertical navigation menu with links for 'Project Overview', 'Consortium', 'Dissemination', and 'News'. Below the menu is a search bar with the text 'Search site' and a magnifying glass icon. At the bottom of the page, there is a dark teal footer with three columns: 'NAVIGATION' with links to 'Project Overview', 'Consortium', 'Dissemination', and 'News'; 'PROJECT SUMMARY' with a short paragraph about the project; and 'ACKNOWLEDGEMENT' with a paragraph about funding from the European Union's Seventh Framework Programme.

www.asap-fp7.eu will be updated continuously during the project, as progress is made and new outputs become available. The content of the Website will include:

- Project overview and objectives;
- Project partner information;
- Downloadable project outputs (publications, deliverables, etc.);
- News and links to dissemination events with deadlines, related projects and research activities, etc.
- Links to related projects and research activities;
- Demos that have been added under scientific results section.

2.1.3. Social Media Channels

Posting project-related information to social media platforms will extend outreach and increase the exposure of project results. The ASAP project has now established a Twitter account (@ASAP_EU), where all project partners will disseminate their publications, events organised and participated in, etc. Project members also disseminate ASAP-specific results in additional social media channels such as LinkedIn and Facebook. The tag #asap-fp7 will be used to tag posts and material published via these channels.



2.1.4 Other Promotional Activities

FORTH organized the printing of T-Shirts with the ASAP logo as an additional promotional item to increase project visibility, for example when attending conferences or showcasing the project at various events.



2.2. Scientific Publications, Demonstrations and Related Public Outreach Activities

To be documented on the ASAP Website, we will publish a growing body of **scientific papers** and **presentations** at **technical and scientific conferences**. The specific outlets will be selected based on their relation to the topics addressed by ASAP, and on their impact. The publications are aimed at sharing the results of ASAP with the European and international scientific community, to encourage their incorporation into the work of other scientists and technologists. The project will also provide online software demonstrations, available from the ASAP Website.

With respect to IPR and promoting take-up of project results, we endeavour to make all **scientific results** of this project available to the research community. In addition, associated tools will be released as open-source software to facilitate take-up.

In order to ensure the widest possible audience, we will use **press releases** and **training activities** aimed at the academic and industrial sectors (e.g., delivery of tutorials), industry-oriented dedicated workshops, etc. In May 2014, WLT has sent out a newsletter (n=642)¹ and published a corresponding document on its Website.² The ASAP project has also been mentioned in the June 2014 issue of “Swiss CIO”, a publication by the Swiss Computerworld (see Section 2.2.1).

2.2.1. Submitted and Planned Papers

The ASAP consortium aims to actively produce and disseminate research results, targeting top-tier **international conferences** and **journals** such as VLDB, SIGMOD, ICDE, CIKM, BigData, Cloud, IEEE TKDE and ACM DKE.

Accepted or Submitted Refereed Publications

1. “I/O Performance Modeling for Big Data Applications over Cloud Infrastructures”, Ioannis Mytilinis, Dimitrios Tsoumakos, Verena Kantere, Anastassios Nanos and Nectarios Koziris. IEEE International Conference on Big Data 2014 (IEEE BigData 2014).
2. “Energy-Efficient In-Memory Data Stores on Hybrid Memory Hierarchies”, Ahmad Hassan, Hans Vandierendonck and Dimitrios S. Nikolopoulos. Submitted to the 2015 Conference on Innovative Data Systems Research (CIDR 2015).

¹ <http://eepurl.com/Tsvln>

² <http://www.weblyzard.com/scalability>

3. “Mapping Construction Compliant with Schema Semantics”, Verena Kantere. International Conference on Database and Expert Systems Applications (DEXA), 2014.
4. “A Holistic Framework for Big Scientific Data Management”, Verena Kantere . IEEE International Conference on Big Data 2014 (IEEE BigData 2014).
5. “Scalable Knowledge Extraction and Visualization for Big Data Applications”, Arno Scharl, Albert Weichselbraun et al. Submitted to IEEE Computer, Special Issue on Big Data Management and Applications.

Trade Publications

1. Scharl, A. and Weichselbraun, A. (2014). “Die Meinungs-Erforscher”, Swiss CIO Magazine, Computerworld, 20 June 2014, 32-35.

Future Publication Plans

ICCS intends to disseminate project results via the submission of research papers and the publication of articles in specialized scientific conferences and journals. The related research areas include Big Data, Distributed Computing, Cloud Computing, Data Management, Performance Modelling, etc. ICCS has already submitted a research paper entitled “I/O Performance Modelling for Big Data Applications over Cloud Infrastructures” to IEEE BigData 2014. In the course of the project the number of submissions will increase, as soon as the first tangible results are available.

IMR mostly participates in workshops and seminars where the R&D team is invited to present technical achievements, performance evaluations, use cases and industrial assessments of open source tools for Big Data management (see, e.g., our presentation at Big’14, a workshop collocated with WWW’14). ASAP will typically supply material and results that could be presented during such events.

UniGe has started disseminating the initial research results to international scientific conferences. Two papers have been accepted for publication: (i) “A Holistic Framework for Big Scientific Data Management” has been accepted to the International Congress on Big Data (IEEE BigData), 2014; (ii) “Mapping Construction Compliant with Schema Semantics” has been accepted to the International Conference on Database and Expert Systems Applications (DEXA), 2014.

webLyzard. Preliminary results of scalability improvements in the knowledge extraction processing pipeline of WLT have been submitted to IEEE Computer. WLT plans to submit further results in the fourth quarter of 2014. In addition to the examples listed at the beginning of Section 2.2.1, we will also target other ACM and IEEE outlets, as well as dedicated journal calls with a focus on scalable analytics, information visualization, dashboard design, and human-computer interaction.

FORTH will disseminate the scientific results of the project via research paper conference publications, journal publications, workshop presentations, and poster presentations in related scientific venues. Moreover, FORTH will organize and participate in sessions in related events such as network-of-excellence meetings. Related areas include High-Performance Computing, Big Data, Distributed Systems, Parallel Programming and Programming Languages. FORTH is currently organizing a thematic session in the HiPEAC Computing Systems Week event in October 2014.

2.2.2. Events Attended and Presentations

1. Katsogridakis, Pratikakis: Poster Presentation on Data Analytics Runtimes at ACACES 2014 (2014-06-20).
2. Vandierendonck, Hans: Lecture on the Swan task dataflow system at Summer School on Advances in Programming Languages (AiPL 2014), Heriot-Watt Univ, UK (2014-08-19).

2.3. Collaboration

ICCS will exploit results from the recently concluded FP7 Project **ARCOMEM** (www.arcomem.eu). Within ARCOMEM, ICCS team members have developed the ARCOMEM database, a component responsible for storing huge amounts of crawled raw data as well as metadata of annotated web archives. The invaluable experience gained in various data stores such as HDFS, HBase, H3RDF (a distributed rdf store developed for the project) will be exploited in accurately modelling such engines.

The **CELAR** project (www.celarcloud.eu) works on elasticity aspects of applications deployed over cloud infrastructures. **ICCS** will exploit its joint involvement in the two projects to evolve research and tools relevant to profiling, modelling and elasticity decisions customized to Big Data and analytics engines. Moreover, **ICCS** plans to utilize ASAP ideas and components in the ongoing **~okeanos** project (okeanos.grnet.gr) of GRNET (Greek Research and Technology Network). Its goal is to develop a public cloud infrastructure, code-named **~okeanos**, offering IaaS to the whole Greek academic community (similar to Amazon WS). Prof. Nectarios Koziris, senior researcher in ICCS, is currently leading the GRNET **~okeanos** team.

IMR has been closely working with TU Berlin, during several years, on the applicability of Stratosphere, a distributed data processing system, to web data management. Since April 14, Stratosphere is an official Apache project, named Apache Flink. We are keen to develop our experience Flink in the context of ASAP.

The **CACTOS Project** (www.cactosfp7.eu) develops the data analytics component of CactoScale, a set of tools and methods to acquire and analyse application behaviour and infrastructure performance data. CactoScale conducts detailed platform and service workload characterisation by mining console and system performance logs. The mined information informs optimisation strategies for VM placement and migration within and across datacentres. The analytic tool of CactoScale can directly feed into ASAP to support the mapping of application components to data centres and services in a way that exploits their topological properties and is aware of application execution patterns and resource demand.

webLyzard technology. ASAP benefits from the results of related national research projects such as **DIVINE** (www.weblyzard.com/divine). We will also pursue active research collaborations, identify synergies to maximize impact, and exchange results with other ongoing projects such as **uComp** (www.ucomp.eu), which develops a human computation framework to blend knowledge extraction and crowdsourcing approaches, as well as **DecarboNet** (www.decarbonet.eu), which builds a collective awareness platform for carbon footprint reductions.

2.4. Software and Data Resources

Software prototypes will be made publicly available for download from the project Website or public repositories such as GitHub.com and SourceForge.org. The list of currently available components includes:

- The **Swan Task Dataflow Scheduler** and its (preliminary) distributed variant www.github.com/hvdieren/swan www.github.com/Richard-Neill/distributedswan
- **easy Web Retrieval Toolkit (eWRT)**
www.github.com/weblyzard/ewrt
Modular open-source Python API to retrieve social data from Web sources, including various helper classes for effective caching and data management.
- **Media Watch on Climate Change (MWCC)**
www.ecoresearch.net/climate
Publicly accessible showcase of the webLyzard dashboard with data export capabilities, focusing on climate science and related environmental issues.
- IMR's policy does not include direct participation to open source projects, although we indirectly contribute to the development and assessment of open source software (e.g., **Apache Flink**, see above).

2.5. Industry-Oriented Materials

As a result of ongoing industry-oriented activities, exploitable results of ASAP will be made public through the ASAP Website as software packages, presentations and dedicated web documents. The industrial partners within the ASAP consortium will seek synergies and emphasize commonalities in industry-oriented presentations. We will pursue a more differentiated approach vis-à-vis specific sectors, depending on which organization is most skilled to address this sector.

The presentations are aimed at helping convince early adopters about the benefits and opportunities that ASAP can offer. Therefore, the presentations will include a comparison of existing solutions and collect specific business needs from industry.

3. Future Exploitation Potential

Due to the generic nature of the project results, especially the ASAP unified programming model and distributed computing engine, exploitation activities will go beyond one specific industry, and beyond the defined uses cases. Specifically tailored exploitation actions are being developed and they will be targeted at the following three key groups:

1. **Companies** already **collaborating with ASAP commercial partners**,
2. **Industrial working groups** and **standardization bodies**,
3. **Companies** in other **domains**, not covered in the ASAP use cases.

Internet Memory Research develops a Web data analytics platform called Mignify. Our goal in ASAP is to extend the streams analytic and notification service of Mignify to Web-scale. In particular, IMR expects to enrich the service scope with a cloud-based approach. Currently, the query-and-annotation pipeline is hosted in the Mignify data centre. Thanks to ASAP and the multi-centre approach it promotes, the Q&A service could be hosted on a dedicated server, possibly for a limited period of time, while provision, storage and indexing services would still be located in IMR's infrastructure. This presents a high additional value for Mignify clients.

WIND, according to the dissemination plan defined with the other partners, will play a fundamental role in communicating the project results, first of all using the internal channels in the company, then using the channels in the Vimpelcom group (typically the company intranet web sites) and in particular cases, via conferences and journals, the dissemination of the ASAP system's value to WIND's stakeholders will raise public awareness for the project and its objectives.

WIND is planning activities to promote the exploitation of the ASAP results. In particular the employees in the technology and marketing area will be involved in this process to design, develop new and improve existing products and services taking into account the ASAP benefits. The dissemination activities in WIND will allow the company to evaluate how to use the new approach to: (i) maximize visibility for brands and WIND products; (ii) enhance relationships through bidirectional interaction channels; (iii) understand market insights to tailor targeted offers measuring efficiency and effectiveness according to the customer requests; (iv) reduce the social network fragmentation and using a new approach to learn from the content of fan pages; and (v) tailor targeted offers understanding better the efficiency and effectiveness factors. Furthermore, the WIND dissemination activities will deal with the impact of the ASAP project in the area of the *Privacy-Aware Mobility Mining* to improve the portfolio of services, taking into account all the acquired knowledge in a big data context.

webLyzard technology. In terms of visibility and engagement in the relevant research communities, and to attract overseas clients (a crucial factor for an SME focusing on large-scale applications of semantic technologies), the communication activities of webLyzard target government organizations and research centres in Europe and the United States, as well as large business-to-consumer brands.

Word-of-mouth and leveraging close ties with large federal agencies such as the National Oceanic and Atmospheric Administration (NOAA) will also be beneficial for disseminating the research results of ASAP to a wide international audience. The improved scalability through the project's distributed computing engine will help to increase the knowledge base of webLyzard, attracting new clients and representing an important competitive advantage - particularly in conjunction with the new visualization components developed in WP6.