Information Society Technologies

IST FP6-027413

12-monthly Activity Report Period 02

Public Executive Summary

MediaCampaign Identifier: MC-T24-JRS-02-PublishableSummary_P02.doc

Author(s) and company: Herwig Rehatschek (JRS), Helen Hasenauer (JRS), Herwig Zeiner (JRS), Mark Guinney (NMR-UK), Julien Nioche (USFD), Borislav Popov (ONTO), Walter Plaschzug (HSA), Marco Masetti (SOF), Roeland Ordelman (UT), David van Leeuwen (TNO)

Work package/task: WP 1 / T1.1

Document status: Final

Confidentiality: public

Keywords: project highlights, project progress

Abstract: This document contains the public executive summary of the periodic activity report for the period 04/2007 (PM 13) – 03/2008 (PM 24) of MediaCampaign project.

DOCUMENT HISTORY

Version	*Date*	*Reason of change*
1	2008-04-28	document created
2	2008-05-05	Input from partners incorporated

Table of Contents

1 Publishable Executive Summary

1.1 Major Achievements within the second project year

In April last year the first MediaCampaign prototype (Milestone II) was set-up. For this prototype all analysis modules were available in a first version and with all necessary communication interfaces implemented. This prototype was fully integrated and supported the full workflow for press and TV. All modules with basic functionality were available. The software was successfully presented during the 1st EC review.

This year a first version - in preparation of the second project review- a first version of the MediaCampaign Integrated Prototype (Milestone III / due to in June 2008) has been installed at NMR-UK offices and presented to NMR-UK operators. This system was also demonstrated during the 2nd EC review. Please refer to the figures at the end of this section for a principle workflow of the system and its associated applications (Figure 1) and for the exact technical configuration of the system currently installed at NMR (Figure 2).

Another main activity within the last reporting period was the first evaluation of the MediaCampaign system. All results have been manifested within D2.3 which was submitted in-time to the EC. Generally positive results have come back from all of the component tests done. Most of the system integration tests performed have been successful within the scope of the requirements at this stage.

A central focus in the last year has been the incremental update of MEdia Presence and Campaign Ontology – MEPCO. It has been designed according to the analysis of the domain and existing terminology and models. MEPCO has been developed as an extension of the basic upper level ontology PROTON[1] reducing the overall effort needed for a domain ontology by reusing the existing entity classes and properties defined there. MEPCO V2.0 has been released as a public Deliverable (D3.4), which can be downloaded from the public website. Major improvements included to establish a mapping to output metadata of the various analysis modules in the first prototype and to allow for an easier discovery of related creatives and thus the discovery of ongoing advertising campaigns in media. Further developments have been focused to the enrichment of the modeling schemata to capture more of the metadata available in MMS, the transmigration of characteristics from Spots to Creatives and eventually to Campaigns, through direct citing, multi-value attributes and statistical summarization of discrete components from the lower levels of abstraction (i.e. from actual media presence to campaign abstractions) and on ensuring compatibility with BigOWLIM (OMS) for enterprise scale representation of large amounts of instance data. Last but not least the SPARQL compatibility of the model was ensured.

During the last year a decision was taken by the consortium to significantly expand the functionality of knowledge fusion and campaign detection to take into account the features generated by the analysis modules at the level of spots. The analysis features that are considered relevant to this task include now a word vector built from all the textual information extracted from spots associated with a creative, the set of logos identified by the visual analysis process, the set of jingles identified by the audio analysis process and the set of entities identified by the text analysis process. Consequently, the Knowledge Fusion sub-system was redesigned to collect all the relevant analysis results from all the spots associated with a given creative and use them to build a creative descriptor that is then forwarded to the Campaign Discovery (CD) module.

As an important step into direction of exploitation preparation the Press workflow subsystem of MediaCampaign consisting of AdClipper, AdComparer and parts of the MMS were installed at NMR. This software is now ready for a first user evaluation which started in March 2008. Next to this important contacts to a Swedish and a Norwegian company could be established which are highly interested in solutions for media monitoring. There were already first meetings with these companies where a demo system was demonstrated. The companies are also interested in other fields of media-analysis (brand detection) and therefore a potential customer for several MC components.

[1] http://proton.semanticweb.org

In connection with dissemination the consortium could successfully publish five scientific publications, two poster sessions and gave presentations at two relevant events. The dissemination report of the activities in year 2 (D7.2.4) has been compiled and submitted to the EC. Further more updates of the public website have been performed and HSA added a description of PrimaRS on their own company public website.

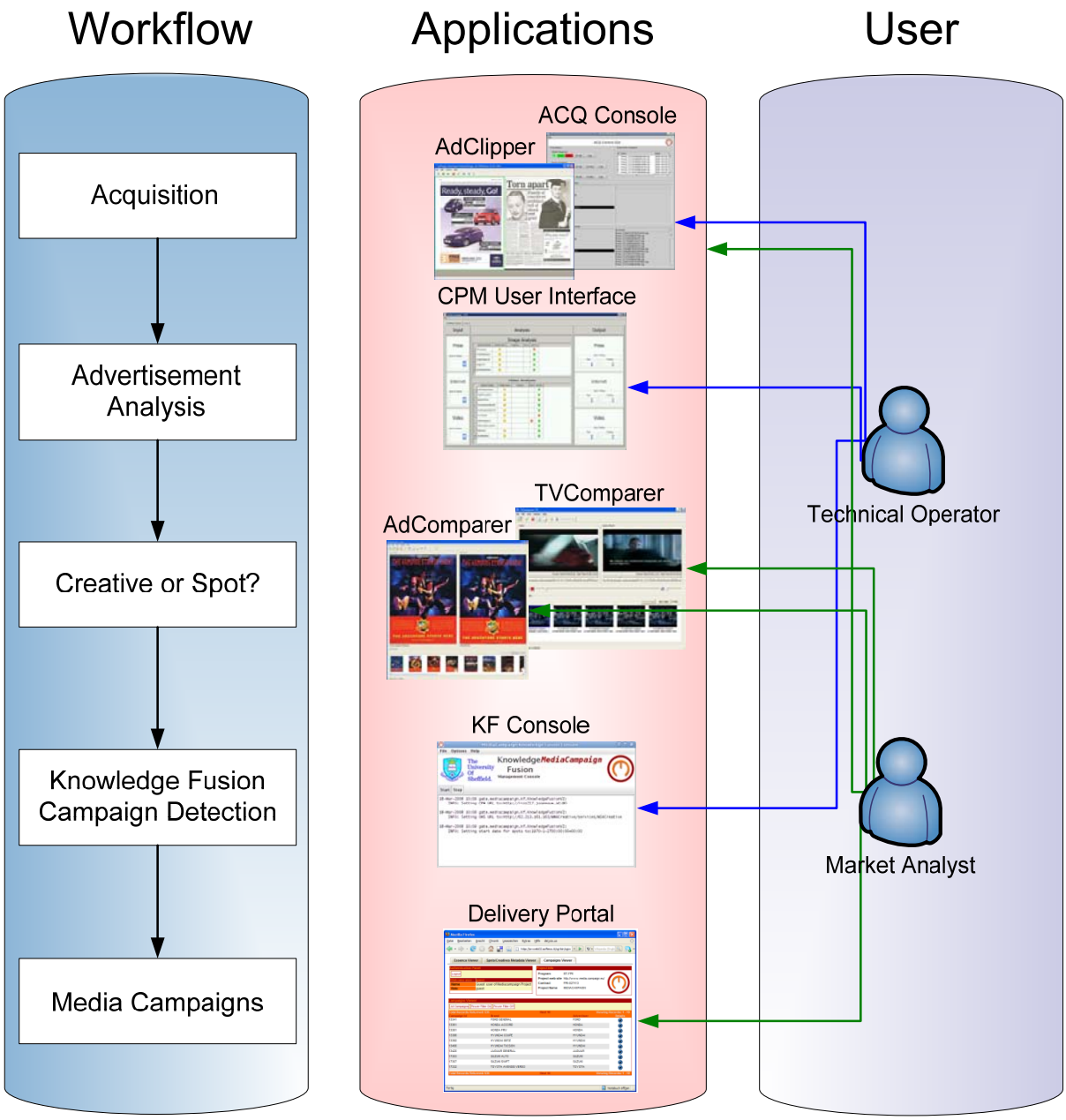

Figure 1: Operator Workflow

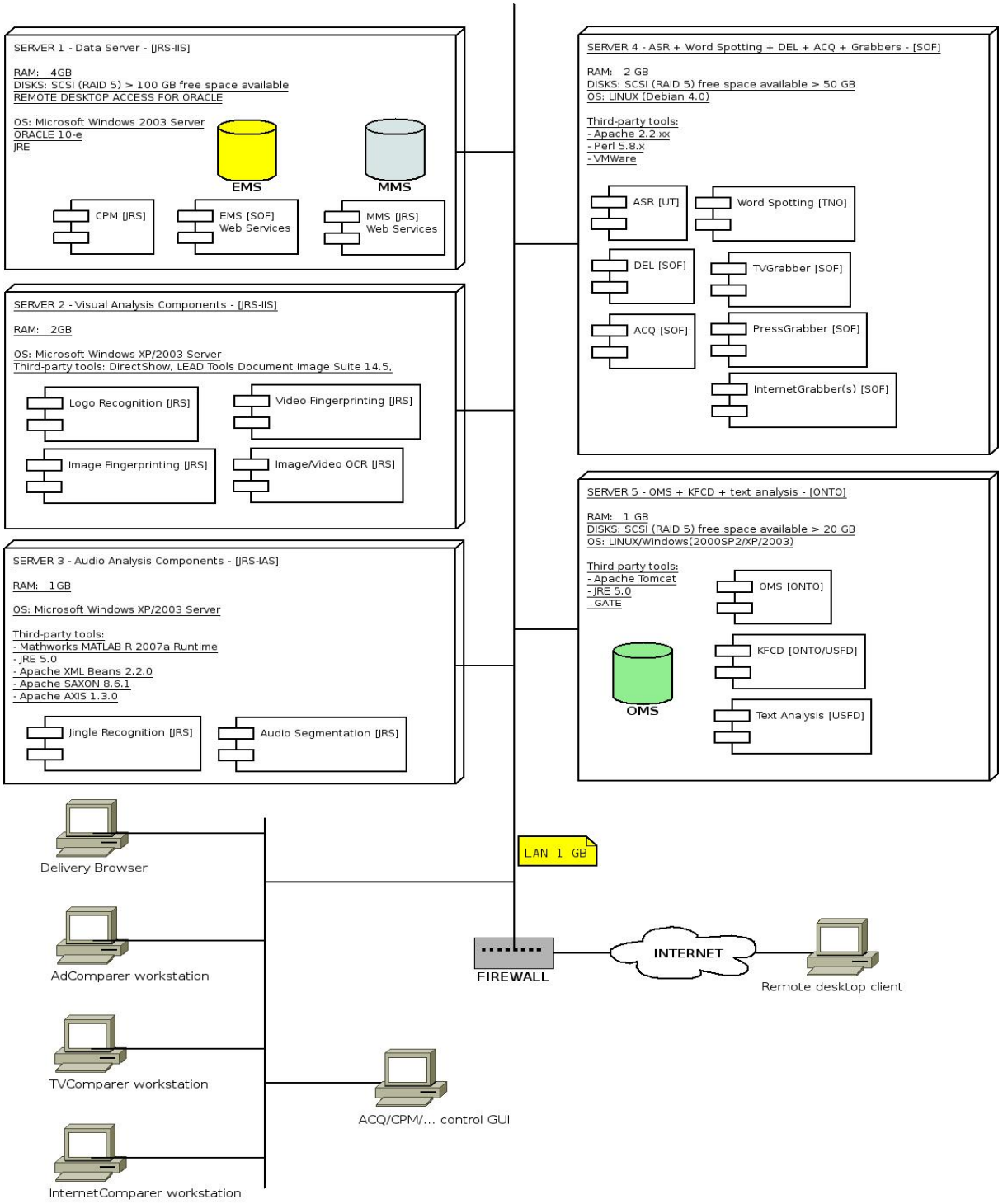

Figure 2: current configuration oft he MediaCampaign integrated prototype installed at NMR

1.2 Next steps

The project is going towards the fulfilment of its next milestone III: "Media Campaign Integrated Prototype". A first set-up was installed in January 2008 at NMR (to be demonstrated during the EC review), however, further developments have still to be conducted. Especially with regards to the Internet workflow some technical open issues have still to be implemented.

Developments on all analysis modules, the knowledge fusion and campaign detection, the acquisition and the delivery system will be continued. User feedback resulting from the evaluation at NMR-UK will be incorporated.

Towards exploitation the consortium will start – based on the dissemination and use plan - with the final exploitation plan. With regards to dissemination the consortium will intensify its scientific publications and also prepare the project show case.

2 Glossary

Terms used within MediaCampaign project sorted alphabetically.

API	Application Programming Interface
Campaign	A campaign represents a number of creatives semantically belonging together. A campaign has a certain duration and can be cross country and cross media. See also "spot" and "creative"
Creative	A Creative represents all the occurrences of a similar spot. A Creative is not cross media (e. g. TV and press) and not cross language. See also "spot"
CPM	Collection processing manager. Central component which controls the analysis workflow via web services.
DB	Database
DBMS	Database Management System
EMS	Essence management store – holds all essence of MC, i.e. TV spots, press images and meta essence such as extracted audio files.
JAPE RULES	Extraction patterns used to identify new entities
HTML	Hypertext Markup Language
HTTP	Hypertext Transfer Protocol
IDE	Integrated Development Environment
IE	Information Engineering
iFS	Internet File System
ISO/ANSI	International Organization for Standardization / American National Standards Institute
META-ESSENCE	Under this term we understand all essence and additional data which is generated by subsystems (e.g. lo-res videos, keyframes, OCR text, ...).
MMS	Metadata Management Store – holds all data extracted by the analysis systems, i.e. low-level and mid-level features such as tagging of spots in connection with logos.
OMG	Object Management Group
OMS	Ontology management store – holds all semantic knowledge of MC, i.e. the MEPCO ontology and the campaign knowledge.
OS	Operating System
PL/SQL	Procedural Language/SQL
Re-detection	The same creative is re-detected in a single media. E.g. the ad of Renault Megane appeared on 5 May 2005 on "TV ORF1" 5.00 pm and on 7 June 2006 on "TV ORF1" at 4.00 pm.
Scene	A consecutive series of shots connected through transitions (hard cut, fade, panning, ...) that constitutes a logical unit of action in a video. It is defined by "bridging features" such as same visual/audio content, music/speech/noise segments or text overlays.
SGML	Standard Generalized Markup Language
Shot	A consecutive series of pictures representing coherent visual content, e.g. when having an interview with two persons, and the camera is changing

between the two faces (depending on who is talking), each face would be a shot.

SOAP Simple Object Access Protocol

SQL Structured Query Language

Spot An occurrence of a Creative, e.g. TV spot seen on a given channel at a specific time. The incoming material obtained from the Media Acquisition step is initially available as a spot and later attached to a new or existing Creative. See also "creative".

S&R Search and Retrieval

Tracking Tracking of creatives by means of history over media and countries. When a creative is re-detected (see also "re-detection"), it is inserted into a data structure which represents the history of this creative. This enables users to query the database for appearances of this creative in the past (e.g. it had appeared on 1 April 2006 in press "The Times" in US p.5 and on 3 April 2006 on "TV Sky 1" at 19.00 in UK.

WP Work Package

UML Unified Modelling Language

URI Uniform Resource Identifier

XML Extensible Markup Language

XPath XML Path Language

XSLT XSL Transformation

Partner Acronyms:

HSA HS-Art Digital Service GmbH

JRS JOANNEUM RESEARCH Forschungsgesellschaft mbH

NMR-UK Nielsen Media Research UK

ONTO Sirma AI EAD

SOF Softeco Sismat SpA

TNO Netherlands Organisation for Applied Scientific Research

USFD University of Sheffield

UT University of Twente

