



Project acronym	SIMBAD
Project full title	Beyond Features: Similarity-Based Pattern Analysis and Recognition Dipartimento di Informatica Università Ca' Foscari di Venezia Via Torino 155 30172 Venezia Mestre Italy <a href="http://www.dsi.unive.it/~pelillo">http://www.dsi.unive.it/~pelillo</a>
Deliverable Responsible	
Project web site	<a href="http://simbad-fp7.eu">http://simbad-fp7.eu</a>
EC project officer	Teresa De Martino
Document title	First Dissemination Plan
Deliverable n.	D8.2
Document type	Report
Dissemination level	Public
Contractual date of delivery	M 7
Project reference number	213250
Status & version	Revised version
Work package	WP8
Deliverable responsible (SHORT NAME)	UNIVE
Contributing Partners (SHORT NAME)	UNİYORK, IST, UNIVR, TUD, ETH Zurich
Author(s)	Marcello Pelillo
Additional contributor(s)	Veronica Giove

## Introduction

This deliverable aims to define the SIMBAD strategy concerning the dissemination activities, as well as to demonstrate how the consortium is ensuring the highest diffusion of the research results, both inside and outside the consortium. This document will cover the activities up to the 22<sup>nd</sup> month of the project (February 2010). Deliverable D8.5 (“Second dissemination plan”) will describe the dissemination activities from the 22<sup>nd</sup> month until the end of the project.

The activities carried out within the first two years of SIMBAD, being theoretical and/or algorithmic in nature, will appeal mainly to the machine learning, pattern recognition and computer vision communities. According to our workplan, we expect to obtain results that will appeal to people outside these audiences (e.g., medical community and even general public) in the final year of the project, when the first results concerning our large-scale biomedical applications will be available.

## Internal dissemination

In order to provide a timely access of information within the consortium, and increase interactions among the SIMBAD partners, we have established the *SIMBAD Technical Report Series*: all partners are asked to publish their latest results, as soon as they are produced, in the form of a Technical Report. The procedure to implement the TR Series will be as follows: as soon as a report is ready to be published within the Series, the TA (Technical Administrator) will provide, upon request, a proper internal reference number that will have to be reported on the front page of the report (the template of the front page will be downloadable from the project website).

All partners will be periodically informed of the latest publication in the series by e-mail.

## Publications

The dissemination of the project’s results during this period will take place mainly through publications in the top-level conferences and journals in the fields of machine learning, pattern recognition and computer vision.

In order to acknowledge financial contribution from the EU Commission, all publications will include the following statement:

*“We acknowledge financial support from the FET programme within the EU FP7, under the SIMBAD project (contract 213250)”.*

In 2009, we will present our results to the following meetings:

- EMMCVPR'09: *The 7th International Conference on Energy Minimization Methods in Computer Vision and Pattern Recognition*, Bonn, Germany, August 2009.  
(Partners involved: UNIVE, TUD, UNIVR)
- ICCV'09: *The 12th IEEE International Conference on Computer Vision*, Kyoto, Japan, October 2009.  
(Partners involved: UNIVE, UNIVR, UNİYORK)
- ICIAP'09: *The 15th International Conference on Image Analysis and Processing*, Vietri sul Mare, Italy, September 2009.  
(Partners involved: UNIVR, UNİYORK)
- CAIP'09: *The 13<sup>th</sup> International Conference on Computer Analysis of Images and Patterns*, Munich, Germany, September 2009.  
(Partners involved: UNİYORK)
- ACCV'09: *The 9<sup>th</sup> Asian Conference on Computer Vision*, Xi' an, China, September 2009.  
(Partners involved: UNİYORK)
- ICIP'09: *IEEE International Conference on Image Processing*, Cairo, Egypt, November 2009.  
(Partners involved: UNİYORK)
- IDMAP: *Workshop on Intelligent Data Analysis in Biomedicine and Pharmacology*  
(Partners involved: UNIVR)
- PMMIA 2009: *Probabilistic Models for Medical Image Analysis (a MICCAI Workshop)*, September 2009  
(Partners involved: UNIVR)

and have already submitted our work to the following conferences:

- NIPS'09: *The Neural Information Processing Systems Conference*, Vancouver, Canada, December 2009.  
(Partners involved: UNIVE, UNİYORK, UNIVR, ETH Zurich)
- CIARP 2009: *The 14th Iberoamerican Congress on Pattern Recognition*, Guadalajara, Jalisco, Mexico., November, 2009  
(Partners involved: TUD)

Further, we plan to submit our work to the following 2010 conferences:

- CVPR 2010: *The 22<sup>nd</sup> IEEE Conference on Computer Vision and Pattern Recognition*, San Francisco, CA, June 2010.  
(Partners involved: UNIVE, UNİYORK, UNIVR)

- ICPR 2010: *The 20<sup>th</sup> International Conference on Pattern Recognition*, Istanbul, Turkey, August 2010.  
(Partners involved: UNIVE, TUD, UNİYORK, IST, UNIVR)
- ECCV 2010: *The 11<sup>th</sup> European Conference on Computer Vision*, Crete, Greece, September 2010.  
(Partners involved: UNİYORK, UNIVR)
- NIPS 2010: *The Neural Information Processing Systems Conference*, December 2010.  
(Partners involved: UNIVE, UNİYORK, IST, ETH Zurich, UNIVR)
- ICML 2010: *The 27<sup>th</sup> international Conference on Machine Learning*, Haifa, Israel, 2010.  
(Partners involved: UNIVE, IST, ETH Zurich)
- MCS 2010: *Multiple Classifier Systems*, Cairo, April 2010  
(Partners involved: TUD)
- SSSPR 2010: *Joint IAPR International Workshops on Structural and Syntactic Pattern Recognition (SSPR 2010) and Statistical Techniques in Pattern Recognition (SPR 2010)*, Turkey, August 2010  
(Partners involved: TUD, UNİYORK, UNIVE)
- MICCAI 2010: *Medical Imaging Computing and Computer Assisted Intervention*, Beijing, China, 2010  
(Partners involved: ETH Zurich, UNIVR)

As for journal publications, by the end of the 22<sup>nd</sup> month of the project we plan to submit our results to:

- *IEEE Transactions on Pattern Analysis and Machine Intelligence*  
(Partners involved: IST, UNİYORK, UNIVE, UNIVR)
- *Journal of Machine Learning Research*  
(Partners involved: IST, UNIVE)
- *Neural Computation*  
(Partners involved: UNIVE)
- *Pattern Recognition*  
(Partners involved: TUD, UNİYORK, UNIVR, UNIVE)
- *Pattern Recognition Letters*  
(Partners involved: TUD)
- *IEEE Transactions on Systems, Man, and Cybernetics -- Part B*  
(Partners involved: UNİYORK)
- *IEEE Transactions on Biomedical Image Analysis*  
(Partners involved: UNIVR)

- *Linear Algebra and Its Applications*  
(Partners involved: UNİYORK)
- *Journal of Chemometrics*  
(Partners involved: TUD)

We foresee joint publication among the following partners: TUD+UNIVR, UNIVE+IST, UNIVE+UNİYORK, UNİYORK+IST, UNİYORK+TUD.

## **Workshops, special sessions, tutorials, special issues**

External dissemination will also take place via the organization of workshops, special sessions and tutorials at international conferences.

In particular, the following activities are planned.

1. We have submitted a proposal to run a workshop on “Learning in non-(geo)metric spaces” at *NIPS 2009* (Vancouver, December 2009) which is of the leading conferences in machine learning and pattern recognition. The workshop is planned to be a one-day meeting. The program will feature a panel discussion, invited oral presentations, a contributed poster session, and poster spotlights. We would like to obtain PASCAL 2 sponsorship for the meeting.

We plan to make the videos of the lectures available on VideoLectures.

2. In August 2009 we will submit a proposal for a poster presentation at the Project Exhibition of ECML PKDD 2009, *The European Conference on Machine Learning and Principle and Practice of Knowledge Discovery in Databases*, which will be held in Bled, Slovenia, in September 2009. The goal of the exhibition is mainly networking and exchanging of the results/experience. The organizers of the conference would like to help in disseminating the results of different projects related to machine learning and knowledge discovery applications and technologies.
3. Partner UNIVE will be running a special session devoted to “Learning and Intelligent Optimization in Structured Domains (LIONS)” within LION 2010: *The 4<sup>th</sup> International Conference on Learning and Intelligent Optimization*, which will be held in Venice, in January 2010 (<http://www.intelligent-optimization.org/LION4>).
4. In February 2009 Partner UNIVE plans to submit a proposal for a tutorial on “Game-theoretic Models in Computer Vision and Pattern Recognition” (the topic of WP5) at

ICPR 2010, *The 20<sup>th</sup> International Conference on Pattern Recognition*, Istanbul, Turkey, August 2010.

Finally, we mention that partners UNİYORK and UNIVE are co-editing a special issue of *Computer Vision and Image Understanding* on “Graph-based representations” (deadline for submission: October 2009).

## **Interactions with other initiatives**

We aim at establishing a close collaboration with external initiatives within the machine learning and pattern recognition communities, with a view to foster cross-fertilization of ideas by actively exchanging achievements and competences.

In particular:

1. The coordinating partner (UNIVE) plans to respond to the “Call for additional partners of PASCAL 2 Network of Excellence” (deadline: August 14, 2009).
2. We plan to request PASCAL 2 sponsorship for the proposed NIPS workshop (see Section 4)
3. The project’s activities will be advertised via the IAPR TC-15 Technical Committee, dedicated to promote the research on Graph-based Representations in the Pattern Recognition field (<http://www.greyc.ensicaen.fr/iapr-tc15/links.html>),
4. We also plan to advertise the project’s activities via the IAPR TC-2 Technical Committee dedicated to promote research on Structural and Syntactical Pattern Recognition (<http://www.rvg.ua.es/tc2/>).

We are also considering the idea of setting up, in a later stage of the project, a “SIMBAD challenge,” in collaboration with PASCAL 2 (details will possibly be described in the D8.2).

## **Web Site**

We have established a dedicated web site (<http://simbad-fp7.eu>), where the project publications, reports, software, data, and results, both theoretical and technological, will be published and illustrated.



An important part of the site will be the collection of similarity data that will be publicly available, thereby creating a common reference for similarity-based computational models and algorithms. An initial collection of datasets and tools that may be used by people outside the SIMBAD consortium will be provided by partner TUD and will be published in the website in October 2009.

Information about project development will be duly updated by each consortium member, who could directly upload the information to be disseminated, thanks to his own restricted web site area.

All publications produced within SIMBAD will be publicly available via the project website.

## Leaflet

We have produced a thousand copies of the following leaflet, summarizing the general objectives of the project, which will be distributed at appropriate conferences, summer schools, etc. A later revision is foreseen with more details about the work performed. This will be described in Deliverable D8.5 ("Second dissemination plan").



The leaflet is a light blue document with a white background for the text. It features several logos and flags, including the Italian flag, the UK flag, the Dutch flag, the Portuguese flag, the Swiss flag, and the European Union flag. The text is organized into columns and sections, providing project details, partner information, and project goals. At the bottom left, there is a small landscape photograph.

Starting date: 01 April 2008  
Duration: 36 months

<http://simbad-fp7.eu>  
contact: [info@simbad-fp7.eu](mailto:info@simbad-fp7.eu)

  
Università Ca' Foscari Venezia - Italy  
<http://www.dsi.unive.it>  
Marcello Pelillo - [pellillo@dsi.unive.it](mailto:pellillo@dsi.unive.it)

  
University of York - United Kingdom  
<http://www.york.ac.uk>  
Edwin Robert Hancock - [erh@cs.york.ac.uk](mailto:erh@cs.york.ac.uk)

  
Technische Universiteit Delft - The Netherlands  
<http://fict.ewi.tudelft.nl/>  
Robert Duin - [r.duin@ieee.org](mailto:r.duin@ieee.org)

  
Instituto Superior Técnico - Portugal  
<http://www.is.utl.pt>  
Mário Figueiredo - [mario.figueiredo@lx.it.pt](mailto:mario.figueiredo@lx.it.pt)

  
Università degli Studi Di Verona - Italy  
<http://vips.sci.univr.it>  
Vittorio Murino - [vittorio.murino@univr.it](mailto:vittorio.murino@univr.it)

  
Eidgenössische Technische Hochschule Zürich - Switzerland  
<http://www.ethz.ch>  
Joachim Buhmann - [jbuhmann@inf.ethz.ch](mailto:jbuhmann@inf.ethz.ch)

COMMISSION OF THE EUROPEAN COMMUNITIES  
  
INFORMATION SOCIETY AND MEDIA DIRECTORATE-GENERAL  
Seventh Framework Programme  
Information and Communication Technologies  
Collaborative Project  
FET Open



Beyond Features  
Similarity-based pattern analysis  
and recognition

### BEYOND FEATURES

Traditional pattern recognition techniques are centered on the notion of "feature". According to this view, each object is described in terms of a vector of numerical attributes and is therefore mapped to a point in a Euclidean (geometric) vector space so that the distances between the points reflect the observed (dis)similarities between the respective objects.

Despite its potential, the geometric approach suffers from a major intrinsic limitation, which concerns the representational power of feature-based descriptions. In fact, there are numerous application domains where either it is not possible to find satisfactory features or they are inefficient for learning purposes. Most commonly, this is typically the case when objects are described in terms of structural properties, such as parts and relations between parts, as is the case in shape recognition.

This project aims at bringing to full maturation a paradigm shift that is currently just emerging within the pattern recognition and machine learning domains, where researchers are becoming increasingly aware of the importance of similarity information *per se*, as opposed to the classical feature-based approach. Indeed, the notion of similarity (which appears under different names such as proximity, resemblance, and psychological distance) has long been recognized to lie at the very heart of human cognitive processes and can be considered as a connection between perception and higher-level knowledge, a crucial factor in the process of human recognition and categorization.

### WHY SIMBAD?

By departing from vector-space representations one is confronted with the challenging problem of dealing with (dis)similarities that do not necessarily possess the Euclidean behavior or not even obey the requirements of a metric. The lack of the Euclidean and/or metric properties undermines the very foundations of traditional pattern recognition theories and algorithms, and poses totally new theoretical/computational questions and challenges.

We aim at undertaking a thorough study of several aspects of similarity-based pattern analysis and recognition methods, from the theoretical, computational, and applicative perspective, with a view to substantially advance the state of the art in the field, and contribute towards the long-term goal of organizing this emerging field into a more coherent whole.

The whole project will revolve around two main themes, which basically correspond to the two fundamental questions that arise when abandoning the realm of feature-based representations:

1. How can one obtain suitable similarity information from object representations that are more powerful than, or simply different from, the vectorial?
2. How can one use similarity information in order to perform learning and classification tasks?

According to this perspective, the very notion of similarity becomes the pivot of non-vectorial pattern recognition in the same way as the notion of feature-vector plays the role of the pivot in the classical (geometric) paradigm.

### SIMBAD AT A GLANCE

From a methodological perspective, SIMBAD will be structured around the following strands:

**Deriving similarities for non-vectorial data (structural kernels)**, to develop computational models for generating similarities for non-vectorial data, with particular emphasis on structured and semi-structured descriptions

**Foundations of non-(geo)metric similarities**, to study both the causes of the lack of (geo)metricity in the similarity data and its effects over traditional machine learning algorithms

**Imposing geometricity on non-geometric similarities (embedding)**, to develop algorithms for transforming the original similarity data into proper vectorial representations suitable for traditional learning algorithms

**Learning with non-(geo)metric similarities**, to develop novel, general learning models which do not require the (geo)metric assumption

An important part of SIMBAD will concern the validation of the developed techniques, focusing mainly on biomedical problems.

**Analysis of tissue micro-array (TMA) images of renal cell carcinoma**, to validate the techniques developed by applying them to the analysis of Tissue Micro Array (TMA) images of renal cell carcinoma

**Analysis of brain magnetic resonance (MR) scans for the diagnosis of mental illness**, to validate the techniques developed by applying them to the analysis of brain MR scans in the context of mental health research

