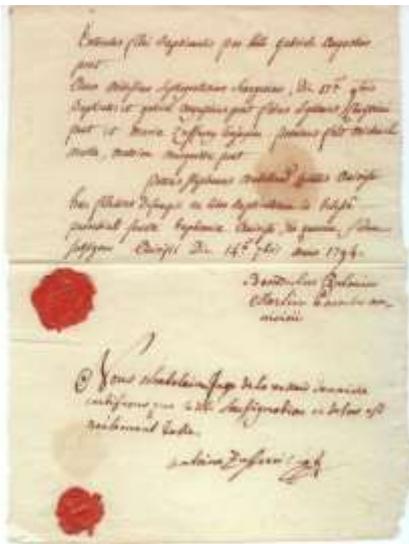


Valais*Wallis Digital

An opportunity for future research in
social media and digital humanities

Andrei Popescu-Belis (with François Foglia)
Idiap Research Institute

February 27, 2014 – Workshop on Historical Documents, U. of Fribourg



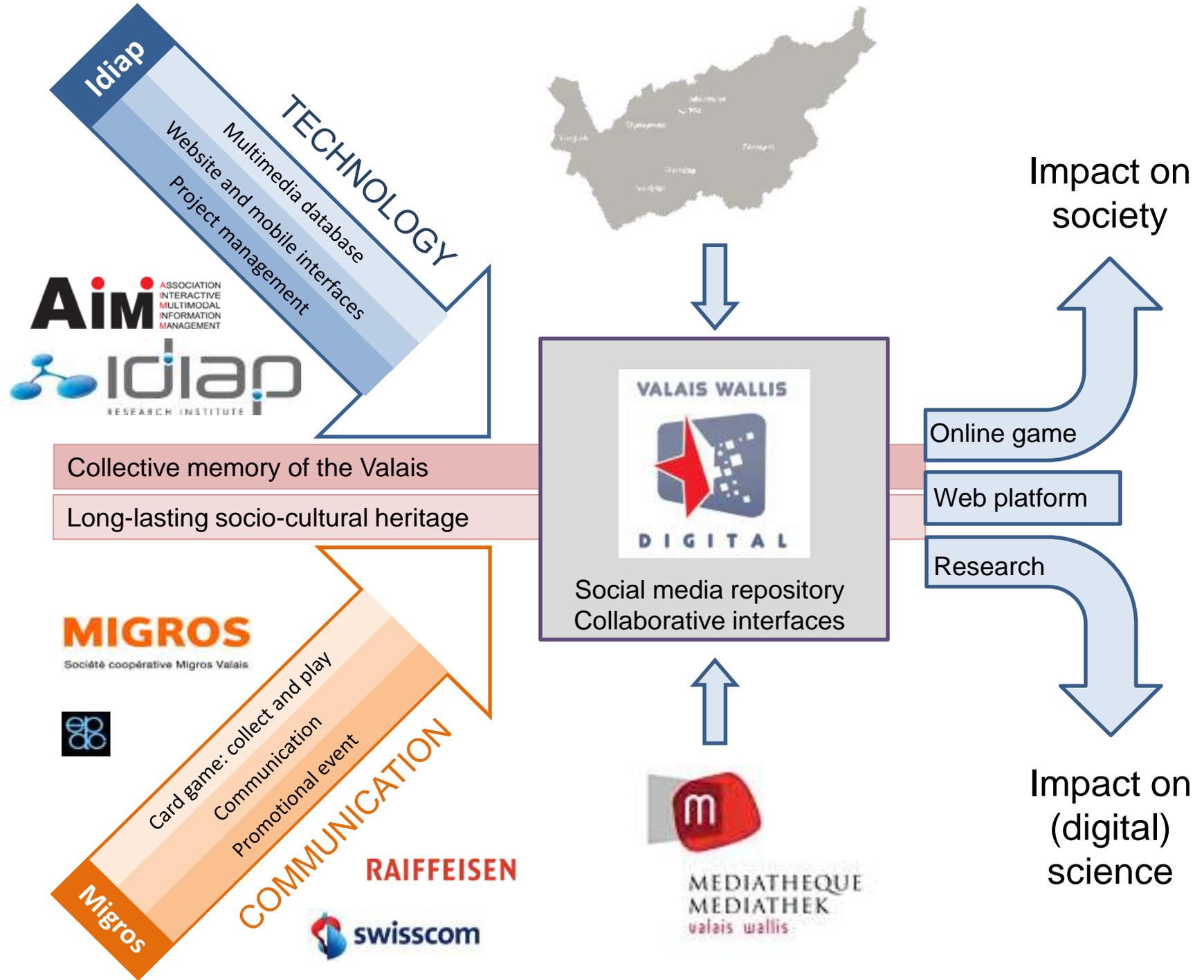
From grandpa's photo
album to resources for (digital)
historians and Wikipedians

Valais*Wallis Digital

<http://www.valais-digital.ch/>

2015* <http://valais2015.ch/>
VALAIS * WALLIS

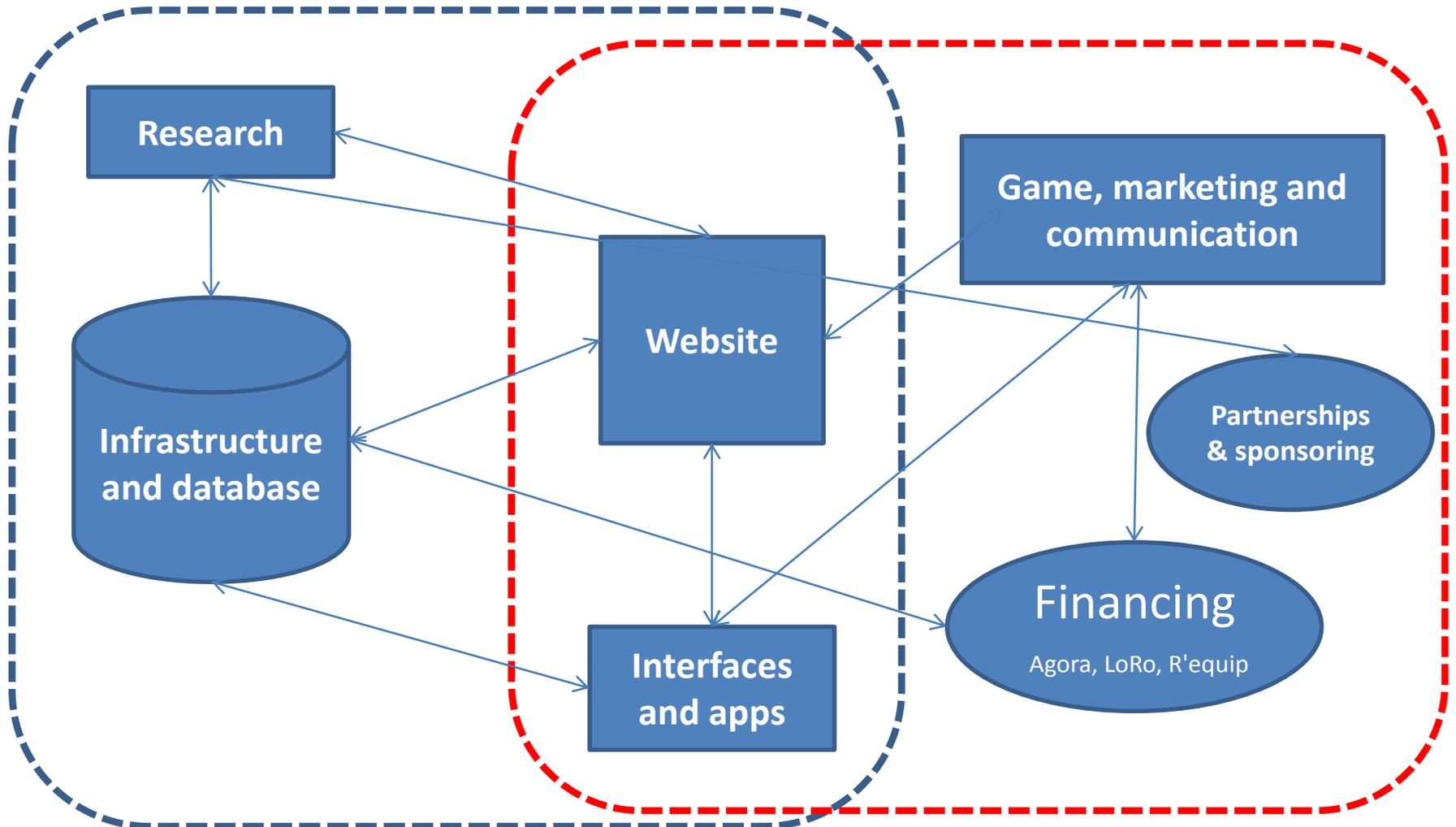
- Celebration of the bicentenary of the State of Valais joining the Swiss Confederation (August 7, 1815)
- Concept
 - bring back to life Valaisan memories from the past 200 years
 - card game (timeline) + social media platform
 - ➔ engage the general public and transfer private documents to an open multimedia repository
- Consortium led by Idiap, with Migros, Swisscom, an art school, a media library, the canton's archives, etc.
 - Head: [Dr. François Foglia](#), deputy director of Idiap



Components

Technology

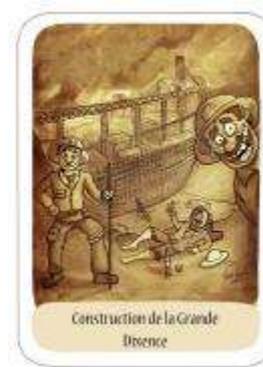
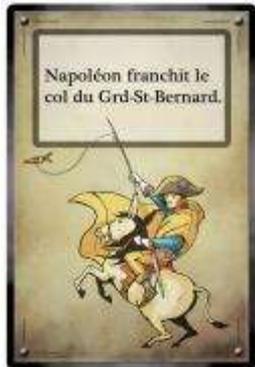
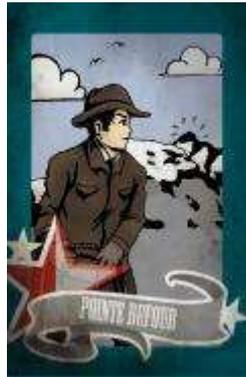
Front-end and financing



Engaging the public

- Valais*Wallis Digital will complement existing media
 - with a well-delimited historical and geographical target
 - e.g. the Canton's archives, Médiathèque, WikiValais, Wikipedia, *Notre Histoire*, *Dictionnaire historique de la Suisse*, and others
- Bootstrapping public interest
 - card game distributed by Migros: place cards on a timeline
 - one card per event | one event per year | 200 cards
- *Physical/digital world interface*
 - scan cards (e.g. with smartphone) to access corresponding events on the social media website
 - peruse or augment the multimedia information
 - digitize legacy supports (paper, film) through Swisscom shops
 - the website is periodically used to generate new cards

Illustration examples - EPAC



Card game: Timeline

- Migros distributing (200) board game cards (“Timeline” game)



- Cards can be scanned to have access to a web site where more information is available (text & multimedia documents, pictures, etc)
- Once on the web site, people are invited to upload additional documents and historical dates: *anybody can scan documents with the free assistance of Swisscom shops*
- Once a year, Migros uses the web site to enrich the board game and keep people motivated.

Core technology and timetable

- Preparation: **May 2013-December 2014**
 - multimedia database
 - relation to existing repositories and standards
 - website and mobile interface
 - relation to physical game
 - procedures for **digitizing legacy content**, licensing
 - game (definition, design, manufacturing)
- Marketing and communication
- Public launch: **January 2015**
 - start of media coverage and card distribution
 - Valaisans play the physical/digital game and upload media to www.valais-digital.ch | www.wallis-digital.ch | www.valais-wallis-digital.ch
- Maintained and enriched over the next 20 years

Research opportunities

- Relevant research fields
 - multimodal content analysis of “historical documents”
 - audio, images, video, text
 - social media and digital humanities
 - physical/digital world interplay
 - multimedia content management & recommendation
 - collaborative interfaces
- Large scale data and user pool
- Follow-up from IM2 work, under AIM2

Specific research needs

- How to federate existing and new resources?
 - transparent reuse and integration of content from existing social media
- How to monitor a social media database?
 - detect redundancy, conflicts, gaps, unauthorized content
- How to enrich a repository of media and user-generated text with suggested multimedia documents?
 - assist users for informative or illustrative purposes
- How to transform a social media repository into a valuable library of sources for historians, genealogists, or anthropologists?
 - large-scale experiments in digital humanities



Joining Valais Digital

Possible partnerships or research projects

Contact F. Foglia or A. Popescu-Belis





Natural Language Processing Group Overview

Head: Andrei Popescu-Belis

Objectives of the NLP group

- Semantic and discourse processing

APPLIED TO

- Statistical machine translation

AND TO

- Information indexing and retrieval,
multimedia recommendation

Research themes and projects (1)

- How can semantic and discourse processing improve statistical machine translation?
 - Projects
 - COMTIS (SNSF Sinergia, 2010-2013): Improving the coherence of machine translation output by modeling intersentential relations
 - MODERN (SNSF Sinergia, 2013-2016): Modeling discourse entities and relations for coherent machine translation
 - People and topics
 - Th. Meyer: disambiguate discourse connectives, pass info on to MT
 - S. Loaiciga (with ThM): predict verb tense translation prior to MT
 - N. Hajlaoui: evaluation with an automatic metric
 - Collaborations
 - U. of Geneva, U. of Zurich, U. of Utrecht, U. of Edinburgh

Research themes and projects (2)

- How can semantic indexing improve information retrieval and recommendation?
 - Projects
 - IM2 (SNSF NCCR, 2002-13, Prof. Bourlard): information interfaces
 - inEvent (EU, 2011-14, by Prof. Bourlard): networked multimedia events
 - AROLES (SNSF, 2012-2014): recommendation of lectures and snippets
 - People and topics
 - M. Yazdani (IM2): learning semantic similarity in networked data
 - C. Bhatt (AROLEs): multimedia segmentation, recommendation, and summarization, participation to MediaEval
 - N. Pappas (inEvent): user preference modeling for multimedia recommendation (e.g. of TED lectures)
 - M. Habibi (IM2): modeling topical diversity for real-time speech-based document retrieval
 - Collaborations
 - U. of Fribourg, U. of Edinburgh, IBM Haifa, HHI Berlin

MEDIEVAL HISTORY: AN OLD HOBBY

Past research on the medieval history of the Sinai peninsula (1998-2006)

- DEA from U. of Paris IV – Sorbonne in 2000
- Articles: editing, translating and criticizing medieval documents from St-Catherine's library about its history
 - Mouton J.-M. & Popescu-Belis A., “Une description du monastère Sainte-Catherine du Sinaï au XIIe siècle : le manuscrit d'Abû I-Makârim”. *Arabica*, vol. 53, n. 1, p. 1-53, 2006.
 - Mouton J.-M. & Popescu-Belis A., “La fondation du monastère Sainte-Catherine de Sinaï selon deux documents de sa bibliothèque : codex Arabe 692 et rouleau Arabe 955”. *Collectanea Christiana Orientalia (CCO)*, vol. 2, p. 141-205, 2005.
 - and 3 other synthesis papers



THANK YOU FOR YOUR ATTENTION!